# DUAL-MICROPHONE ECHO CANCELLER FOR SUPPRESSING LOUD NONLINEAR ECHO

*Osamu HOSHUYAMA*

Information and Media Processing Laboratories, NEC Corporation

## ABSTRACT

This paper proposes a new dual microphone echo canceller (DMEC) for suppressing loud nonlinear echo common with hands-free talk on small terminals such as cellphones. The proposed DMEC has an adaptive null former whose nulls by the dual microphones focus on cancelling the nonlinear echo which can not be cancelled by ordinary linear echo canceller. To focus on the nonlinear echo, an adaptive filter eliminates the linear echo components which may perturb the adaptive null former from the second microphone signal before the adaptive null former. The following linear echo canceller and echo suppressor eliminate the linear echo and the residual nonlinear echo in the first microphone signal. Nearend speech degradation by the echo suppressor is reduced because the nonlinear echo is already reduced by the adaptive null former. Evaluations with real cellphones demonstrate that the proposed DMEC can cancel out the loud nonlinear echo almost completely and obtain stable nearend speech quality.

***Index Terms*—** Echo Cancellation, Nonlinear Echo, Dual Microphone, Microphone Arrays, Beamforming

## 1. INTRODUCTION

Hands-free phone is a basic and essential application with small information terminals such as cellphones, smart phones, and tablet PCs. For hands-free communication, echo cancellation is common but still a difficult function. An echo canceller has an adaptive filter to emulate the echo path between the input of amplifier to drive a loudspeaker and the microphone. Even with echo cancellers, suppression of echo is very difficult because the loudspeaker is small and close to the microphone, and the sound from the loudspeaker is very loud [1]. The echo at the microphone is larger than the nearend speech which is the target signal to be sent to the farend. Especially, nonlinear echo generated by loud sound from small loudspeakers and parts' vibration in the echo path can not be cancelled out by only a linear adaptive filter. To cancel the nonlinear echo, nonlinear adaptive filters such as Volterra filters and neural filters have been proposed, however, they cannot suppress nonlinear echo efficiently, despite their large computational complexity. These nonlinear adaptive filters are rarely used in commercial products.

To suppress the nonlinear echo subjectively, echo suppressors (post filters) are widely used in actual products. They are based on multiplying operations such as gain control in frequency domain [2, 3, 4, 5]. In echo suppressors, degradation of the nearend voice is inevitable when suppressing loud nonlinear echo. There are a lot of techniques to reduce the degradation [4, 6, 7, 8], however, they have limitations for very loud nonlinear echo. When the amplitude of the nonlinear echo is comparable to that of the nearend speech, the nearend speech is chopped and becomes unintelligible.

Introducing null forming (beamforming) using an additional microphone [9, 10] to suppress the nonlinear echo is a natural idea. However, degrees of freedom to form nulls by the dual microphones are limited. To suppress the nonlinear echo through complicated echo paths, the nulls should be assigned efficiently. Simple null forming tend to use the null to suppress only the linear echo, because the linear echo is quite larger than the nonlinear echo. The nonlinear echo still remains resulting in serious nearend speech degradation by the echo suppressor. Another problem with the simple null former is attenuation of the nearend speech along with the echo.

This paper proposes a new dual microphone echo canceller structure to suppress loud echo. Target signal-to-echo ratio is -10 dB or less. The new structure has 3 adaptive filters and appropriate control for efficient null assignment to concentrate on suppressing loud nonlinear echo. The smaller residual nonlinear echo leads to less nearend speech degradation by the following echo suppressor.

## 2. PROBLEM ANALYSIS

In echo cancellation, the quality of the output signal sent to the farend is discussed from the view points of the amount of the residual echo and the quality of the nearend speech. Ambient noise is out of focus in this paper, because there are a lot of good solutions. Figure 1 illustrates a typical structure of a nonlinear echo canceller (NLEC) including an adaptive filter (AF) and an echo suppressor (ES) [2, 4, 5]. Farend speech is emitted by loudspeaker. Linear and nonlinear echo generated from the loudspeaker sound, and nearend speech are picked up by microphone 1. Adaptive filter 1 (AF1) is a classical and essential part of echo canceller, which estimates the echo path and generates echo replica to cancel linear echo. In this paper, nonlinear echo means the residual echo components which can not be cancelled out by the classical linear echo canceller. These nonlinear echo components are generated by various nonlinearities such as distortion in the amplifier to drive loudspeaker, the loudspeaker itself, and distorted vibration in the echo paths between the loudspeaker and the microphone.

Most of echo suppressors for the residual nonlinear echo use gain reduction processing in frequency domain based on rough estimation of echo and nearend speech. They uses state-of-the-art techniques to obtain acceptable speech qual-

