SUBBAND DOUBLETALK DETECTOR FOR ACOUSTIC ECHO CANCELLATION SYSTEMS

Tao Jia[1], Ying Jia[2], Jian Li[2], Yongge Hu[2]

[1]Department of Electronics, Peking University, Beijing, 100871, China
[2]Intel China Research Center, Beijing, 100020, China
E-mail< tao.jia@263.net, ying.jia@intel.com>

ABSTRACT

Though subband adaptive filter has been studied for several years and already widely used in real-time acoustic echo cancellation (AEC) systems, the study of doubletalk detector in subband AEC has not been reported. In this paper, two subband doubletalk detectors are proposed and simulated. An objective measurement is also employed to quantify the performance of the subband doubletalk detectors in contrast with fullband implementation. The simulation shows the proposed subband doubletalk detectors can increase the detection probability of doubletalk at least 3 percent over traditional fullband implementation under same false alarm probability.

1. INTRODUCTION

Driven by increasing demand for hands-free audio input for communication and speech recognition applications on desktop and mobile PCs, acoustic echo canceller has been a must component in audio driver to support multimedia applications. An acoustic echo canceller is usually composed of an adaptive filter and a doubletalk detector.

Doubletalk detector is to detect the presence of near-end signals and stop the adaptation of echo path filters to keep it from convergence. The research on doubletalk detection has more than decades history and several practical algorithms have been proposed.

Adaptive filters with a subband structure has shown the advantages of accelerated converging rate and reduced computation complexity [3][5][9]. So the subband adaptive filtering has been widely applied in real-time signal processing systems. However the advantages of subband structure has not been investigated and validated for doubletalk detector under the framework of subband adaptive filtering. It's probably because of the lack of objective evaluation method. An objective evaluation technique was published in [7], which could help us to analyze the performance of subband doubletalk detector.

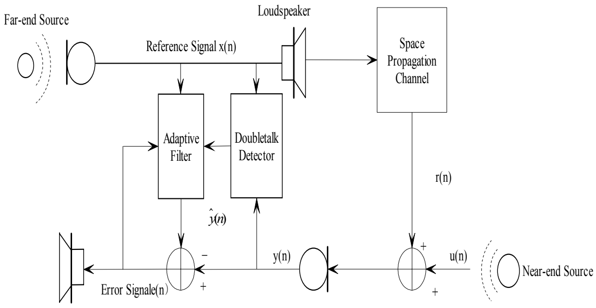

Figure.1. The structure of AEC systems

To combine the doubletalk detector with subband adaptive filter, we have to answer these two questions: 1) how to design the doubletalk detector at each subband? 2) how to make the global decision to achieve best performance? In this paper we propose and simulate two normalized cross-correlation vector based subband doubletalk detectors that are integrated with subband adaptive filter.

2. SUBBAND AEC & FULLBAND DOUBLETALK DETECTOR

2.1. Two Components of AEC

An acoustic echo canceller is shown in figure 1. The reference signal x(n) and the error signal e(n) are used to update the coefficients of the echo path filters to minimize the least square error or mean square error. The received signal of microphone y(n) is mixed by echo signal r(n) and near-end speech signal u (n). We have

$$y(n) = \mathbf{h}^T \mathbf{x}_L(n) + u(n) \qquad (1)$$

where

$$\mathbf{x}_L(n) = [x(n) \quad x(n-1) \quad \cdots \quad x(n-L+1)]$$

and

$$\mathbf{h} = [h_0 \quad h_1 \quad \cdots \quad h_{L-1}]^T$$

is echo path coefficient vector of length L.

Doubletalk detector is to detect the presence of u(n) in y(n) with knowledge of the reference signals x(n). Most doubletalk detection algorithms are to define a decision variable ξ based on the available signals x(n) and y(n). Usually we declare the presence of doubletalk when ξ < T where T is a predefined threshold. The decision on the presence of near-end speech signal (hypothesis 1) is made when ξ < T. When there is no near-end speech, we call the probability of ξ less than T the false alarm probability under hypothesis h0 (Pf). When near-end speech presents, the probability of ξ great than T is called detection probability under hypothesis h1 (Pd).

2.2. Subband Adaptive Filter

The structure of subband adaptive filter in AEC systems [3] is shown in figure 2.

