J. Acoust. Soc. Jpn. (*E*) **10**, 1 (1989)

Acoustic echo canceller with multiple echo paths

Nobuo Koizumi,* Shoji Makino,* and Hiroshi Oikawa**

**NTT Human Interface Laboratories,*
3–9–11, Midori-cho, Musashino, Tokyo, 180 Japan
***NTT Customer Equipment Division,*
Yamato Seimei Building, 1–1–7, Uchisaiwai-cho, Chiyoda-ku, Tokyo, 100 Japan

(*Received 31 May 1988*)

A new configuration of acoustic echo canceller for multiple microphone teleconferencing systems is proposed. It is designed for use with microphones whose gains switch or vary during teleconferencing according to the talker. This system requires memory for multiple echo paths, which enables the updating of filter coefficients when an echo path is changed due to the switching of the actuated microphone during talker alternation. In comparison to the single echo path model which uses only adaptation, this method maintains echo cancellation during abrupt changes of the echo path when the microphone alternates between talkers. Also in comparison to direct microphone output mixing, this method reduces the stationary residual echo level by the reduction of acoustic coupling.

PACS number: 43. 55. Lb, 43. 88. Md, 43. 88. Si

1. INTRODUCTION

In teleconferencing, hands-free telephony through multiple microphones is required to allow the participation of many people in each room. In such a system, acoustic feedback compensation and room echo suppression are necessary for natural two-way conversations.

To achieve this, the voice switching method, which controls the send and receive signal levels, is commonly used. Although this method is reliable, it causes the chopping of speech when the send/receive status alternates. An echo canceller can eliminate this problem, and the recent implementation of digital signal processing developments makes use of this method for treating long delay acoustic feedback practical. In the conventional configuration, however, this method is applied to a single echo path model. This means that it contains only a single echo path filter to compensate for the total acoustic feedback. When many microphones are used, direct mixing of multiple microphone outputs decreases the signal-to-noise ratio, and degrades speech quality. To eliminate this problem, individually controlled microphone gains are desirable. This requires the use of multiple acoustic echo paths, as stable operation cannot be expected by a single echo path configuration.

This paper proposes a configuration which achieves stable operation during the gain switching of multiple microphones, and the transient and stationary performance of this configuration is investigated.

2. DESIGN OF THE MULTIPLE MICROPHONE SYSTEM

Examples of the microphone arrangement for teleconferencing with many participants at each location is shown in Fig. 1. For best speech quality, the microphone should be set nearest the talker to keep the direct voice signal level higher than the reverberant portion of the signal. Therefore microphones should be set up for each participant, or at least one for every two participants. If the outputs of these microphones are simply mixed into a send signal, the signal-to-noise ratio will be reduced by the ambient room noise, and distortion will

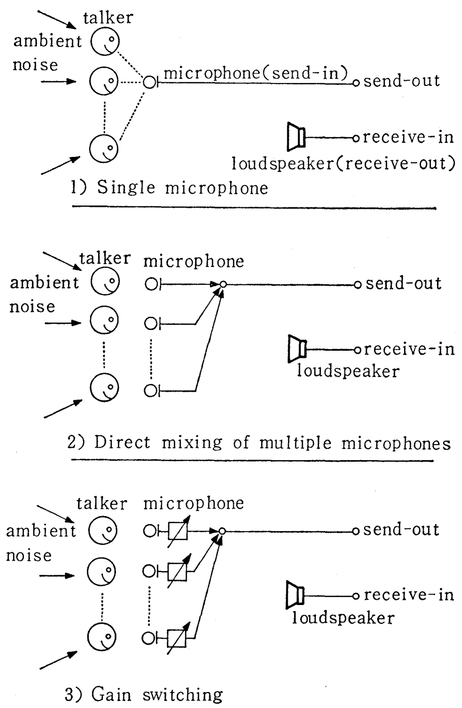

Fig. 1 Arrangement of microphone systems.

occur when the reverberant portions are added. As the returned echo path signal level and the ambient noise signal level increase, acoustic coupling increases and the system will become unstable. Degradation of the S/N or increase of acoustic coupling is estimated as $10 \log_{10} N$(dB), where N is the number of the connected microphones. Also, it is known that the speech transmission index, the objective measure of speech signal intelligibility, will be reduced by approximately 0.1 when six microphones are connected in a room with 1 s reverberation time.[1] However, with regard to microphone usage, usually there is only one talker and occasionally two or more when another talker interrupts.

