Nonlinear Echo Suppression Technology Enabling Quality Handsfree Talk for Compact Equipment

HOSHUYAMA Osamu

Abstract

This paper introduces a nonlinear echo suppression technology that enables compact equipment such as a notebook PC or cellular phone to be employed for handsfree use without a headset. In the suppression of echo, (audio feedback from the loudspeaker to the microphone) the distorted echo produced by vibrations of the compact loudspeaker or cabinet has been particularly difficult in the past. However, the newly developed frequency domain processing technology makes it possible to cancel the echo at a level that permits high-quality talk.

Keywords

cellular phone, handsfree call, echo canceller, distortion, nonlinear

1. Introduction

Recently, the dissemination of remote conferences in enterprises, VoIP calls on PC, video phone on cellular phones and the prohibition of the use of cellular phone during driving have increased the demand for a handsfree capability without using a headset or handset. The handsfree devices for this purpose currently use a linear echo canceller based on a linear adaptive filter in order to improve the communication quality. However, this technology still poses a problem, which is the incapability of canceling the nonlinear echo that is produced due to the nonlinear characteristics of a compact loudspeaker and the vibrations propagated through the cabinet from the loudspeaker to the microphone.

Techniques to cancel the nonlinear echo using a nonlinear adaptive filter or neural network have been studied, but these devices do not offer a realistic solution because they require a huge amount of computations, more than 10 times that of the linear echo canceller. Therefore, there is a strong need to cancel the nonlinear echo with small computations.

The nonlinear echo canceller proposed in this paper has overcome the above problem and enables the cancellation of nonlinear echo with a small amount of computation. The proposed method allows compact equipment such as a notebook PC or cellular phone to easily perform a quality two-way communication at a similar level to the high-quality TV conference system.[3,4]

2. Principles of Nonlinear Echo Generation

Let us assume that you are in Tokyo as shown in **Fig. 1**(a) and that you are holding a handsfree talk with Osaka, and see how the voice is heard. First, voice "Hello" spoken at Osaka is sent through the communication line and reproduced as "Hello" to the loudspeaker in Tokyo. Since the loudspeaker reproduced volume is high in handsfree talk, "Hello" output from the loudspeaker is fed back to the microphone and returned to Osaka with a delay. As a result, the person at Osaka hears his own voice like an echo. The echo not only makes the pronunciation of voices difficult but may also in the worst case

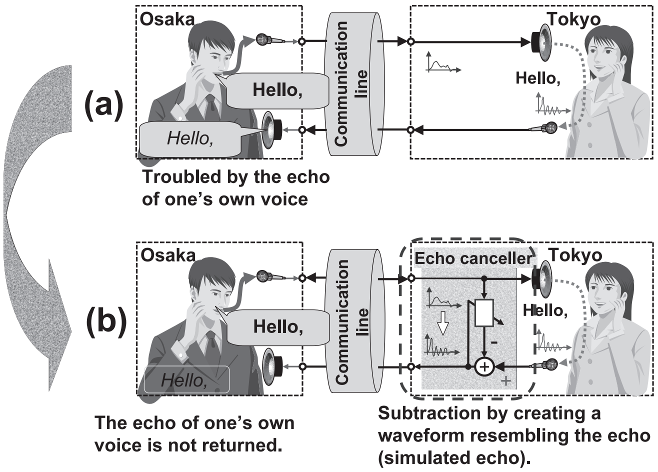

Fig. 1 (a) Principles of echo generation,
(b) Principles of the Echo Canceller.

produce a howling tone, in which case a satisfactory talk would no longer be possible.

The problem of echo is old and has usually been solved with the echo canceller technology. As shown in Fig. 1(b), the echo canceller regards the feedback path from the loudspeaker to the microphone as a filter. It generates a replica signal that resembles the signal fed back to the microphone, using a technology called the linear adaptive filter and uses it to cancel the signal from the microphone. The linear adaptive filter generates the echo replica by modeling the feedback path by applying a linear computation (the delay, addition and multiplication with a constant) in order to reproduce the signal sent to the loudspeaker. The linear echo canceller is used widely as an indispensable technology for the videoconferencing system, etc.