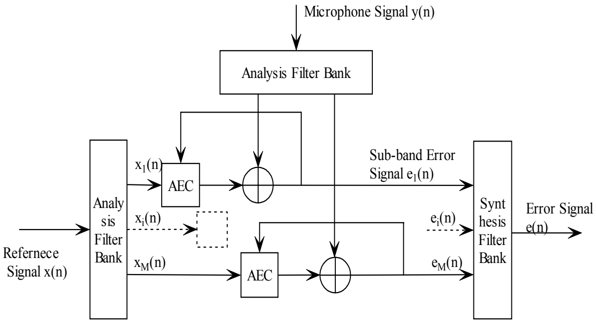

Figure.2. Structure of subband adaptive filter

The reference signal and microphone signal is passed through analysis filter banks and downsampled to get subband signals. Suppose there are M subbands. When down-sampling rate r=M, i. e. critical downsampling is used, the alias due to the imperfection of the analysis filter will decrease the performance of adaptive filtering significantly [5]. So we choose to use noncritical downsampling i.e. r<M. Because each of the analysis filters is the complex modulated form of prototype low-pass filter, we need only calculate M/2+1 channels of subband signal [9].

In subband adaptive filtering, the reduced sampling rate in each subband results in reduced computation complexity. On the other hand, reduced filter length relaxes the requirement of computation precision. Because the eigenvalue spread of subband signal is less than that of the fullband signal, the convergence rate of some adaptive filtering algorithm such as NLMS is accelerated.

2.3. Fullband Doubletalk Detector

For line echo canceller, an early doubletalk detector was proposed by A. A. Geigel [1]. It was based on power of the microphone signal which is mixture of echo signal and near-end speech signal. The optimal threshold of this algorithm may change dramatically in different situation so that it has to be re-measured in a specific situation. In [4], a doubletalk detection algorithm based on orthogonal theorem is proposed. This algorithm is more robust than first one but its decision variable is also not well normalized [8]. In [8], J. Benesty et al proposed a new doubletalk detection algorithm based on normalized cross-correlation vector which can avoid the problem with previous two algorithms. The decision variable of algorithm in [8] is shown as

$$\xi = \sqrt{\mathbf{r}_{xy}^T \mathbf{R}_{xx}^{-1} \mathbf{r}_{xy} / \sigma_y^2} \qquad (2)$$

where rxy is the cross-correlation vector of scalar x(n) and vector y(n), Rxx is the auto-correlation matrix of reference signal x(n) and σy2 is the variance of microphone signal y(n). This decision variable is well normalized so that

This decision variable is well normalized so that when there is no near-end speech it is equal to 1[8].

To reduce the computation of equation (2), we can substitute the wiener solution $\mathbf{R}^{-1}\mathbf{r}_{xy}$ with its estimation $\hat{\mathbf{h}}$ and rewrite (2) as

$$\xi = \sqrt{\mathbf{r}_{xy}^T \hat{\mathbf{h}} / \sigma_y^2} \qquad (3)$$

In [8], ξ is defined using the signals x(n) and y(n) with a fullband form.

To compare the performance of these algorithms, an objective evaluating method adopting receiver operation character (ROC) curve which is used in radar detection was proposed in [5]. In terms of ROC curve, the normalized cross-correlation vector based doubletalk detector was shown better than other algorithms above [7].

3 THE SUBBAND DOUBLETALK DETECTOR

T. Gänsler has proposed a doubletalk detection algorithm based on coherence [6] in which the coherence function of the input signal x(n) and microphone signal y(n) is calculated. Due to the character of speech, the coherence function's value was only calculated in a specific frequency interval to get the decision variable or more generally the value of coherence function is weighed differently according to the frequency in order to get a better decision variable. Although this algorithm may have improved performance, it is hard to be implemented in real-time systems because of the extensive complexity in computing coherence function [4].

The reference signal and microphone signal for doubletalk detector in an AEC system with subband adaptive filter are of the subband forms. To construct the doubletalk detector using subband x(n) and y(n), we can realize M/2+1 independent subband doubletalk detectors in M/2+1 subbands. Though independent controlling of subband adaptive filters may be suitable in practice, to compare Pd with fullband detector, a global decision about the presence of near-end signal is obtained by combining local decisions made by subband doubletalk detectors in this paper.

In this paper, we proposed two combination methods. One is to select a subband with best detection performance and use its local decision variable as the global decision. This method will be called optimal subband selection method in this paper. Another intuitive idea is to weigh the subband decision variables according to the signal power of their bands and get the global decision variable as following

$$\xi = \sum_{i=0}^{M/2} \xi_i \sigma_i^2 \Big/ \sum_{i=0}^{M/2} \sigma_{yi}^2 \qquad (5)$$

where ξ_i is the local decision variable at ith subband, and σ_{yi}^2 is the variance of the microphone signals at ith subband. It will be called weighted method in this paper.

