Available online at www.sciencedirect.com

ScienceDirect

Signal Processing 86 (2006) 3752–3760

SIGNAL PROCESSING

www.elsevier.com/locate/sigpro

A robust transform domain echo canceller employing a parallel filter structure

Jiaquan Huo[a,b], Ka Fai Cedric Yiu[b,*], Sven Nordholm[a], Kok Lay Teo[c]

[a] Western Australian Telecommunications Research Institute (WATRI), A joint venture between The University of Western Australia and Curtin University of Technology, Perth, Australia

[b] Department of Industrial and Manufacturing Systems Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong, PR China

[c] Department of Mathematics and Statistics, Curtin University of Technology, Perth, Australia

Received 9 August 2005; received in revised form 5 March 2006; accepted 8 March 2006
Available online 6 April 2006

Abstract

The proper control of acoustic echoes is of vital significance in modern communication systems. Acoustic echoes are commonly combatted by means of acoustic echo cancellation. A critical issue in acoustic echo cancellation is the control of adaptation of the echo cancellation filter with the possible presence of near-end speech activities. In this paper, a novel echo cancellation structure is proposed. The proposed structure makes use of two parallel adaptive filters. Echo of the two parallel adaptive filters is tailored for a specific operating situation. By so doing, the contradicting requirements of robustness against near-end disturbance and fast convergence and tracking speed are achieved simultaneously.

Keywords: Echo cancellation; Double-talk; Filter bank; Subband adaptive filter

1. Introduction

The proper control of acoustic echoes is of vital importance in modern communication system. Acoustic echoes arise due to the acoustic coupling between the loud-speaker and the microphone at user terminals. With the long round-trip delay typical in today's mixed signal networks, unsuppressed echoes would be very annoying to the end-users, and to the extreme, can make a conversation impossible [1,2].

Echo cancellers (EC) have been developed to suppress echoes in communication networks [3]. An EC is essentially an adaptive filter, as illustrated in Fig. 1, that generates an echo estimate from the far-end signal. This echo estimate is subsequently subtracted from the microphone signal. The most attracting feature of an EC is its capability of providing a natural conversation pattern, i.e., allowing the users at both ends interrupting each other.

Nevertheless, during time periods in which users at both ends are active (known as double talk, or DT), the adaptation of the EC filter would be subject to perturbation of strong near-end speech signal. Such a strong perturbation would set the EC filter off its already converged state, resulting in an annoyingly high level of returned echo to the far-end user.

*Corresponding author. Tel.: +852 22415956; fax: +852 28592583.
E-mail address: cedric@hkucc.hku.hk (K. Fai Cedric Yiu).

0165-1684/$ - see front matter
doi:10.1016/j.sigpro.2006.03.004

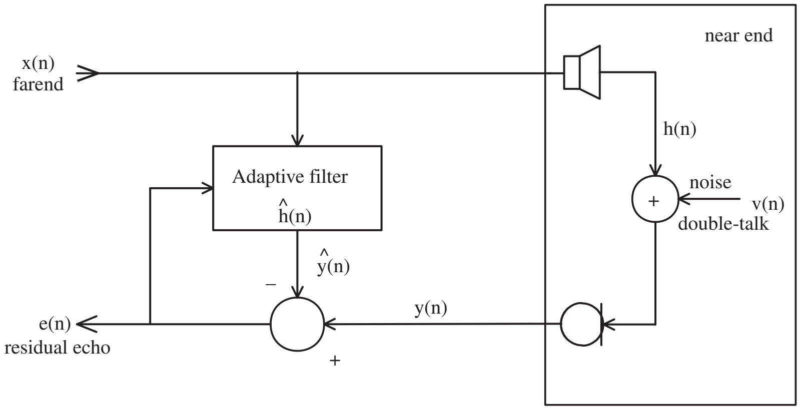

Fig. 1. An illustration of an echo canceller for acoustic channels.

A common way to alleviate the DT problem is to halt the adaptation of the EC filter whenever near-end speech activity is detected [4]. Due to statistical fluctuation and delay in parameter estimation, detection errors are inevitable. In [5], an adaptive filtering algorithm based on robust statistics is developed to cope with such detection errors.

In this paper, a novel parallel adaptive filter structure is proposed which is robust against double-talk, fast in tracking echo path variations due to random or systematic changes, and does not rely on a double-talk detector (DTD). It comprises of two FIR filters, each tailored for a specific operating situation. A controlling algorithm is employed to govern the exchange of the filter coefficients between the two filters, and to decide the filter to be used for cancelling the echoes.

One feature of the acoustic echo path is its extreme length [6]. The impulse response of a typical office requires an FIR filter of about 1000 taps at 8 kHz sample rate to model. The implementation cost of such a long adaptive filter is high. More importantly, when adapted in time domain with input signals of high spectral dynamics, the convergence speed of such a long adaptive filter is very slow. Adapting the EC filter in frequency domain or in subbands [7] can substantially speed up the convergence. Moreover, frequency domain and subband algorithms generally employ block processing and fast Fourier transform, which results in a considerable reduction in computational complexity. In this paper, we implement the echo canceller in subband with a delayless structure. Clear benefits of the proposed algorithm in terms of convergence rate, DT robustness and tracking are demonstrated with simulation results.

2. Robust statistics based adaptive filtering

Fig. 1 illustrates the basic idea of acoustic echo cancellation. The echo path is considered as a linear time-invariant system. The microphone signal $y(n)$ can therefore be written as

$$y(n) = \sum_{k=0}^{\infty} h_k x(n-k) + v(n), \tag{1}$$

where h_k is the kth tap of the impulse response of the echo path, $x(n)$ is the loudspeaker signal, $v(n)$ is the near-end signal. The echo canceller models the room impulse response with an N tap FIR echo cancellation filter (EC filter) and generates an estimate of the echo signal as

$$\hat{y}(n) = \sum_{k=0}^{N-1} \hat{h}_k x(n-k), \tag{2}$$

where $\hat{h}_k$ is the kth tap of the impulse response estimate. The echo estimate $\hat{y}(n)$ is then subtracted from the microphone signal and the error signal

$$e(n) = y(n) - \hat{y}(n) \tag{3}$$

is transmitted to the far-end. The EC filter $\hat{h}_k$ is made adaptive so that it can be adjusted to match vastly different room impulse responses and to track time variation of the echo paths.

Typically, the EC filter is adjusted to minimize the power of the error signal. It is well-known that an EC filter so adjusted is sensitive to extraneous disturbances. A single large uncorrelated error is sufficient to send the EC filter far off its optimal state. This makes the operation of the echo canceller unreliable in the presence of near-end speech. In order to reduce this sensitivity, the notion of robust