Special Feature

Echo Canceller with Noise Reduction Provides Comfortable Hands-free Telecommunication in Noisy Environments

Sumitaka Sakauchi†, Yoichi Haneda, Manabu Okamoto, Junko Sasaki, and Akitoshi Kataoka

Abstract

Audio-teleconferencing, video-teleconferencing, and hands-free telecommunication using personal computers in noisy offices are on the increase. To achieve hands-free telecommunication that is pleasant even in such environments, we have developed technology for an echo canceller that includes noise reduction. Here we describe a mechanism for suppressing ambient noise and the echoes embedded in that noise, and we introduce practical devices that implement that mechanism.

1. Introduction

Hands-free telecommunication uses a microphone to pick up the near-end[*1] speech and a loudspeaker to reproduce the far-end[*1] speech, thus eliminating the need for the handset of conventional telephones and the bother of wearing a headset. In a teleconference, this hands-free technology provides comfortable speech communication by letting the participants take notes and operate personal computers in the same way as in a conventional ordinary conference. In hands-free telecommunication, however, the microphone picks up the far-end speech from the loudspeaker in addition to the near-end speech. That may result in echoing that interferes with conversations or produces howling due to feedback (**Fig. 1**). An echo canceller cancels such echoing and prevents howling, so it is necessary technology for hands-free telecommunication.

Conventional echo cancellers have been developed for video teleconferencing assumed to be conducted in relatively quiet conference rooms. Recently, however, there has been a trend toward diversification in how hands-free telecommunication is used. Examples include telephone conferencing or speech communication via personal computers in noisy offices or

factories and hands-free telephoning while driving a vehicle. In environments such as these, office air-conditioner noise, office machinery noise, the sounds within a moving automobile, and other such ambient noise negatively affect the canceling effect of conventional echo cancellers and prevent communication. Echo cancellation with noise reduction can solve these problems. This technology provides comfortable hands-free telecommunication even in noisy environments where conventional echo cancellers are ineffective.

2. System description

Remote teleconferencing using an echo canceller with noise reduction is shown in **Fig. 2**. "Hello?" from the far-end (room A) emerges from the loudspeaker in room B. The listener hears it, but it is also picked up by the microphone as an echo. Thus, besides picking up the listener's reply "Yes?", the microphone in room B also picks up the echo and the noise of the air conditioner and office machinery. The echo canceller with a noise reduction function suppresses the echo and noise from the mixed signal and only the speech is transmitted to the far end. As a result, only clear transmitted speech, without echoes and ambient noise, is reproduced by the loudspeaker

† NTT Cyber Space Laboratories
Musashino-shi, 180-8585 Japan
E-mail: sakauchi.sumitaka@lab.ntt.co.jp

*1 Near-end and far-end are terms used to describe the locations involved in a teleconference.
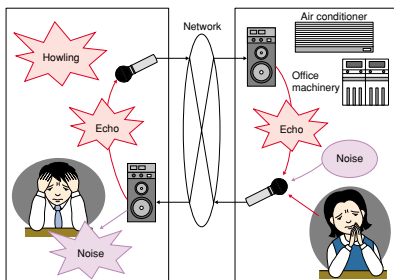
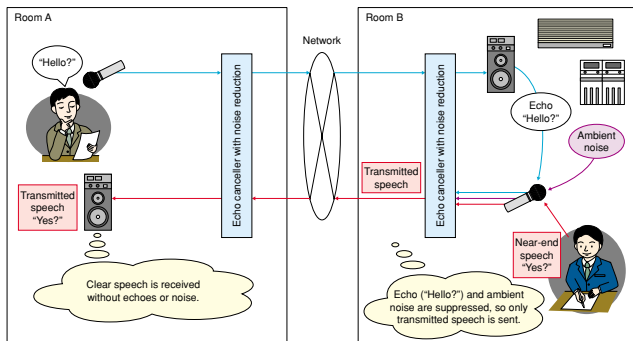Fig. 1. Various problems in hands-free telecommunication.

Fig. 2. Remote teleconferencing using echo canceller with noise reduction.

at the far end.

Only when there is no ambient noise can conventional echo cancellers [1] successfully learn the relationship between the received speech and the echoes caused by it and estimate how the speech will be changed into an echo. Then, the signal picked up by the microphone at the near end is processed to remove the estimated echo. When that is done, however, the ambient noise at the near end cannot be esti-

mated by the same method as used to estimate the echo, so it has been impossible to remove the noise from the microphone input signal. A further problem is that low-level echoes embedded in the noise cannot be removed either.