The problem of echo is more serious with compact equipment such as a notebook PC or cellular phone. **Fig. 2** shows the principles of nonlinear echo generation. When the voice from the other party is reproduced at a high volume through the loudspeaker on a compact cabinet, the loudspeaker and other parts in the cabinet vibrate, generating distorted sounds like a buzzer. This echo including the distortion is called the nonlinear echo in contrast to the linear echo that can be defined as a linear computation applied to the loudspeaker signal. A linear echo canceller cannot suppress the nonlinear echo completely, so the residual echo degrades quality of conversation. This problem is noticeable especially when a cheap loudspeaker or compact equipment is used.

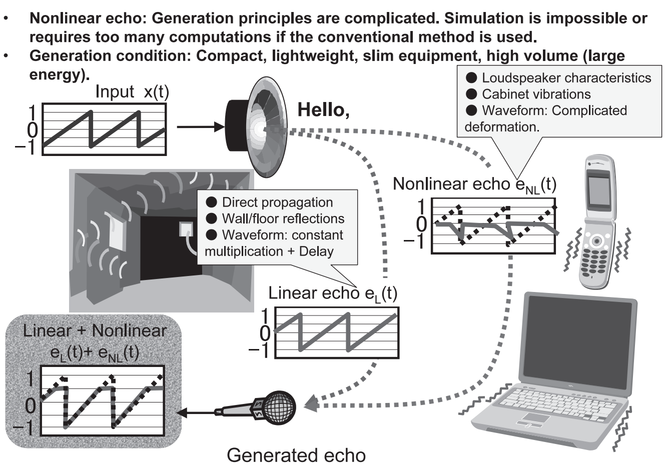

Fig. 2 Principles of nonlinear echo generation.

3. Conventional Nonlinear Echo Countermeasures

Many notebook PCs, cellular phones and car-use handsfree kits equipped with a handsfree talk function use a technique called the voice switch to cancel the nonlinear echo. This technique is based on the fact that in most of a conversation, only one of the parties is speaking. It reduces the reproduced volume from one party while the other party is speaking. However, when both parties speak at the same time or the surrounding noise is significantly loud, the voices tend to be intermittent and a high-quality two-way communication becomes impossible.

In order to suppress the nonlinear echo, nonlinear adaptive filters such as a Volterra filter and a neural network are used, which are theoretically capable of expressing any nonlinearity. However, they need more than ten times more computations than the linear echo canceller and still cannot offer satisfactory echo cancellation even with the huge computations.

4. Proposed Method[2)]

Fig. 3 shows the structure of the nonlinear echo canceller proposed in this paper. The nonlinear echo canceller first uses a linear echo canceller to suppress linear echo, it then applies FFT (Fast Fourier Transform) to convert the signal into a frequency spectrum, and then applies IFFT (Inverse Fast Fourier Transform) to re-synthesize the audio waveform, which is transmitted to the other party of the communication. The frequency spectrum processing is based on the newly discovered “correlation in spectrum amplitude between the nonlinear echo and the linear echo replica.”

Fig. 4 shows an example of spectrum amplitude correlation. The linear echo and nonlinear echo have completely different waveforms. However, as a result of surveys of qute a few types of compact equipment, it was found that their spectrum ampli-

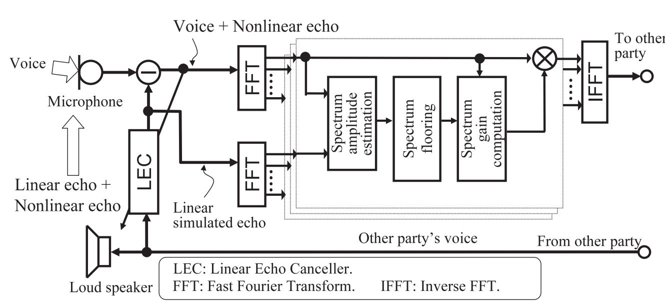

Fig. 3 Structure of the Nonlinear Echo Canceller.

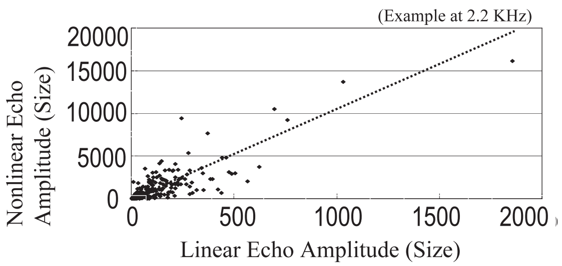

Fig. 4 Correlation of nonlinear and linear simulated echo spectrum amplitudes.

tudes at corresponding frequencies are correlated, that is, the spectrum amplitude of a frequency in the nonlinear echo increases when that of the corresponding frequency in the linear echo replica increases as shown in Fig. 4. The discovery of the correlation has made it possible to estimate the amount of nonlinear echo efficiently. In the spectrum estimation block shown in Fig. 3, the spectrum amplitude of the desired audio signal can be estimated based on the estimated nonlinear echo amplitude.

