# ACOUSTIC ECHO CANCELLATION WITH POST-FILTERING IN SUBBAND

*Xiaojian Lu and Benoît Champagne*

Department of Electrical & Computer Engineering, McGill University
3480 University Street, Montreal, Quebec, H3A 2A7, Canada
{xlu, champagne}@tsp.ece.mcgill.ca

**ABSTRACT**

In this paper, a new structure of acoustic echo cancellation (AEC) system exploring the post-filtering technique in subband is proposed. The post-filtering approach used here seeks to minimize near-end speech distortion while further suppressing the residual echo to a predefined level. The new AEC system under study seamlessly combines this post-filtering technique and the adaptive filters in a computationally efficient subband structure based on oversampled uniform DFT banks. Compared to its full-band counterpart, the proposed subband algorithm has a much lower computational complexity while achieving a comparable performance in terms of acoustic echo suppression. These observations are supported by experimental results described in the paper. An improved practical design of the prototype filter for use with the DFT filter banks is also introduced.

## 1. INTRODUCTION

In practice, echo suppression and computational complexity are two fundamental aspects of an acoustic echo cancellation (AEC) system. Ideally the acoustic echo can be entirely removed if the unknown echo path can be perfectly modelled by an adaptive filter. However, the performance of conventional AEC systems that mainly employ a linear adaptive filter is limited by numerous factors, including: background noise, quantization error, nonlinearity of the audio devices, and channel nonlinearity. The latter is especially serious in centralized AEC applications when vocoders are present along the echo path in the digital networks.

The post-filtering technique, which consists in applying various gains in the transform domain, e.g. frequency domain, has been widely used in speech enhancement [1]. Recently, this technique was also employed in a combined AEC-noise reduction approach [2]. Furthermore, it was explored to aggressively attenuate the notable residual echo resulting from the channel nonlinearity to a satisfactory level in a nonlinear channel [3].

The use of the post-filter can remarkably improve the performance of the AEC system in terms of high echo suppression. However, the computational complexity is also increased. For AEC applications requiring very long filters, it is thus necessary to reduce the computational load so as to be amenable to low-cost real-time implementation. To this end, the subband adaptive filtering approach in AEC has received considerable attention in the recent years, e.g. [4]. Compared with the full-band case, subband adaptive filtering achieves a gain in computational complexity as a result of a sampling rate reduction of the subband signals.

In this work, the post-filtering technique is explored and seamlessly combined with the adaptive filtering algorithm in a subband structure based on oversampled uniform DFT filter banks. An improved practical method for designing the prototype filter in the filter banks is also introduced. In addition to a very low computational complexity, experimental results show that the acoustic echo is significantly suppressed by the proposed AEC system based on combined subband adaptive filtering and post-filtering.

## 2. POST-FILTERING TECHNIQUE IN AEC

Consider the situation depicted in Fig. 1, where a linear filter (estimator) with impulse response $h(n)$ is used to remove the unwanted component $\delta(n)$ from a contaminated signal $e(n)$, thereby producing a signal estimate $\hat{v}(n)$ at its output. In the present AEC context, the near-end speech $v(n)$ needs to be extracted from the mixed signal $e(n)$ which may consist of the near-end speech and the residual echo in AEC. Because this estimator, which is usually realized by a filter, is placed after the conventional acoustic echo canceller, it is also called post-filter in this context.

Suppose the signal picked up by the microphone of a hands-free telephone is written as

$$d(n) = v(n) + y(n) \tag{1}$$

where $v(n)$ and $y(n)$ denote the near-end speech and the acoustic echo, respectively. The estimated echo produced by a conventional AEC algorithm (e.g. NLMS), denoted by $\hat{y}(n)$, is then subtracted from $d(n)$, resulting in a residual signal $e(n)$

$$e(n) = v(n) + \delta(n) \tag{2}$$

where, $\delta(n) = y(n) - \hat{y}(n)$ is the residual echo.