For faire comparison of various doubletalk detection algorithms, an objective evaluating technique is proposed in [5], in which the performance of doubletalk detector is measured by Pd under a given Pf. The first step of this technique is to obtain a threshold under a given Pf in the absence of near-end speech signal and this threshold is used to measure the Pd in different far-end to near-end speech ratio (NFR, σ_u^2/σ_x^2)[8]. The measured Pd as a function of NFR is known as receiver operating characteristic (ROC) in detection theory.

This objective evaluating technique is employed to compare the two subband doubletalk detection algorithms we proposed here with the fullband detector. The normalized cross-correlation vector based doubletalk detector was used for fullband and subband implementation in our simulation because of its simplicity in computation and excellent performance in practice.

4. SIMULATION AND RESULTS

4.1 Simulation Setup

In our simulation, the impulse response of the propagation channel is simulated using an image mode method [9], and then truncated to a length of 2048 points at the sampling rate of 8 kHz. The fullband echo signal is generated by filtering the reference signal through the impulse response of the propagation channel. The fullband microphone signal is the sum of the echo signal and near-end speech.

In subband structure the fullband reference signal and subband microphone signal are divided into 16 subbands and downsampled at a rate of 12 to obtain corresponding subband signals. The synthesis filter banks and analysis filter bands are designed according to [1]. The amplitude-frequency response of prototype low-pass base-band analysis filter is shown in figure 3.

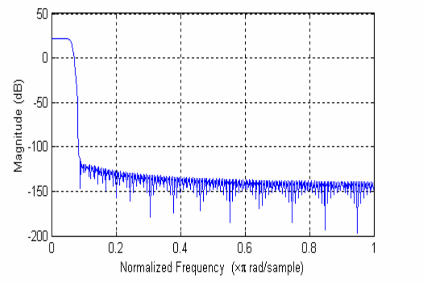

Figure.3. Amplitude-frequency response of low-pass prototype filter.

To get the estimation of echo path, a 20,000-points-sequence of Gaussian white noise is generated as reference signal in training. NLMS algorithm is used for fullband and subband adaptive filters to get the estimation of echo path because of its simplicity. The adaptive filter has a length of 384 in fullband and 32 in subband. The coefficients of adaptive filters used as the estimation of echo path are recorded after they had converged sufficiently.

Another variable to be estimated is $\mathbf{r}_{xy}$ and it was estimated by using sliding window method as

$$\hat{\mathbf{r}}_{xy}(i,n) = \lambda \hat{\mathbf{r}}_{xy}(i,n-1) + (1-\lambda)x(n-i)y(n), \quad i = 0...L-1 \; (5)$$

here λ is called forgetting factor. In our simulation, $\lambda = e^{-1/L}$ and L is the length of the adaptive filter. The power of the subband signals is estimated in the same way.

In order to meet with practical situation, the fullband reference signal and near-end signal are both speech signal which are used to measure Pd under a given Pf. Each of them is a passage with a length of 175s at the sampling rate of 8 kHz. The NFR is adjusted by changing the amplitude of the near-end speech signal. For the first proposed subband detector the Pd and Pf are measured independently in each of the subbands.

4.2 Simulation Results

In the simulation of optimal subband selection method, the measured detection probability of doubletalk detectors in fullband and subbands numbered from 0 to 8 varies according to the varying of NFR when Pf=0.1 and Pf=0.3 as shown in Figure 4 and 5.

It can be seen from above results that the performance of doubletalk detector in subband0 is the best and the performance of doubletalk detectors in other subbands is worse than the fullband's. It can be explained by the strong coherence between reference signal and echo signal in the lowest frequency band. Therefore in optimal subband selection method only the decision variable at lowest frequency band is used. It is similar to [6] in which only the values of coherent function in frequency of interest were used. But the doubletalk detector at subband0 is of less computation complexity.

The ROC curve of power weighted method is shown in figure. 6. We can see that the performance of weighted method is worse than that of the detector in subband0 i. e. optimal subband selection method but is better than that of the fullband one. This algorithm will be useful when the near-end interfering signal is band-passed and its main power is not located in low frequency band.

The simulation result shows that optimal subband selection method can improve Pd at least 3 percent over fullband implementation.

In our simulation, the two proposed subband doubletalk detectors outperform fullband ones in detection stability. In these two subband methods, optimal subband selection method can improve Pd at least 3 percent over fullband implementation. This result shows that subband technique not only improve adaptive filtering but also enhance the performance of doubletalk detector in AEC systems.