As this correlation is not precise, the spectrum amplitudes of the estimated audio signal may contain errors. We succeeded in implementing nonlinear echo cancellation at a practical quality by developing methods to conceal the estimation errors for ears. If the estimated spectrum amplitudes of an audio signal are excessively low and lower than those of the background noise, the signal level varies depending on whether or not echo is present, resulting in unnatural sounds to the ear. To prevent variation in the signal level, the spectrum flooring block in Fig. 3 estimates the background noise level and sets it as the lower limit of the estimated spectrum amplitudes. On the other hand, if a large amount of echo remains in the estimated spectrum amplitude due to estimation errors, the residual echo varies intermittently and suddenly so that an artificial sound called musical noise is heard. To prevent this by suppressing the echo, the gain computation block shown in Fig. 3 multiplies the gain so that the amplitude of the estimated nonlinear echo becomes appropriate instead of subtracting it. The intermittent variation of the residual echo can be reduced by smoothing for inhibiting sudden changes in the gain.

The accumulation of the above technologies has enabled quality two-way communications even with compact equipment such as notebook PCs and cellular phones. The nonlinear echo canceller can be implemented by simple computations such as constant number multiplications and smoothing operations. It can be implemented by a small number of additional computations between 20% and 100% compared to the computations for the linear echo canceller.[3)]

5. Evaluation with Simulation

Fig. 5 shows the results of evaluation with a cellular phone mockup, which is the most compact equipment with a small loudspeaker. Fig. 5(a) is the audio signal from the other party that is reproduced at the loudspeaker, and Fig. 5(b) is the signal waveform of the echo captured by the microphone. In Fig. 5(c), the echo is canceled using only the linear echo canceller. However, the nonlinear echo is still remaining and the sound of the spoken voice is still audible. On the other hand, with the output waveform of the proposed method, shown in Fig. 5(d), any echo including the nonlinear echo is cancelled to an inaudible level.

6. Conclusion

A new nonlinear echo suppression technology is presented, which enables compact equipment such as a notebook PC or cellular phone to support high-quality handsfree talk. We believe that the proposed method contributes to high-quality handsfree communications using a variety of equipment including PCs, softphones, VoIP handsets and cellular phones. We are planning to utilize the technology in commercial products in the near future.

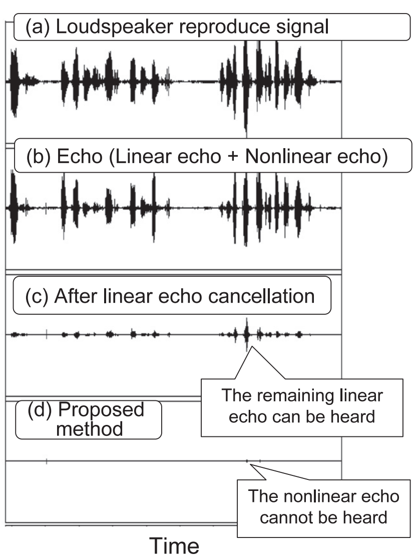

Fig. 5 Example of evaluation with simulation.

References

1) A. N. Birkett and R. A. Goubran: "Limitations of Hands-free Acoustic Echo Cancellers Due to Nonlinear Loudspeaker Distortion and Enclosure Vibration Effects," IEEE Proc. WASPAA'95, pp. 13-16, 1995.
2) O. Hoshuyama and A. Sugiyama: "An Acoustic Echo Suppressor Based on a Frequency-Domain Model of Highly Nonlinear Residual Echo," IEEE Proc. ICASSP2006, pp. V-269-272, 2006.
3) O. Hoshuyama and A. Sugiyama: "Evaluations of an Echo Suppressor Based on a Frequency-Domain Model of Highly Nonlinear Residual Echo," Proc. IWAENC'06, A54.

Author's Profile

HOSHUYAMA Osamu
Principal Researcher,
Common Platform Software Research Laboratories,
NEC Corporation
Member of IEEE, IEICE (Institute of Electronics, Information and Communication Engineers) and ASJ (Acoustic Society of Japan).