Subband acoustic echo canceller
using two different analysis filters and 8th order projection algorithm

Akira Nakagawa, Yoichi Haneda, and Shoji Makino
NTT Human Interface Laboratories
3-9-11, Midori-cho, Musashino-shi, Tokyo, 180 Japan
Tel: +81 422 59 3546, Fax: +81 422 60 7811, Email: nakagawa@splab.hil.ntt.co.jp

ABSTRACT

This paper presents a subband acoustic echo canceller (SBEC) using two different analysis filters and an 8th order complex affine projection algorithm (APA). This SBEC uses different analysis filters to divide the input signals and echo signals. The analysis filter for the input signals is designed to improve the convergence speed, and the analysis filter for the echo signals is designed to achieve high echo return loss enhancement. This SBEC also uses an 8th order complex APA to further improve the convergence.

We implemented the proposed SBEC on digital signal processors and evaluated it. The results show that our SBEC converges five times faster than the conventional SBEC.

1. INTRODUCTION

Acoustic echo cancellers are widely used for teleconferencing and hands-free telecommunication systems to overcome acoustic feedback, making conversation more comfortable. The fullband acoustic echo canceller has the advantage of having no delay, but it involves much computational complexity. The subband acoustic echo canceller (SBEC) has some delay, but lower computational complexity than the fullband one.

An SBEC reduces the eigenvalue spread by dividing the input signals into smaller frequency subbands. This whitens the input signals, resulting in fast convergence [1]. Furthermore, computational complexity is reduced because downsampling expands the sampling interval and reduces the number of taps needed for the adaptive digital filters (ADFs).

The use of critical sampling in the SBEC results in undesirable aliasing effects due to a non-ideal low pass filter (LPF). To avoid aliasing due to non-ideal LPFs, the downsampling rate is chosen to be smaller than the number of subbands. However, when a small downsampling rate is chosen, the frequency characteristics of the LPF remain in the frequency characteristics of each subband signal and the eigenvalue spread of each subband input signal is increased. This causes the slow aymptotic convergence of a conventional SBEC [2], [3].

To overcome this, one suggestion is to increase the bandwidth of the analysis filters relative to the synthesis filters [4], and another is to use an affine projection algorithm (APA) for each subband [5].

We propose changing the analysis filter characteristics for input signals and echo signals separately to improve convergence. We also use the APA to further improve the convergence. This made it possible to achieve fast convergence and high ERLE by using a lower-order projection algorithm than the one used in the previous study [5].

We implemented these techniques on digital signal processors (DSPs). This paper describes these investigations and presents the results of the evaluation of the implemented system.

2. PROPOSED SUBBAND ACOUSTIC ECHO CANCELLER

In this section, we describe the main features of our proposed echo canceller. Figure 1 shows a block diagram of

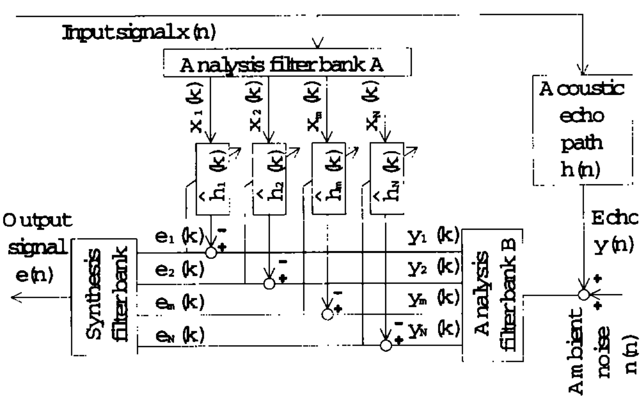

Fig. 1. Block diagram of proposed subband echo canceller.

the echo canceller. This echo canceller uses two different analysis filters to improve the convergence speed with high ERLE and a complex APA to further improve convergence.

2.1 Two different analysis filters

In the SBEC, an analysis filter to divide the echo signals is desired to reduce the aliasing effect and to maintain high speech quality. The conventional SBEC uses the same analysis filter to divide the input signals. As a results, the characteristics of input signals for the adaptive filters are not fully whitened, that is, the frequency characteristics of LPF remains. This causes slow asymptotic convergence.

To overcome this problem, we used two different analysis filter banks as shown in Fig. 1. The analysis filter for input signals is designed to reduce the eigenvalue spread in order to achieve fast convergence. The analysis filter for echo signals is designed to reduce the aliasing due to downsampling in order to maintain high speech quality and achieve high ERLE.