The signal $e(n)$ can be expressed in the frequency domain:

$$E(\omega) = V(\omega) + \Delta(\omega) \tag{3}$$

where $E(\omega)$, $V(\omega)$ and $\Delta(\omega)$ represent the Fourier transforms of $e(n)$, $v(n)$ and $\delta(n)$, respectively.

Let $\hat{V}(\omega) = H(\omega)E(\omega)$ be an estimate of $V(\omega)$ where $H(\omega)$ is a real-valued frequency weighting function. The error signal

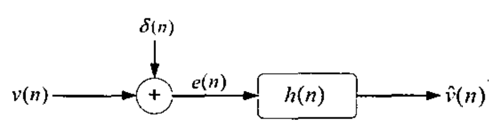


Figure 1: Diagram of the estimator.

associated to this estimator is given by $\epsilon(n) = \hat{v}(n) - v(n)$, or equivalently in the frequency domain:

$$\begin{aligned} \varepsilon(\omega) &= \hat{V}(\omega) - V(\omega) \\ &= [H(\omega) - 1]V(\omega) + H(\omega)\Delta(\omega) \end{aligned} \tag{4}$$

Assuming $v(n)$ and $\delta(n)$ are uncorrelated, and letting $S_{\varepsilon\varepsilon}(\omega)$ be the power spectral density (PSD) of the error signal $\epsilon(n)$, it can be verified that

$$S_{\varepsilon\varepsilon}(\omega) = \underbrace{[H(\omega) - 1]^2 S_{vv}(\omega)}_{J_v(\omega)} + \underbrace{H(\omega)^2 S_{\delta\delta}(\omega)}_{J_\delta(\omega)} \tag{5}$$

where $S_{vv}(\omega)$ and $S_{\delta\delta}(\omega)$ denote the PSDs of $v(n)$ and $\delta(n)$, respectively. Note that the first term $J_v(\omega)$ describes the distortion of the near-end speech $v(n)$, while the second term $J_\delta(\omega)$ represents the attenuation of the residual echo $\delta(n)$. Using a similar approach as in [5], and aiming at minimizing the near-end speech distortion while suppressing the residual echo to a predefined level, $H(\omega)$ can be obtained as the solution to

$$\min_{H(\omega)} J_v(\omega) \quad \text{subject to} : J_\delta(\omega) \leq \beta S_{\delta\delta}(\omega) \tag{6}$$

where $\beta \in [0, 1]$ describes the amount of the echo attenuation.

The optimal estimator in the sense of (6) can be found by applying Karush-Kuhn-Tucker necessary conditions for inequality constraints [6]. A practical solution then follows by making suitable approximations as described in [3], resulting in

$$H(\omega) \approx \frac{S_{ed}(\omega)}{S_{ed}(\omega) + \alpha S_{\hat{y}\hat{y}}(\omega)}, \tag{7}$$

where $S_{\hat{y}\hat{y}}(\omega)$ is the PSD of $\hat{y}(n)$ and $S_{ed}(\omega)$ is the cross spectrum density of $e(n)$ and $d(n)$. The attenuation factor $\alpha$ can be adjusted to achieve the proper trade-off between the echo attenuation and the distortion of the near-end speech, as needed in various situations.

## 3. A COMBINED ALGORITHM IN SUBBAND

In a typical subband filtering scheme for AEC, signals are partitioned into subbands; adaptive filters are then applied in each subband at a decimated rate. Since the microphone signal may contain a near-end speech component that will go through the cascade of an analysis and a synthesis bank prior to its transmission, a perfect-reconstruction (PR) or near-PR property is needed for the filter banks so that the near-end speech will not be distorted. Furthermore, due to aliasing effects in subband, critically sampled filter banks, which have received considerable attention in the literature, can not be applied directly in a subband system for AEC [7]. Accordingly, an oversampling scheme is often used in subband to keep aliasing distortion in the full-band output of the synthesis bank below an acceptable level [4, 8].

Among various subband structures, the uniform DFT filter banks [9] are motivated by providing DSP engineers with simple and concrete methods for the rapid design and prototyping of filter banks suitable for subband AEC applications. The details about the implementation of the simple uniform DFT filter banks in the AEC application can be found in [4].