In contrast, an echo canceller with noise reduction [2] suppresses both the ambient noise and the echoes that are embedded in it by the method described below (**Fig. 3**).
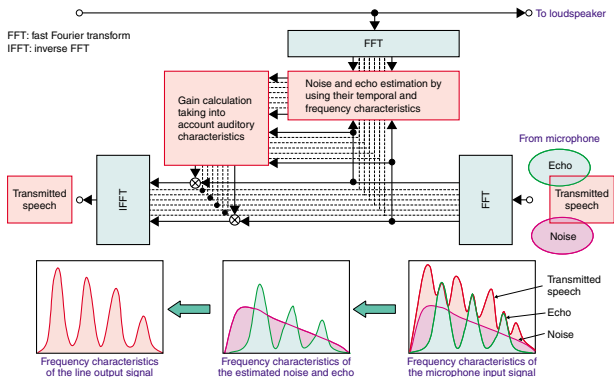
Fig. 3. Noise and echo reduction in the frequency domain.

2.1 Highly precise noise reduction using differences in the temporal and frequency characteristics of speech and noise

The stationary noise in offices (e.g., air-conditioner noise) differs from speech in that the changes in the signal over time are more moderate and the signal includes a wider range of frequency components than speech. These characteristics can be used to extract and accurately estimate the stationary noise. Then, the noise can be suppressed by subtracting the amount of estimated noise at each frequency from the microphone input signal including speech and noise. This technique achieves a maximum noise reduction of 20 dB in terms of sound pressure level. That level permits hands-free telecommunication with about the same degree of quietness as in a quiet video conferencing room, even when it is actually conducted in a noisy office for example.

2.2 High-performance echo elimination using adaptive filtering and echo reduction

The echo canceller with noise reduction first cancels out echoes that are louder than the ambient noise by using adaptive filtering, which is also used in conventional echo cancellers [3]. Next, the small residual echoes that are embedded in the ambient noise and

could not be eliminated in the first stage are estimated using the frequency domain amplitude characteristics of the echoes and of the speech that was the source of the echoes. Then, the echoes are eliminated by subtracting the volume that corresponds to the volume at each frequency of the estimated echo from the speech mixed with echoes that were picked up by the microphone.

2.3 Preventing distortion of processed speech by using human auditory characteristics

By using masking, an auditory phenomenon in which the minimum sound pressure level of sound that can be heard by the human ear is increased by the presence of other sounds, the echoes and noise are reduced to just below the level of audibility. In this way, excessive suppression processing of the input signal is avoided and communication by natural sounding speech is made possible.

3. Examples of echo cancellers with noise reduction

Two practical devices that implement an echo canceller with noise reduction are shown in **Fig. 4**.
(a) Audio-teleconferencing terminal: EC-13 (NTT

(a) Audio-teleconferencing terminal:
EC-13
(NTT East and NTT West)

(b) Echo canceller with noise reduction unit:
RealTalk 70
(NTT Advanced Technology Corp.)

Fig. 4.   Examples of echo cancellers with noise reduction.

East and NTT West)

Audio-teleconferencing among multiple participants can be accomplished easily by simply connecting this product to an analog telephone line. In addition to the echo canceller with noise reduction, the terminal is equipped with built-in microphones in the four corners of the unit and a built-in loudspeaker situated in the raised central part to reduce acoustic coupling [4], thus producing high-quality, easy-to-understand audio-teleconferencing, free of echoes and noise. There is a choice of two modes to suit different forms of teleconferencing. In the "receive-priority mode", the received speech volume is made stable for conferences where speech in that direction is dominant, such as a report. In the "bilateral mode", speech can be heard equally well in both directions for discussion-type conferencing.

This terminal can be plugged into the audio-in and audio-out connectors of a videophone or personal computer soundcard for use as a speakerphone. It is also possible to connect two of these terminals to allow even more people to participate in the conference. Furthermore, external microphones and loudspeakers can be connected to the audio-in and audio-out lines of the system to allow the terminal to be used by a large number of people in a large conference room.

(b) Echo canceller with noise reduction unit:
RealTalk 70 (NTT Advanced Technology Corp.)

This product is a stand-alone echo canceller for incorporation in video conferencing systems and other such equipment. It is designed for ITU-T G.722 7-kHz-band speech. Even with existing video conferencing systems, comfortable hands-free telecommunication can be achieved by turning off the built-in echo canceller and using this product to eliminate noise from the air conditioner and projector fan. It is possible to handle many talkers in a large conference room by connecting multiple microphones to a mixing amplifier.