Considering these parameters, the best speech quality can be attained by actuating only the microphone nearest the talker and attenuating the rest. When the talker changes, microphone gains must also change to maintain a constant total signal level. Also, we should avoid the chopping of the voice signals. This means that an attenuated microphone gain should maintain some value.[2] Though this switching of microphone gains will cause a decrease of individual talker's speech signal level when two or more participants speak simultaneously, it is acceptable, since in a real situation, this status will not last for long.

Therefore, let the gain for an attenuated microphone be normalized to unity. Represent the gain of the actuated microphones by a_k', where k is the number of actuated microphones, and then assume that their gains have equal values according to each k. When the total mixed signal level is kept constant, the relation between the gain for when a single microphone is actuated and for when k microphones are actuated is

$$a_1'^2+(N-1)=ka_k'^2+(N-k)$$

or

$$a_k'=\sqrt{\frac{a_1'^2+k-1}{k}} \tag{1}$$

where we assume that there are no correlations between microphone signals. We consider this the ideal gain switching condition.

3. CONFIGURATION OF THE ECHO CANCELLER

An echo canceller is used to eliminate the portion of the send signal which comes from a loudspeaker as a returned echo. It is constructed of an echo path filter which simulates the total of all acoustic echo paths from the loudspeaker to each microphone. However, when a conventional echo canceller that is unable to memorize several sets of filter coefficients is applied to a switched gain microphone system, the amount of cancellation abruptly decreases when the microphone gain switches, and recovery is delayed until the convergence of the adaptation. Therefore, a system is proposed with sets of acoustic echo path transfer functions contained in memory. The filter coefficients of the total acoustic echo path are calculated by a linear combination of the values stored in the memory. Filter coefficients are changed when talkers alternate and microphone gain changes. Because this gain switching is distinguished from environmental variations in the acoustic echo path, cancellation degradation due to change of speaker can be avoided. The system configuration is shown in Fig. 2. Though it is assumed that the number of echo path sets is equal to the number of microphones, it is not necessary that the impulse response stored in a memory is that of the echo path to a single microphone.

In this configuration, when the gain of the

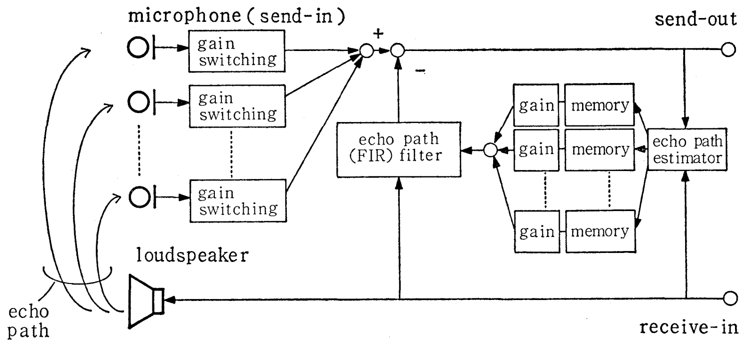

Fig. 2 Configuration of echo canceller for multiple echo paths.

attenuated microphones is set to zero, the total echo path can be formulated as a linear combination of each microphone's echo path for arbitrary gain settings. However, when attenuated microphones have gain values such as those found in practical situations, some constraints are required to obtain the total echo path by a linear combination.

Consider the impulse response of the acoustic echo path from the loudspeaker to the j-th microphone, H_j. When the system is operating, H_j cannot be observed or estimated because attenuated microphones still keep some gain. The total impulse response from the loudspeaker to the mixed output of the microphone system when k microphones are actuated can be described as

$$H=(a_k-1)\sum_{j:\mathrm{on}}^{k} H_j+\sum_{i=1}^{N} H_i \tag{2}$$

where attenuated microphone gain is normalized to unity and a_k indicates the actuated microphone gain when k microphones are actuated. In this equation, the first summation on the right side represents the summation of impulse responses for k actuated microphones. When only a single j-th microphone is actuated, total impulse response, which is indicated as H^*_j, becomes

$$H^*_j=(a_1-1)H_j+\sum_{i=1}^{N} H_i \tag{3}$$

If we take the summation for all actuated microphones, we have

$$\sum_{j:\mathrm{on}}^{k} H^*_j=(a_1-1)\sum_{j:\mathrm{on}}^{k} H_j+k\sum_{i=1}^{N} H_i \tag{4}$$

Therefore, if the relation between gains a_1 and a_k is given as

$$a_k-1=\frac{a_1-1}{k} \tag{5}$$

the total impulse response for when an arbitrary number of microphones is actuated can be obtained by the following equation.

$$H=\frac{1}{k}\sum_{j:\mathrm{on}}^{k} H^*_j \tag{6}$$

If only single microphone is actuated when there is no near-end talker, the impulse response H^*_j can be identified during operation. Therefore, if they are stored in memory, H can be estimated on demand.