### 3.1. A simple design of prototype filter

The design of the prototype filter plays a critical role in the implementation of DFT filter banks, since the properties of the prototype filter remarkably affect the performance of the filter banks, such as the reconstruction error and the aliasing in the subband. Instead of using complex computer-aided optimization techniques, e.g. [10], a simple way to design the prototype filter for use in oversampled uniform DFT filter banks is proposed in [4]. In this approach, the prototype filter $h(n)$ in a $K$-channel DFT structure is obtained by performing a $K/2$-point interpolation of a two-channel QMF filter $h_0(n)$. The simplicity of this design procedure brings great convenience in practical engineering application.

Here we propose an improved method to obtain a prototype filter by interpolating the QMF filter in the frequency domain. This method, simply referred to as the DFT method, explores the properties of the DFT to carry out the interpolation. Specifically, suppose that we have a QMF filter $h_0(n)$ with $L_0$ samples, and that we plan to obtain a prototype filter used in a subband structure with $K$ channels. First apply an $L_0$-point DFT to the coefficients of $h_0(n)$, resulting in a sequence in the frequency domain. Then expand this sequence to a length of $KL_0/2$ by padding $(K/2 - 1)L_0$ zeros. Finally apply $(KL_0/2)$-point inverse DFT to the new sequence, so that the prototype filter $h(n)$ is obtained.

It has been verified (results are not shown here due to the limited space) that using the DFT method can lead to a better prototype filter in terms of higher attenuation (more than 10dB) in the stopband, resulting in turns in a smaller reconstructed error and less aliasing in the subbands.

### 3.2. Proposed subband combination algorithm

As pointed out, the post-filter can significantly suppress the residual echo resulting from an insufficient attenuation of a conventional AEC system. However, the main drawback of this system is its high computational complexity, which is caused not only by the considerable computation requirement of the full-band adaptive filtering algorithm itself, but also by the use of the DFT for transforming the signals back and forth between the time domain and the frequency domain, since the post-filter is usually realized in the frequency domain while most popular AEC algorithms are implemented in the time domain.

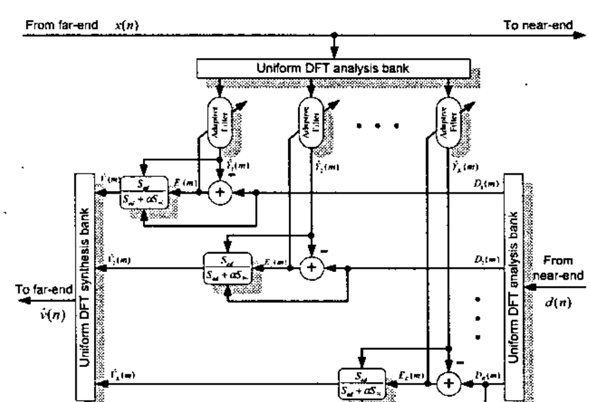


Figure 2: The diagram of the combined AEC in subband.

In order to reduce the computational complexity in AEC applications, where the power of the processor is usually limited, the adaptive filtering algorithm can be implemented in a subband structure. Furthermore, we note that when the signals are mapped into the subbands, the decimated outputs of the subband analysis filters actually provide a frequency domain representation of their corresponding input. Accordingly, the post-filter shown in (7) can be realized in the subband. We will show that the combination of the post-filter and the adaptive filter in the subband results in significant computational savings.

The diagram of the combined AEC in subband is depicted in Fig. 2, where both the far-end signal $x(n)$ and the near-end signal (i.e. the microphone signal) $d(n)$ are split into subbands by analysis banks. The adaptive filtering algorithm is applied in each subband at a decimated rate. The subband residual echo signals are further attenuated by individual (scalar) post-filters before they are recombined by a synthesis bank to compose a full-band output signal $\hat{v}(n)$ at the original rate. The uniform DFT filter banks have been employed in the combined AEC system since they provide a simple and concrete method for the rapid design and prototyping of filter banks suitable for subband AEC applications.