Figure 2 shows the frequency characteristics of these two analysis filters after downsampling. These downsampling rates are half of the number of subbands, so the cutoff frequency is $\pi/2$ and the needed frequency band is from $-\pi/2$ to $\pi/2$. The filter for input signals has broad characteristics, but the filter for the echo signals has sharp characteristics. These analysis filters can be obtained by truncating the ideal LPF to different lengths.

Because the filter characteristic for the input signals includes the aliasing, it becomes flatter. The ADFs converge faster. The aliasing does not affect the pass band of the signals, so aliasing does not affect the synthesized error.

In this filter bank system, the ADF has to estimate the impulse response added characteristic of LPF for echo signals. The adequate ADF length in each subband is given by the following equation considering the difference in the two LPF lengths.

$$L_m = \frac{L}{R} + \frac{L_B - L_A}{R} \tag{1}$$

where

L_m : ADF length for each subband,
L : impulse response length,
R : downsampling rate,
L_A : LPF length for input signal,
L_B : LPF length for echo signal.

We confirmed the relationship between the ERLE and the adequate ADF length for various LPF lengths using computer simulation as shown in Fig. 3. The original impulse response length is $L = 512$, the number of subbands is $N = 64$ ($0\sim2\pi$), and downsampling rate is $R = 32$. The sampling frequency is 16 kHz. This shows that the shortest ADF length to achieve the highest ERLE can be calculated from the difference between the two LPF lengths as in equation (1).

2.2 Complex affine projection algorithm

The APA has been used for the acoustic echo canceller because it converges faster than the normalized least-mean-squares (NLMS) algorithm for speech input. This fast convergence is achieved by whitening the speech input in the fullband echo canceller.

On the other hand, in the SBEC, auto-correlation of the speech remains in the subband signals. Therefore, if we

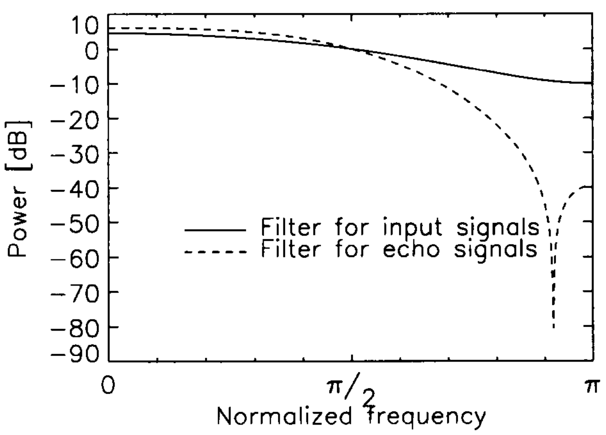

Fig. 2. Frequency characteristics of two different analysis filters. The cutoff frequency of both filters is $\pi/2$.

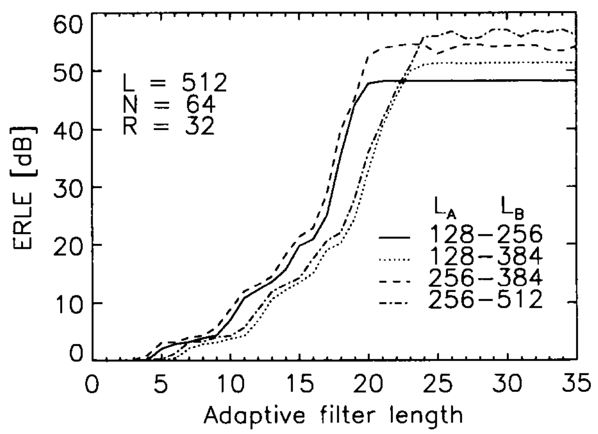

Fig. 3. Required ADF length for various LPF lengths, L_A and L_B for input and echo signals.

apply the APA to the SBEC [5], the resulting convergence should be faster than the conventional SBEC, that uses the NLMS algorithm.

Because the subband signals are complex (owing to complex modulation), the ADF in each subband must handle complex signals. Therefore, we extended the APA to a complex APA. Furthermore, we use intermediate variable z instead of ADF coefficients in order to reduce the computational complexity.

