

A MODIFIED OPEN-LOOP DELAYLESS SUBBAND ADAPTIVE ECHO CANCELLER

Peter Eneroth and Tomas Gänsler

Department of Applied Electronics, Lund University, Sweden
pe@tde.lth.se, tg@tde.lth.se

ABSTRACT

In the delayless subband adaptive echo canceller, the estimated subband impulse responses are non-causal and the non-causal parts are truncated. In the closed-loop structure, the feed-back of the error-signal compensate for the truncation, which is not the case in the open-loop structure. A modified open-loop structure is proposed, in which the truncation error is reduced. The performance of the proposed structure approaches the closed-loop structure, and has the advantage of a higher convergence rate.

1. INTRODUCTION

Due to long duration of acoustic impulse responses, an ordinary echo canceller utilizing the normalized least mean square (NLMS) algorithm for adaptation, needs to estimate an FIR filter with a large number of filter taps. This results in a high computational complexity. Furthermore, the convergence rate of the NLMS algorithm is slow due to the large eigenvalue spread of the speech signal correlation matrix.

To reduce these drawbacks, the input signal, $x(n)$, and the return signal, $y(n)$, possible with the addition of double-talk, may be decomposed into M down-sampled narrow-band signals $x_m(k)$ and $y_m(k)$, where m denotes the index of the subband, Fig. 1. Since adaptation in the subbands are performed at a lower rate and the subband impulse responses are shorter, the computational complexity is reduced. Convergence rate is improved because the eigenvalue spread in the *effective* frequency range of the subbands is reduced [1], where "effective" is the frequency range used in the estimation of the impulse response.

In a perfect reconstruction subband algorithm, the echo cancelling is performed in subbands and the fullband residual echo signal is reconstructed from the subband residual echo signals. A disadvantage with this structure is the inherent delay in the signal path. A delayless subband adaptive filter architecture was introduced in [2]. In this structure the impulse response is estimated in the subbands, but the sub-

band estimates are then transformed into a fullband impulse response. Echo cancellation is performed in the time domain, without introducing any delay in the signal path.

The delayless subband echo canceller may be configured in two ways, open or closed loop. In the open-loop configuration, the NLMS subband algorithm have the far end signal, $x(n)$, and the echo signal, $y(n)$, as input and error signals are generated internally in the subband NLMS algorithms. In the closed-loop configuration the fullband error signal, $e(n)$, is decomposed into subbands and used as an input signal instead of the echo signal, as in the open-loop configuration. The closed loop variant has a smaller estimation error after convergence, at the cost of slower convergence.

In this paper, the subband impulse responses are shown to be "non-causal", even though the fullband impulse response is causal. This is discussed in [3]. In the closed-loop configuration, the feed-back of the error signal will compensate for the truncation of the non-causal part. In the open-loop configuration, the truncation will increase the estimation error. Because of the faster convergence rate, the open-loop configuration is to be preferred. Also presented is, a modified open-loop structure, that reduces the misalignment of the estimated subband impulse response.

2. SYSTEM DESCRIPTION

The echo canceller is illustrated in Fig. 1. The far end signal, $x(n)$, is returned as an echo, $y(n)$, because of the acoustic coupling between the loudspeaker and the microphone. The echo canceller adaptively estimates the echo path impulse response, $\hat{h}(n)$, and cancels the echo, $e(n) = y(n) - \hat{h}(n) * x(n)$, where $*$ denotes the convolution operator. Due to impulse response estimation error, the residual echo signal, $e(n)$, is non-zero.

The echo path transfer function is estimated in subbands, depicted in the bottom half of Fig. 1. In an open-loop configuration, a NLMS algorithm estimates the subband impulse response $\hat{h}_m(k)$ using the decomposed and down-sampled far end signal, $x_m(k)$,