### 3.3. Example: NLMS combined with post-filter in subband

The normalized least-mean-squared (NLMS) algorithm is very popular in practical AEC applications due to its predictable robust behaviour, simple implementation and low computational requirement. In order to compensate the echo attenuation loss caused by the nonlinearities of the echo path and to reduce the computational complexity for the use of low-cost real-time DSP processors, we integrate NLMS and the post-filter in the subband structure. The algorithm is summarized in Algorithm 1, where all the variables are the components of signals in subband, and the adaptive filter in the $k$th bank is assumed to have a length of $L_k$, which may be different for each bank. The parameters $\mu$, $\gamma$ and $\alpha$ denote the step-size, the forgetting factor and the attenuation factor, respectively. A small constant $\rho$ is introduced to prevent the division from overflow. Note that other adaptive filtering algorithms, e.g. affine projection algorithm, may also be employed in the structure shown in Fig. 2. However, more computational capacity may be required.

**Algorithm 1** NLMS plus post-filtering in subband

**Initialization:** $\mathbf{W}_k(1) = \mathbf{0}$, $k = 1, 2, \ldots, K$, where

$\mathbf{W}_k(m) = [W_k^{(1)}(m), W_k^{(2)}(m), \ldots, W_k^{(L_k)}(m)]^T$

**Recursion:**

**for** $m = 1, 2, \ldots$ **do**

**for** $k = 1, 2, \ldots, K$ **do**

$\mathbf{X}_k(m) = [X_k(m), X_k(m-1), \ldots, X_k(m - L_k + 1)]^T$

$\hat{Y}_k(m) = \mathbf{W}_k^H(m)\mathbf{X}_k(m)$

$E_k(m) = D_k(m) - \hat{Y}_k(m)$

$$\mathbf{W}_k(m+1) = \mathbf{W}_k(m) + \frac{\mu}{\mathbf{X}_k^H(m)\mathbf{X}_k(m) + \rho} E_k^*(m)\mathbf{X}_k(m)$$

$S_{ed}^{(k)}(m) = \gamma S_{ed}^{(k)}(m-1) + (1-\gamma)E_k(m)D_k^*(m)$

$S_{\hat{y}\hat{y}}^{(k)}(m) = \gamma S_{\hat{y}\hat{y}}^{(k)}(m-1) + (1-\gamma)|\hat{Y}_k(m)|^2$

$$\hat{V}_k(m) = \frac{S_{ed}^{(k)}(m)E_k(m)}{S_{ed}^{(k)}(m) + \alpha S_{\hat{y}\hat{y}}^{(k)}(m)}$$

**end for**

**end for**

Note that the subband signals at the output of the uniform DFT analysis banks satisfy a conjugate symmetric property: the $i$th subband signal is the complex conjugate of the $(K + 2 - i)$th for $1 < i \leq K/2$ (assuming the number of the filter banks, $K$, is even), while the signals in the first and the $(K/2 + 1)$th banks are real. Thus, there are a total of $K/2 + 1$ independent subband signals: two real and $K/2 - 1$ complex, which only need $K/2 + 1$ adaptive filters running in these subbands. Furthermore, FFT is used to perform DFT, resulting in more computational savings.

In this work, we measure the computational complexity in terms of *operations*, where one operation is defined as one real multiplication plus one real addition. Accordingly, 4 operations are basically required to realize one complex multiplication and one complex addition. The total computational complexity of the proposed AEC system is made up of three distinct components, namely, analysis/synthesis filtering, adaptive filtering and post-filtering in each subband.