The gain a_k given by Eq. (5) does not coincide with the ideal constant level gain a_k' given by Eq. (1). However, the total mixed signal level can be interpreted as a constant if the difference is small. The values of Eq. (1) and Eq. (5) are compared in Fig. 3. When single actuated gain a_1 is set to 2 (6 dB) or 3 (9.5 dB), the difference between a_k and a_k' is within 1 dB. This implies that using Eq. (5) as the gain setting is acceptable if a_1 is appropriately given, and acoustic coupling can be maintained approximately

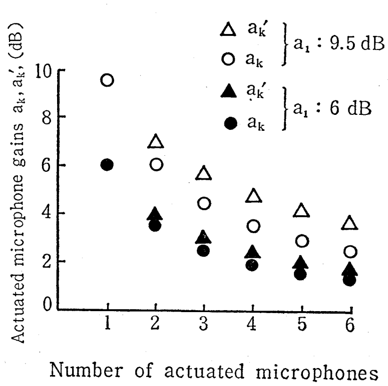

Fig. 3 Values of actuated microphone gain.

constant. Therefore we adopt a_k defined by Eq. (5) as the gain setting of this proposed configuration.

The block diagram of the operation is shown in Fig. 4. At the startup, microphone gains are switched serially to actuate each microphone. Using a test signal, initial values of H^*_j $(j=1, \ldots, N)$ are identified and stored in memory. When the conferencing starts, one stored response is called up as the set of filter coefficients for a single talker, and for multiple talkers, filter coefficients are calculated from stored responses. When speech stops, the impulse response is adaptively identified using the speech signals of the far-end talker at the opposite site. In this case, the microphone of the last talker remains actuated. The adaptation will work well, and more adaptation time will be assigned to frequently actuated microphones. Because this scheme only requires the addition of memory to the hardware configuration and a simple calculation scheme, it can be easily implemented.

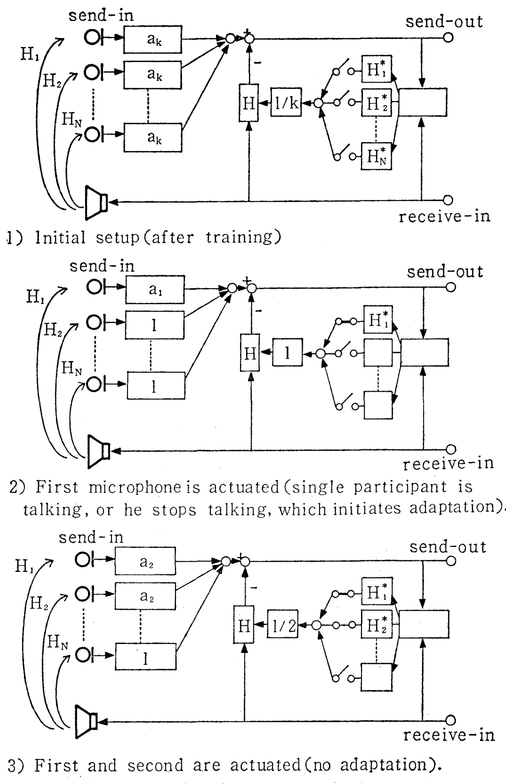

Fig. 4 Operational block diagram.

4. SIMULATION STUDY

The convergence of echo cancellation is simulated using real acoustic echo path data. For the measurement of impulse response data, 3 microphones are arrayed in a line 50 cm apart as shown in Fig. 5, in a room with a reverberation time of 0.46 s. The stored impulse response word length is 4,000 samples, which can deal with 0.5 s data at an 8 kHz sampling rate.

The change of echo cancellation due to microphone switching using the conventional single echo path model is shown in Fig. 6. The horizontal axis is the iteration, and the microphone switching occurs at a constant rate. The first six switches are by single microphone and last three are by two actuated microphones. The echo cancellation is defined as 10 log [(power of returned signals)/(power of the residual signals)]. For the adaptation algorithm, the adaptive learning method[3] is used and white noise is applied as the test signal. The figure shows that cancellation decreases abruptly when the distance from the loudspeaker to the microphone varies only slightly. Though the convergence performance depends on the adaptation algorithm, the convergence will deteriorate when a speech signal is used for adaptation during real use.