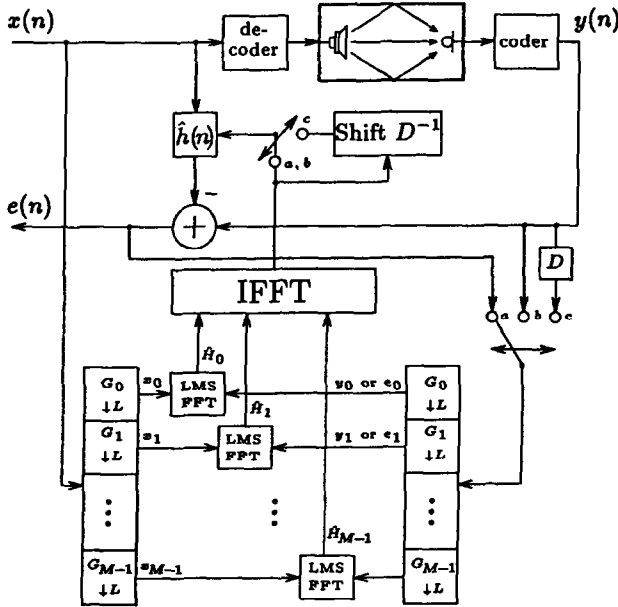


Figure 1: Delayless subband echo canceller and model of acoustic echo path. The echo canceller have three configuration possibilities. When the two switches are in position a, the echo canceller operates in closed-loop mode. In position b or c it operates in open-loop or modified open-loop, respectively.

and the corresponding echo signal, $y_m(k)$. The error signal, $\hat{e}_m(k)$, needed in the NLMS algorithm is generated internally, $\hat{e}_m(k) = y_m(k) - \hat{h}_m(k) * x_m(k)$. In a closed-loop configuration, the decomposed fullband error signal, $e_m(k)$, is used instead of the echo signal $y_m(k)$. The following steps are performed on segment basis, with a typical segment size $Seg = 256$, which reduces the calculation complexity. The subband transfer function is estimated as the fast Fourier transform (FFT) of the subband impulse response. Finally the estimated fullband impulse response is found as the inverse FFT of the frequency-stacked subband impulse responses. Echo cancellation is then performed in the time domain, without introducing a delay to the return signal transmission path. A detailed description and analysis can be found in [2].

2.1 Subband impulse response analysis

The subband filter bank can be realized with modulators and low-pass filters, $\tilde{x}_m(n) = (x(n)e^{-j2\pi nm/M}) * h_{LP}(n)$, where h_{LP} is a low pass prototype filter for filterbank G . $\tilde{x}_m(n)$ is then down-sampled as $x_m(k) = \tilde{x}_m(Lk)$, k integer. $y_m(k)$ and $e_m(k)$ are defined similarly. In order to reduce downsampling aliasing, non-critical downsampling is used, typically $L = M/2$. In other words, there is a 50% overlap between two adjacent subband transfer function estimates.

In order to calculate the optimal open loop subband impulse response in the least square error sense, we start without any downsampling, that is $L = 1$. Let

$$\mathbf{x}(i) = [x(i) \ x(i+1) \ \dots \ x(N-K+i)]^T,$$

$$\mathbf{X} = [\mathbf{x}(K) \ \mathbf{x}(K+1) \ \dots \ \mathbf{x}(N)]^T,$$

$$\mathbf{y} = [y(K) \ y(K+1) \ \dots \ y(N)]^T,$$

$$\mathbf{e} = [e_m(K) \ e_m(K+1) \ \dots \ e_m(N)]^T,$$

$$\hat{\mathbf{h}} = [\hat{h}_0 \ \hat{h}_1 \ \dots \ \hat{h}_{K-1}]^T,$$

$$\mathbf{D} = \begin{bmatrix} \mathbf{h}_{LP}^T & \mathbf{0}^T & \dots & \mathbf{0}^T \\ \mathbf{0}^T & \mathbf{h}_{LP}^T & \dots & \mathbf{0}^T \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}^T & \mathbf{0}^T & \dots & \mathbf{h}_{LP}^T \end{bmatrix}$$

where N denotes the number of samples, K the length of the estimated impulse response $\hat{\mathbf{h}}$, \mathbf{h}_{LP} the impulse response of the low pass filter of length P and \mathbf{D} a lowpass matrix of size $(N-K-P+2) \times (N-M+1)$. The subband residual echo can then be expressed as,

$$\mathbf{e} = \mathbf{D}\mathbf{y} + \mathbf{v}_y - (\mathbf{D}\mathbf{X} + \mathbf{V}_s)\hat{\mathbf{h}},$$

where \mathbf{v}_y and \mathbf{V}_s denotes the quantization error vector and matrix, respectively. The optimal subband impulse response is determined by minimizing the cost function $\epsilon = \mathbf{e}^H \mathbf{e}$, where H denotes the hermitian transpose. The optimal subband impulse response in the least square sense may then be written as,

$$\hat{\mathbf{h}} = (\mathbf{X}^H \mathbf{D}^2 \mathbf{X} + \mathbf{V}_s^H \mathbf{V}_s + \mathbf{X}^H \mathbf{D}^H \mathbf{V}_s + \mathbf{V}_s^H \mathbf{D} \mathbf{X})^{-1} \cdot (\mathbf{X}^H \mathbf{D}^2 \mathbf{y} + \mathbf{X}^H \mathbf{D}^H \mathbf{v}_y + \mathbf{V}_s^H \mathbf{D} \mathbf{y} + \mathbf{V}_s^H \mathbf{v}_y).$$