Suppose that the decimation factor is $M$ ($M < K$) and that the number of taps of all subband filters is chosen equal to $L/M$, where $L$ is the length of the full-band adaptive filter, which is assumed to match the duration of the echo path. Then the operations per every $M$ input samples for the above three components are respectively $3K(L_0/2 + 2\log_2 K)$, $4(K-1)L/M$ and $6(K-1)$. Therefore, the computational complexity of the proposed algorithm (based on NLMS), in common units of operations per sample (OPS) is obtained

$$\frac{3K(L_0/2 + 2\log_2 K) + 4(K-1)L/M + 6(K-1)}{M} \tag{8}$$

We note that the new algorithm achieves remarkable computational savings, if compared to the full-band counterpart which usually needs $2L + O(L_p)$ OPS, where $L_p$ denotes the size of the DFTs, as required in the practical full-band realization of the post-filter (7). Furthermore, the seamlessly integrated post-filter in subband only requires slightly extra computational capacity of $6(K-1)/M$ OPS, compared to the pure subband NLMS.

## 4. EXPERIMENT RESULTS

Three AEC algorithms, namely, the subband NLMS, and NLMS plus post-filter, both in full-band and subband, were tested on a platform where a low-cost loudspeaker played in a high volume so that the entire echo path presented certain nonlinearity.

In the implementation of the subband adaptive filtering algorithms, the number of filter banks and the decimation factor were $K = 16$ and $M = 12$, respectively. The prototype filter was obtained by using the DFT method to interpolate the QMF filter 16A [9], which has $L_0 = 16$ parameters, by a factor $K/2 = 8$. Hence, the length of the prototype filter is 128. For simplicity, the adaptive filters in subband were set to the same length, i.e. $L/M = 300/12 = 25$, where $L = 300$ was the length of the adaptive filter in full-band. The parameters of $\alpha$ and $\gamma$ in Algorithm 1 were 5.0 and 0.8, respectively. The step-size $\mu$ was set to 0.9 for all the algorithms.

Under the above conditions, the computational complexity of the proposed algorithm, as shown in (8), can be calculated, resulting in about 200 OPS, that is: about 4% more than the pure subband NLMS, but only about 1/3 the complexity of the its full-band counterpart (NLMS with post-filtering).

The results shown in Fig. 3 reveal that the proposed algorithm outperforms the pure subband NLMS in terms of significant acoustic echo suppression, and especially when the nonlinearity of the echo path cannot be neglected. Furthermore, the proposed algorithm exhibits a level of echo attenuation that is comparable to its full-band counterpart. A little distortion of the near-end speech is introduced by the use of the post-filtering technique, both in the full-band and in the subband, when the double-talk occurs. However, the distortion is not significant, as illustrated in Fig. 4, and thus it is almost imperceptible for the far-end user when he/she is talking at the same time. Listening demos can be found on the web at http://www.tsp.ece.mcgill.ca/~xlu/waspaa03.htm.

As it is shown, the proposed algorithm significantly reduces the computational complexity, compared to its full-band counterpart. However, similarly to other subband adaptive filtering algorithms, some processing delay is introduced by the proposed scheme, which is 16ms in the above experimental implementation. We note that this amount of delay is acceptable in most AEC applications.

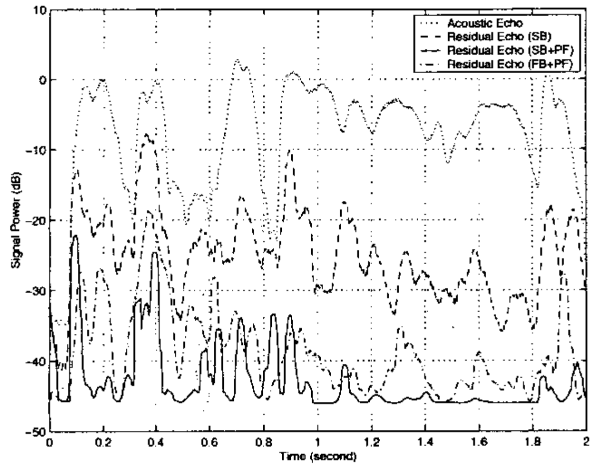


Figure 3: Performance comparison of AEC system (based on NLMS): NLMS in subband (SB), NLMS plus post-filter in full-band (FB+PF) and NLMS plus post-filter in subband (SB+PF).