5. CONCLUSION

The subband structure has been known for reducing the complexity of adaptive filter and improving its convergence rate. In this paper, we have shown that subband structure can also improve the performance of doubletalk detector. The two proposed subband doubletalk detectors in our simulation are better than fullband ones when they are all based on normalized cross-correlation vector.

7. REFERENCES

[1]. D. L. Duttweiler, “A twelve-channel digital echo canceler”, IEEE Trans. Comm. vol. 26 No. 5, May 1978.

[2]. J. B. Allen and D. A. Berkeley, “Image method for efficiently simulating small-room acoustics”, J. Acoust. Soc. Am., Vol. 65. no. 4, pp. 943-950 (1979)

[3]. W. Kellermann, “Analysis and design of multirate systems for cancellation of acoustic echoes”, Proc. ICASSP 1988, pp.2570-2573.

[4]. Hua Ye and Bo-Xiu Wu, “A new doubletalk detection algorithm based on the orthogonality theorem”, IEEE Trans. Comm., vol. 39, pp. 1542-1545, Nov. 1991

[5]. 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, pp1862-1875, Aug. 1992.

[6]. T. Gänsler, M. Hansson, C.-J. Ivarssson and G. Salomonsson, “A double-talk detector based on coherence”, IEEE Trans. Comm. pp1421-1427 vol.44, No. 11, Nov 1996.

[7]. J. H. Cho, D. R. Morgan and J. Benesty, “An objective technique for evaluating doubletalk detectors in acoustic echo cancellers”, IEEE Trans. Speech and Audio Processing, vol. 7, No. 6, Nov 1999.

[8]. J. Benesty, D. R. Morgan and J. H. Cho, “A new class of doubletalk detectors based on cross-correlation”, IEEE Trans. Speech and Audio Processing, vol. 8, No. 2, Mar 2000.

[9]. P. Eneroth, S. L. Gay, T. Gänsler and J. Benesty, “A real-time implementation of a stereophonic acoustic echo cancler”, IEEE Trans. Speech and Audio Processing, vol. 9, No. 5, July 2001.

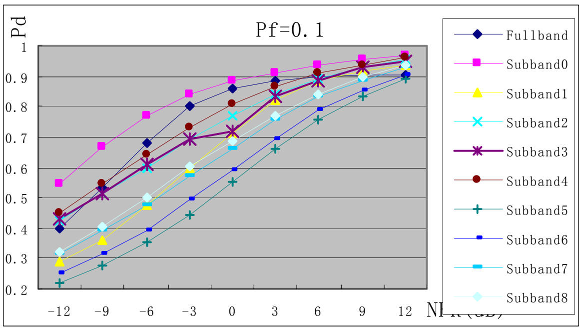

Figure.4. ROC curve of the first proposed subband doubletalk detector when Pf=0.1

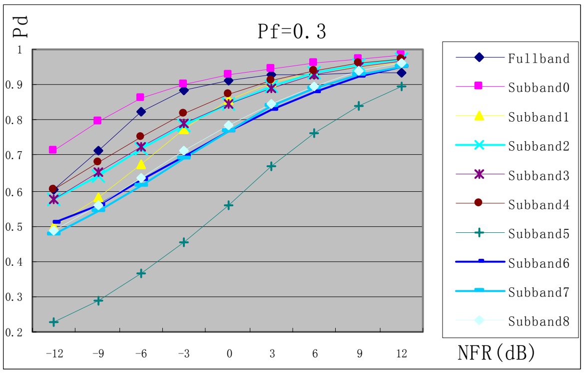

Figure.5. ROC curve of the first proposed subband doubletalk detector when Pf=0.3

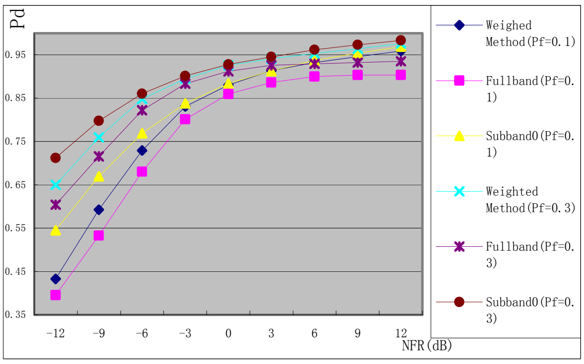

Figure.6. ROC curve of the second proposed subband doubletalk detector