By using the independence theory, the expected subband impulse response can be reduced to,

$$\begin{aligned} E[\hat{\mathbf{h}}] &= (E[\mathbf{X}^H \mathbf{D}^2 \mathbf{X}] + E[\mathbf{V}_s^H \mathbf{V}_s])^{-1} E[\mathbf{X}^H \mathbf{D}^2 \mathbf{y}] \\ &= const \cdot (\mathbf{R}_{\mathbf{D}\mathbf{X}} + \sigma_v^2 \mathbf{I})^{-1} \mathbf{r}_{\mathbf{D}\mathbf{y}, \mathbf{D}\mathbf{X}}, \end{aligned} \quad (1)$$

where the quantization error is modelled as white noise with variance σ_v^2 , $\mathbf{R}_{\mathbf{D}\mathbf{X}}$ denotes the correlation matrix of the subband signal x_m and the cross-correlation vector between x_m and y_m is defined as $\mathbf{r}_{\mathbf{D}\mathbf{y}, \mathbf{D}\mathbf{X}} = [r_{y_m x_m}(0) \ \dots \ r_{y_m x_m}(K-1)]^T$. The cross-correlation between the fullband signals y and x , $r_{yx}(l) = E[y(n)x^*(n-l)]$, is zero when $l < 0$, since the fullband impulse response is casual. The subband crosscorrelation may be expressed as

$$r_{y_m, x_m}(l) = E \left[\sum_{a=0}^{P-1} h_{LP}(a) y(n-a) \sum_{b=0}^{P-1} h_{LP}^*(b) x^*(n-a-l) \right]$$

where P is the lowpass filter length. This can be simplified to,

$$\begin{aligned} r_{y_m, x_m}(l) &= \sum_{i=-(P-1)}^{P-1} r_{y_m}(i+l) \sum_{j=0}^{P-1-|i|} h_{LP}(j) h_{LP}^*(j+|i|), \\ &= \sum_{i=-(P-1)}^{P-1} r_{y_m}(i+l) r_{h_{LP}}(i), \end{aligned} \quad (2)$$

where $r_{h_{LP}}(i)$ denotes the autocorrelation of the low-pass filter h_{LP} . If we denote the length of the echo path flat delay by Δ , then $r_{yx}(l) = 0$ when $l < \Delta$. According to eq. 2, the crosscorrelation between the subband signals x_m and y_m has its first non-zero value at position $\Delta - P + 1$, due to the lowpass filter h_{LP} . If $\Delta - P + 1 < 0$ the subband impulse response is *non-causal*. As shown in eq. 1, the non-causal part of the subband impulse response is not taken into account when the estimated impulse response \hat{h} is calculated. Therefore, the performance of the echo canceller will decrease when the flat delay Δ is short compared to the lowpass filter length P .

Even if we consider subband impulse response estimation with non-critical down-sampling, $L < M$, eq. 1 is still relevant, we only need to down-sample \mathbf{R}_{DX} and $\mathbf{r}_{Dy,DX}$. The crosscorrelation may then be written as

$$r_{y_m, x_m}(l) = \sum_{i=-(P-1)}^{P-1} r_{y_m}(i+Ll)r_{h_{LP}}(i),$$

where L is the down-sampling factor. The subband impulse response is still non-causal when $\Delta - P + 1 < 0$.