4. Future developments

A software echo canceller that runs on a general-purpose CPU and is currently under development is shown in **Fig. 5**. If this software is incorporated into telecommunication applications for personal computers, it will enable comfortable hands-free telecommunication using personal computers.

Our technology for an echo canceller with noise reduction that provides comfortable hands-free telecommunication in noisy environments has currently reached a practical level for monaural speech up to the 7-kHz band for video-teleconferencing. In the future, we plan to work first on stereophonic
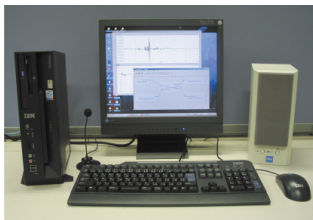

Fig. 5.   Software echo canceller for hands-free communication using a PC.

speech and then continue our research and development towards practical technology for multi-channel, multipoint, broadband signals, including music.

References

[1] J. Kojima, S. Makino, Y. Haneda, and S. Shimauchi, "High-Performance Acoustic Echo Canceller Development," NTT R&D, Vol. 44, No. 1, pp. 39-44, 1995 (in Japanese).

[2] S. Sakauchi, M. Tanaka, Y. Haneda, and K. Yamamori, "Ambient Noise and Residual Echo Reduction Using Short-time Spectral Amplitude Estimation," NTT R&D, Vol. 50, No. 4, pp. 246-252, 2001 (in Japanese).

[3] S. Makino and Y. Kaneda, "Exponentially Weighted Step-Size Projection Algorithm for Acoustic Echo Cancellers," NTT R&D, Vol. 44, No. 1, pp. 45-52, 1995 (in Japanese).

[4] A. Nakagawa, Y. Haneda, and K. Yamamori, "A Structure of Hands-free Communication Unit Using Microphones and Loudspeakers with a Flat Directional Pattern," NTT R&D, Vol. 50, No. 4, pp. 240-245, 2001 (in Japanese).

Sumitaka Sakauchi
Research Engineer, Media Processing Project, NTT Cyber Space Laboratories.
He received the B.S. degree in physics from Yamagata University, Yamagata in 1993 and the M.S. degree in physics from Tohoku University, Sendai, Miyagi in 1995. In 1995, he joined NTT. Since then, he has been engaged in R&D of acoustic echo cancellation and noise reduction. He is a member of the Acoustical Society of Japan (ASJ) and the Institute of Electronics, Information, and Communication Engineers (IEICE). He received the paper award from IEICE in 2001.

Yoichi Haneda
Senior Research Engineer, Media Processing Project, NTT Cyber Space Laboratories.
He received the B.S., M.S., and Ph.D. degrees in information science from Tohoku University, Sendai, Miyagi in 1987, 1989 and 1999, respectively. Since joining NTT in 1989, he has been investigating acoustic signal processing and acoustic echo cancellers. He is a member of IEEE, ASJ, and IEICE. He received paper awards from ASJ and IEICE in 2001.

Manabu Okamoto
Senior Research Engineer, Media Processing Project, NTT Cyber Space Laboratories.
He received a master's degree in design from Kyushu Institute of Design, Fukuoka in 1991. In 1991 he joined NTT Electrical Communication Laboratories. Since then, he has been researching the acoustic design of various kinds of teleconference systems. He is a member of ASJ and IEICE.

Junko Sasaki
Senior Research Engineer, Media Processing Project, NTT Cyber Space Laboratories. She received the B.E. and M.E. degrees in engineering from Osaka University, Osaka in 1990 and 1992, respectively. Since joining NTT Human Interface Laboratories in 1992, she has been investigating active noise control, noise reduction, and acoustic echo cancellers and teleconference systems. She is a member of ASJ.

Akitoshi Kataoka
Manager, Acoustic Information Group, Media Processing Project, NTT Cyber Space Laboratories.
He received the B.E., M.E., and Ph.D. degrees in electrical engineering from Doshisha University, Kyoto in 1984, 1986, and 1989 respectively. Since joining NTT Laboratories in 1986, he has been engaged in research on noise reduction, acoustic arrays, speech-signal processing, and medium-bit-rate speech and wideband coding algorithms for ITU-T standards. He is a member of ASJ, IEICE, and IEEE. He received the Technology Development Award from ASJ in 1996 and the Prize of the Commissioner of the Japan Patent Office from the Japan Institute of Invention and Innovation in 2003.