This kind of degradation can be avoided if the proposed configuration is applied with the same adaptation algorithm, and only the time period for the convergence of the variation of the room acoustics is required. This feature is shown in Fig. 7. In this figure, the horizontal axis represents

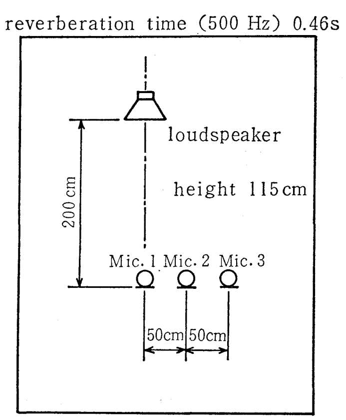

Fig. 5 Location for impulse response measurement.

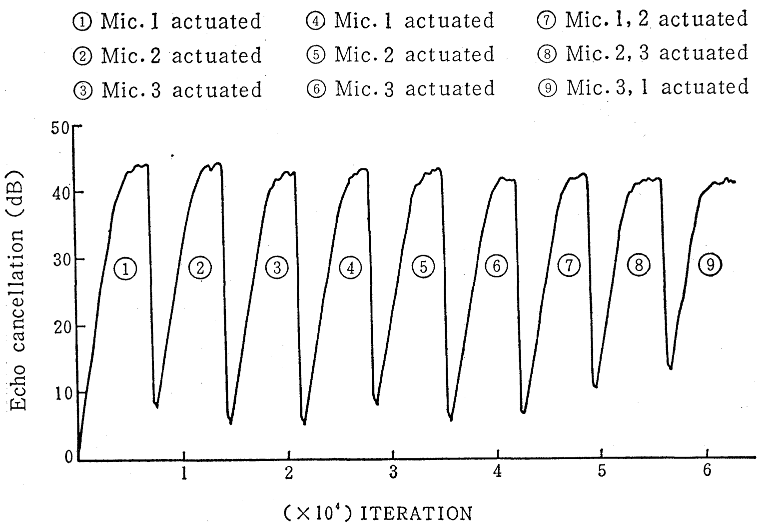

Fig. 6 Variation of echo cancellation.

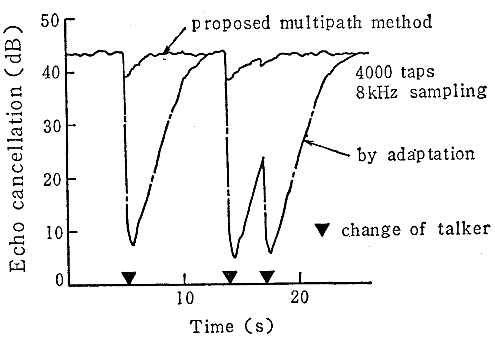

Fig. 7 Variation of echo cancellation over time.

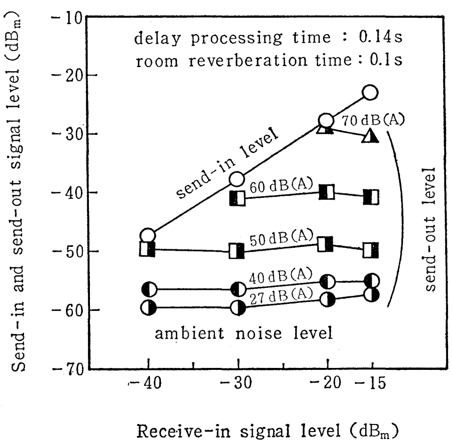

Fig. 8 Send-in and send-out signal level of echo canceller in room environment.

time. In a real environment, after a talker or actuated microphone changes, the adaptation stops and restarts after the talker stops talking. However, to show only the difference of the convergence, in the figure, this interval is abridged, and adaptation is shown to start immediately after the alternation. Even though the conventional adaptation method greatly reduces cancellation when talker alternates frequently, the proposed method maintains a stable cancellation performance during talker alternation.

5. STATIONARY PERFORMANCE

If it is assumed that the delay processing time of the echo path filter, which is determined by the order of finite impulse response filter, sufficiently exceeds reverberation decay time, the cancellation performance in a real room's acoustics depends on ambient noise level. Here, some stationary characteristics from the experimental results of a single microphone system are considered. These results were obtained from equipment which used digital signal processor chips.[4] The set has a passband of 4 kHz and a delay processing time of 144 ms. It is designed for a conventional single echo path configuration. The room used for this experiment

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**

fastcase®
Smarter legal research.