The *p-th* order complex APA updates the ADF coefficient vector z as follows:

$$z^*(k+1) = z^*(k) + s_p(k)\boldsymbol{x}^*(k-p+1)$$
$$\boldsymbol{s}(k) = [0 \quad s_1(k) \quad s_2(k) \cdots s_{p-1}(k)] + \alpha \boldsymbol{g}^*(k)$$
$$\boldsymbol{s}(k) = [0 \quad s_1(k) \quad s_2(k) \cdots s_{p-1}(k)]^T + \alpha \boldsymbol{g}^*(k)$$
$$\boldsymbol{R}(k) = \boldsymbol{x}(k)\boldsymbol{x}^H(k) + \delta \boldsymbol{I}$$
$$\boldsymbol{e}(k) = \boldsymbol{y}(k) - \hat{\boldsymbol{y}}(k) + \boldsymbol{n}(k)$$
$$\hat{y}(k) = \sum_{i=1}^{p-1} \{s_i(k-1) r_{i+1}(k)\} + \boldsymbol{z}^*(k)\boldsymbol{x}(k)$$
$$\boldsymbol{R}(k) = \boldsymbol{x}(k)\boldsymbol{x}^H(k)$$
$$\boldsymbol{x}(k) = [x(k) \quad x(k-1) \cdots x(k-p+1)]^T$$
$$\boldsymbol{e}(k) = [e(k) \quad (1-\alpha)e(k-1) \cdots (1-\alpha)^p e(k-p+1)]^T$$
$$\boldsymbol{r}(k) = x(k)\boldsymbol{x}^*(k)$$
$$\boldsymbol{n}(k) = [n(k) \quad n(k-1) \cdots n(k-p+1)]^T,$$

where

I : unit matrix,

α : scalar step size ($0<\alpha<2$),

δ : small positive constant,

^{H} : complex conjugate and transpose,

* : complex conjugate.

The small positive constant δ is introduced to stabilize the inverse matrix to calculate *g(k)* [6]. The computational complexity can be more reduced by using FTF (fast transversal filter) [7].

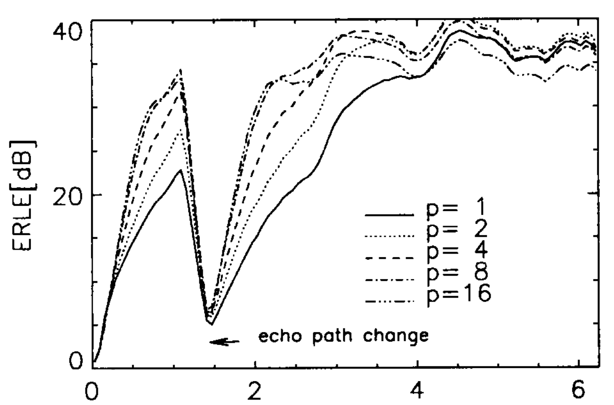

Fig. 4. Convergence curves for various projection order in the proposed subband echo canceller. The convergence speed of *p=8* is almost the same as that of *p=16*. (computer simulation)

3. COMPUTER SIMULATION

To confirm the advantage of our proposed SBEC, and to determine the hardware specifications, we evaluated the convergence speed by using computer simulations, using the impulse response measured in a conference room. The number of subbands was set to 64 (0~2π), and the downsampling rate was set to 32 to avoid aliasing. The ADF length of each subband was set to 44. The length of the echo path impulse response was 1280. The sampling frequency was 16 kHz. Ambient noise of -40 dB was added to the echo signal.

Figure 4 shows the ERLE curves for the average of 30 trials speech. The convergence speed with projection order $p=8$ is nearly equal to one with $p=16$. These results show that the saturation order of the APA in the proposed SBEC is half that in the previous report [5]. Note that the spectral envelope of input signals is almost flat in our proposed system, so there is no need for the extra projection order to whiten its envelope.

4. IMPLEMENTATION AND EVALUATION

Based on the simulation results, we implemented our system on DSPs. The sampling frequency was 16 kHz and the number of subbands was 64. The ADF lengths can be set to upto 100 taps. The projection order can be set to a maximum of 8th order for each.

We evaluated the convergence speed of the implemented system in a conference room, which has a reverberation time of 200 ms. The evaluation used 48 Japanese and English female and male speech signals. The convergence curves were obtained by averaging the 48 trials.

First, we confirmed the effect of using two different analysis filter bank system compared with the conventional method. Both SBECs used the complex NLMS algorithm, which corresponds to the projection order $p=1$. Figure 5 shows the convergence curves of each SBEC. After 6 seconds the mean squared error (MSE) for the conventional filter bank system is down about -23 dB but the proposed one is down about -28 dB.