In order to study the spectral behavior of the estimated impulse response, let us define a orthonormal matrix \mathbf{Q} as,

$$\mathbf{Q} = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & w^1 & \dots & w^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & w^{N-1} & \dots & w^{(N-1)^2} \end{bmatrix}, \quad w = e^{-j\frac{2\pi}{N}}$$

The subband autocorrelation \mathbf{R}_{DX} may then be approximately decomposed as $\mathbf{R}_{DX} \approx \mathbf{Q}\mathbf{\Lambda}_{DX}\mathbf{Q}^H$, giving

$$\mathbf{E}[\hat{\mathbf{h}}] = \mathbf{Q}^H (\mathbf{\Lambda}_{DX} + \sigma_v^2 \mathbf{I})^{-1} \mathbf{Q} \mathbf{r}_{Dy,DX}.$$

where $\mathbf{\Lambda}_{DX}$ is a diagonal matrix. Premultiply a vector with \mathbf{Q} is equivalent to perform the Discrete Fourier Transform and the expected subband transfer function is therefore,

$$\mathbf{E}[\hat{\mathbf{H}}] = \mathbf{E}[\mathbf{Q}\hat{\mathbf{h}}] = (\mathbf{\Lambda}_{DX} + \sigma_v^2 \mathbf{I})^{-1} \mathbf{Q} \mathbf{r}_{Dy,DX}.$$

$\mathbf{\Lambda}_{DX}$ may be identified as the power spectral density of subband input signal, $S_{x_m x_m}(k) = |H_{LP}(k)|^2 S_{xx}(k - mK/2)$, and the Discrete Fourier Transform of the crosscorrelation $\mathbf{r}_{Dy,DX}$ as $S_{y_m x_m}(k) = |H_{LP}(k)|^2 S_{yx}(k - mK/2)$. The least square solution to the expected transfer function is,

$$\mathbf{E}[\hat{H}(k)] = \frac{|H_{LP}(k)|^2 S_{yx}(k - mK/2)}{|H_{LP}(k)|^2 S_{xx}(k - mK/2) + \sigma_v^2}$$

In the lowpass region, $|H_{LP}(k)|^2 S_{xx}(k - mK/2)$ is the dominant term in the denominator and the estimate converges to the Wiener solution. In the highpass

region, the quantization and aliasing error dominate and the estimate converges to zero.

As an improvement of the open-loop architecture, we propose the *modified open-loop* architecture illustrated in Fig. 1, switch position c. A delay D is introduced to the signal $y(n)$ before the subband decomposition. The subband impulse responses are delayed, and errors due to truncation of the non-causal part is reduced. To compensate for the introduced delay, the estimated fullband impulse response is shifted. Notice that no delay is introduced to the fullband signal transmission path.

2.2 Closed-loop convergence rate

The closed-loop configuration almost eliminates the problem due to truncation of the non-causal part of the estimated subband impulse responses, because the fullband error signal is minimized. Convergence rate of the closed-loop is lower than the open-loop configuration. This can be seen in [4], since the closed-loop echo canceller has the same properties as the Block LMS structure analyzed in [4]. It is shown in [4] that to guarantee stability, the standard LMS (open-loop) step size parameter is bounded as $0 < \mu < \lambda_{max}^{-1}$, where λ_{max} denotes the largest eigenvalue of the far end signal correlation matrix. The corresponding bound for Block LMS (closed-loop) is $0 < \mu < (Seg\lambda_{max})^{-1}$, where Seg is the segment size introduced above. Calculation complexity is reduced when Seg is increased, since the fullband impulse response estimate is calculated once for every segment. Therefore low calculation complexity configurations needs to be open-loop.

3. RESULTS

A fullband transfer function, estimated with the three different configurations, is shown in Fig. 2. A 17 taps long band-pass filter with 10 taps flat delay and white noise as input signal, $x(n)$, was used to simulate the system. Notice the large bias error at the subband frequency edges of the standard open-loop system. The modified open-loop, $D = 90$ taps, and the closed-loop structures almost eliminates the problem. Figure 3 shows convergence rates. In this example, the echo path was simulated using a measured room impulse response with 21 taps flat delay and white noise as the input signal, $x(n)$. No system noise was added. The modified open-loop configuration has the fastest convergence rate.

In the last two figures, real life data was recorded in a stereo video conference situation. Two microphones in the far-end room transmits speech to two loudspeakers in the near-end room, the signal is also

recorded and used as the far-end signal. The echo is recorded with two microphones in the near-end room. Two parallel systems are used to cancel the echo, one for each channel. Each system has both far-end channels as input signal, and the algorithms do not take in account the high correlation between the two channels. Figure 4 shows the echo return loss enhancement (ERLE) for the modified subband algorithm (solid line) and a NLMS system (dotted line). In both systems, the estimated impulse response is 1536 taps long. Figure 4 indicates only a small improvement of the subband algorithm over the standard NLMS algorithm. This result is because the NLMS converges well in the frequency segments with high power, but in other segments convergence is slow. A spectral analysis of the two methods of the segment 9.5 second to 11 second is plotted in Fig. 5. The echo spectrum is plotted with dotted line and the residual echo spectrum from the subband and the NLMS algorithm with solid and dashed line, respectively. In the higher frequency regions the subband algorithm outperforms the standard NLMS algorithm, which is confirmed in informal listening test.