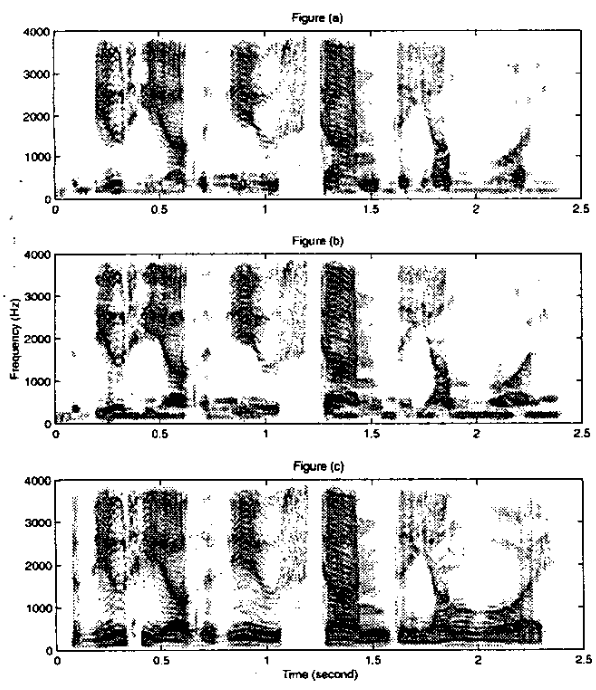


Figure 4: Spectrograms: (a) output of combined subband system; (b) output of combined full-band system; (c) original (clean) near-end speech (*The beauty of the view stunned the young boy*).

## 5. CONCLUSIONS

An efficient AEC system that combines the post-filtering technique with the conventional adaptive filtering algorithm in the subband structure is proposed. The acoustic echo is significantly suppressed by the new AEC system even if the nonlinearity of the echo path cannot be neglected. Furthermore, the seamless combination of the post-filter and the adaptive filter in the subband greatly reduces the computational complexity. Therefore, the proposed AEC system is very attractive for the practical use with real-time implementation.

## 6. REFERENCES

[1] D. O'Shaughnessy, *Speech Communications: Human and Machine*, IEEE Press, 2nd edition, 2000.

[2] R. Le Bouquin-Jeannès, P. Scalart, G. Faucon, and C. Beaugeant, "Combined noise and echo reduction in hands-free systems: a survey," *IEEE Trans. Speech and Audio Processing*, vol. 9, no. 8, pp. 808–820, Nov. 2001.

[3] X. Lu and B. Champagne, "A centralized acoustic echo canceller exploiting masking properties of the human ear," in *Proc. ICASSP'03*, Apr. 2003, vol. 5, pp. 377–380.

[4] Q.-G. Liu, B. Champagne, and K. C. Ho, "Simple design of oversampled uniform DFT filter banks with applications to subband acoustic echo cancellation," *Signal Processing*, vol. 80, pp. 831–847, Jun. 2000.

[5] Y. Ephraim and H. L. Van Trees, "A signal subspace approach for speech enhancement," *IEEE Trans. Speech and Audio Processing*, vol. 3, no. 4, pp. 251–266, Jul. 1995.

[6] S. G. Nash and A. Sofer, *Linear and Nonlinear Programming*, McGraw-Hill, New York, 1996.

[7] A. Gilloire and M. Vetterli, "Adaptive filtering in subbands with critical sampling: analysis, experiments, and application to acoustic echo cancellation," *IEEE Trans. Signal Processing*, vol. 40, pp. 1862–1875, Aug. 1992.

[8] Z. Cvetkovic and M. Vetterli, "Oversampled filter banks," *IEEE Trans. Signal Processing*, vol. 46, no. 5, pp. 1245–1255, May 1998.

[9] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*, Prentice Hall, Englewood Cliffs, N.J., 1983.

[10] C. D. Creusere and S. K. Mitra, "A simple method for designing high-quality prototype filters for m-band pseudo QMF banks," *IEEE Trans. Signal Processing*, vol. 43, no. 4, pp. 1005–1007, Apr. 1995.