Next we evaluated the convergence speed of proposed

SBEC with different projection orders. Figure 6 shows the evaluation results. The convergence speed with projection order p=8 achieved four times faster convergence than one with p=1 until the MSE reaches -25 dB. The convergence speed with projection order p=8 achieved five times faster convergence than one which uses the conventional filter bank and NLMS.

We also evaluated the convergence speed using white noise for input signals shown in Fig. 6. When using projection order p=8, the convergence speed for speech input is the same as the one for white noise input.

This successful implementation makes possible high-quality echo cancellation systems.

5. CONCLUSION

We implemented an SBEC that uses two different analysis filters and the 8th ordered complex APA on DSPs. Using two different analysis filters improved the convergence speed with high ERLE. The 8th order complex APA further improved the convergence speed. The proposed SBEC achieved five times faster convergence than the conventional system until the MSE reaches -25 dB. This fast-convergence SBEC should lead to high-quality teleconferencing systems.

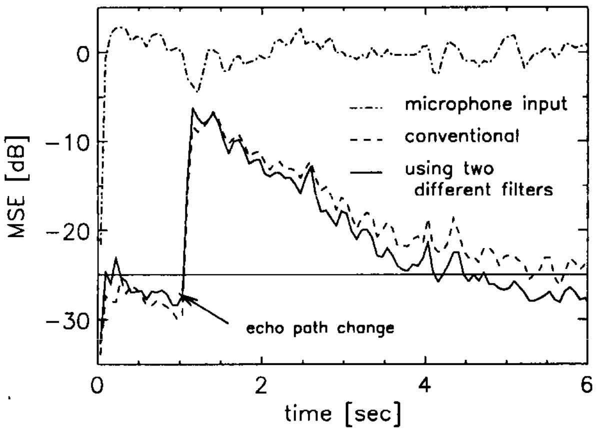

Fig. 5. Convergence curves of two different analysis filters and conventional system. The proposed filter bank system achieves faster convergence and a higher echo reduction level.

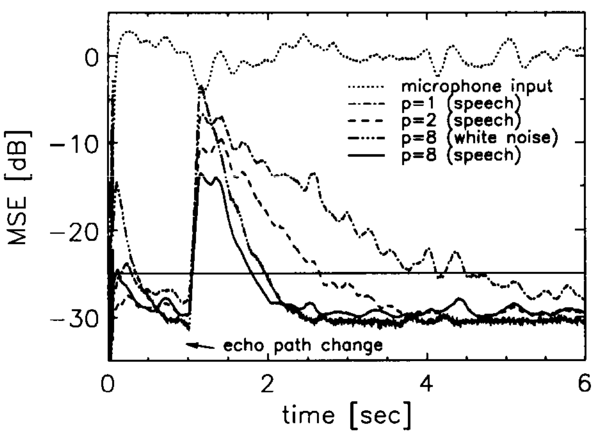

Fig. 6. Convergence curves for various projection orders.

ACKNOWLEDGMENTS

We would like to thank Dr. N. Kitawaki, Mr. J. Kojima, and Dr. Y. Kaneda for their fruitful discussions.

REFERENCES

[1] A. Gilloire, "Experiments with sub-band acoustic echo cancellers for teleconferencing," *Proc. ICASSP87*, pp. 2141-2144, 1987.

[2] W. Kellermann, "Analysis and design of multirate systems for cancellation of acoustical echoes," *Proc. ICASSP88*, pp. 2570-2573, 1988.

[3] D. R. Morgan et al., "Slow asymptotic convergence of LMS acoustic echo cancelers," *IEEE Trans. SP* vol. 3, No. 2, 1995.

[4] D. R. Morgan et al., "Experimental results with increased bandwidth analysis filters in oversampled, subband acoustic echo cancelers," *IEEE Signal Processing Letters*, vol. 2, No. 1, Jan. 1995.

[5] S. Makino et al., "SSB subband echo canceller using low-order projection algorithm," *Proc. ICASSP96*, pp. 945-948, May 1996.

[6] Y. Haneda et al., "Implementation and evaluation of an acoustic echo canceller using the duo filter control system," *Proc. International Workshop on Acoustic Echo Control*, June, 1995.

[7] M. Tanaka et al., "Fast projection algorithm and its step size control," *Proc. ICASSP95*, pp. 945-948, May 1995.