4. DISCUSSION

We have proposed a modification to the open-loop delayless subband echo canceller, to reduce the estimation error due to truncation of non-causal subband impulse responses. The modified structure reduces the residual echo signal when the echo source flat delay is short. It has faster convergence rate than the closed-loop configuration, since stability is guaranteed for larger stepsize parameter μ .

REFERENCES

- [1] A. Gilloire and M. Vetterli, "Adaptive filtering in subbands with critical sampling: Analysis, experiments, and application to acoustic echo cancellation," *IEEE Trans. on Signal Processing*, vol. 40, no. 8, pp. 1862-1875, August 1992.
- [2] D. R. Morgan and J. C. Thi, "A delayless subband adaptive filter architecture," *IEEE Trans. on Signal Processing*, vol. 43, no. 8, pp. 1819-1830, 1995.
- [3] M. M. Sondhi and W. Kellermann, *Advances in Signal Processing*, chapter 11, pp. 327-356, Marcel-Dekker, 1992.
- [4] S. K. Mitra G. A. Clark and S. R. Parker, "Block implementation of adaptive digital filters," *IEEE Trans. on Circuits and Systems*, vol. CAS-28, no. 6, pp. 584-592, June 1981.

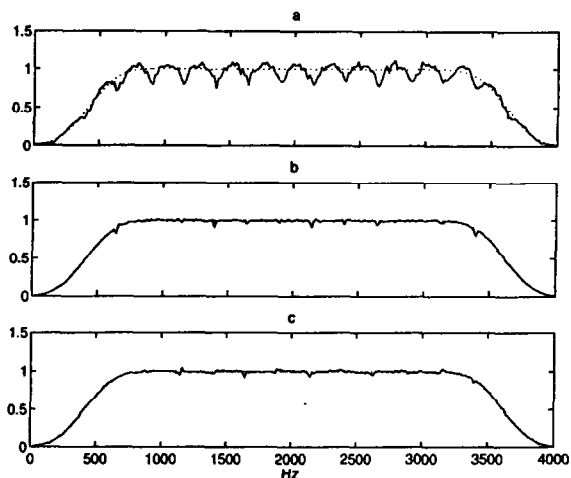


Figure 2: (a) The open-loop estimated transfer-function (solid line). The true transfer-function (dotted line). (b) The modified open-loop estimate (solid line). (c) The closed-loop estimate (solid line).

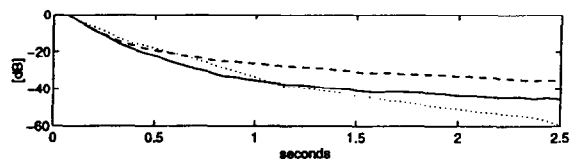


Figure 3: Misalignment. The modified open-loop convergence (solid line). The open-loop convergence (dashed line). The closed-loop convergence (dotted line).

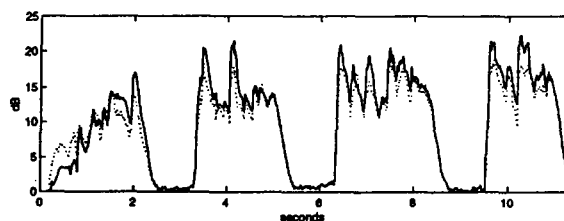


Figure 4: Echo return loss enhancement on recorded speech data in a stereo environment. Solid line: modified subband algorithm. Dotted line: NLMS algorithm.

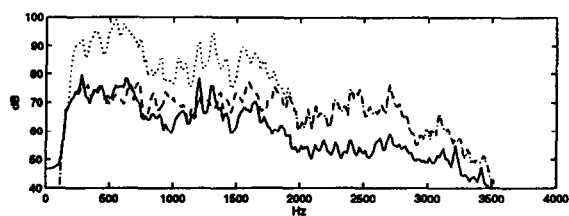


Figure 5: Spectrum analysis of echo and residual echo. Dotted line: echo signal spectrum. Solid line: modified subband residual echo spectrum. Dashed line: NLMS residual echo spectrum.