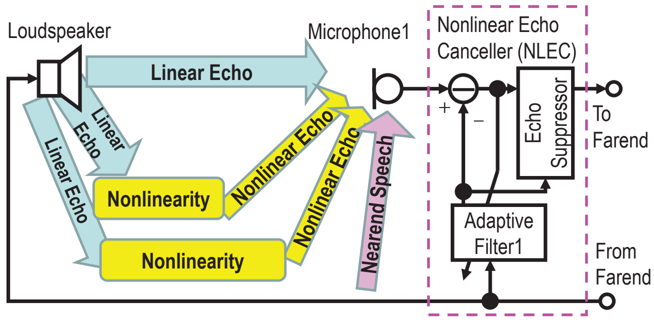


**Fig. 1**. Single Microphone Nonlinear Echo Canceller.

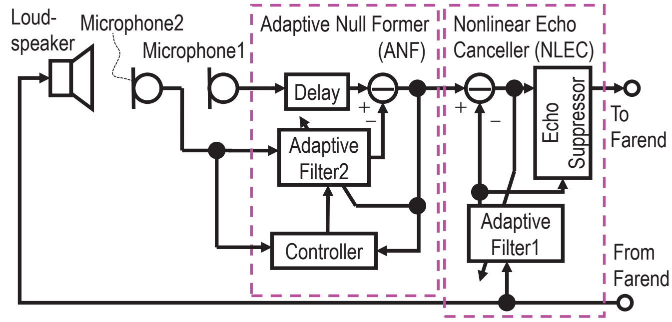


**Fig. 2**. Simple Dual Microphone Echo Canceller.

ity, however, when the amplitude of nonlinear echo is comparable to or larger than the nearend speech, degradation of the nearend speech is terrible, and sometimes the speech becomes unintelligible.

## 3. DUAL MICROPHONE APPROACH TO ECHO CANCELLATION

Dual microphone null forming is a simple microphone array proposed for noise reduction [9, 10]. Introducing this technique for echo cancellation is natural. A dual microphone echo canceller (DMEC) with a simple adaptive null former (ANF) is shown in Fig 2. Additional adaptive filter 2 (AF2) following the additional microphone-2 minimizes the output of subtractor. Directivity (spatial nulls) formed by AF2 reduces the echo. Influences of the directivity on the nearend speech should also be considered as well as the echo.

In this approach, performance is limited by the characteristics of nulls, therefore, microphone positions which dominate the null characteristics is an important issue. In our experiments, endfire arrays, where microphone-2 is located closer to the loudspeaker than microphone-1, is better to avoid nearend speech attenuation than broadside arrays, where microphone-2 is located at the same distance as microphone-1. The nulls formed by a broadside array tend to be wide, and easily attenuate the nearend speech with the echo. In this paper, endfire arrays are assumed.

Performances of DMECs also depend on the type of nonlinear echo. When there is one dominant nonlinearity (e.g. distortion with amplifier or loudspeaker), dual microphone approach can be effective. If the dominant nonlinear echo is picked up by both the microphones, cancellation of the nonlinear echo is easy for the nulls by the adaptive null former.

When the sources of nonlinearity are distributed and there is no dominant nonlinearity, the performance improvement by DMEC is not so high as expected. Most of the nonlinear echoes belong to this category, because nonlinear echo propagation paths are complicated with light-body small terminals. The degrees of freedom by the dual microphones are not enough to assign the nulls to so many distributed nonlinearities. In this case, maximization of null effect is an important issue.

The simple DMEC is effective for echo suppression, however, not so efficient for nonlinear echo suppression. When the linear echo is larger than the nonlinear echo and both are in the microphone-2 signal, minimizing the output power of the adaptive null former means minimization of linear echo. The simple DMEC can efficiently suppress the nonlinear echo only when the nonlinear echo paths are similar to the linear echo paths, which can not be expected in most cases.

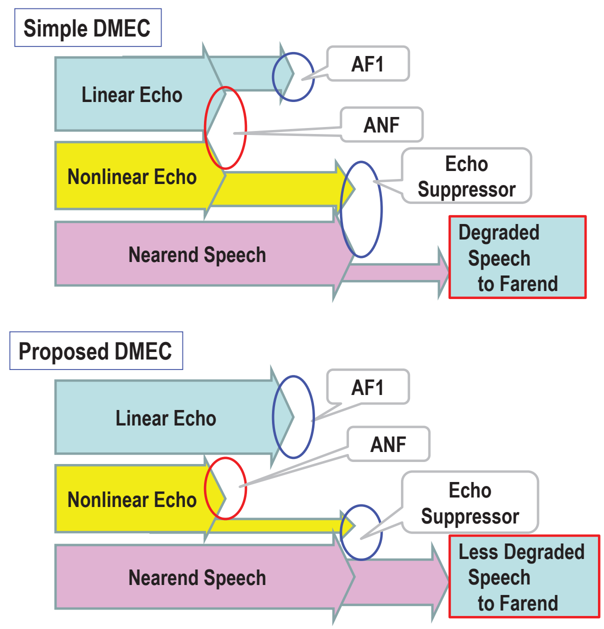


**Fig. 3**. Comparison of Simple Dual Microphone Echo Canceller and Proposed Dual Microphone Echo Canceller from View Point of Signal Suppression.

Another artifact with the simple adaptive null former is degradation of the nearend speech. Explicit directional nulls formed by the simple adaptive null former to suppress directional propagation of the linear echo may cause serious degradation of the nearend signal arrived from the null direction.

## 4. PROPOSED DUAL MICROPHONE ECHO CANCELLER

Figure 3 illustrates the idea of the proposed DMEC in comparison with the simple DMEC from a view point of signal suppression. The width of an arrow indicates the amount of the corresponding signal, and an oval means signal suppres-

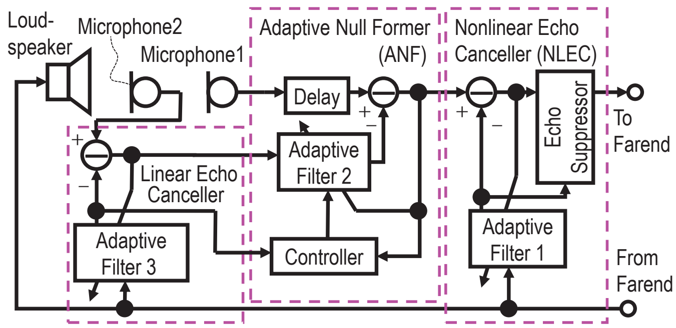


**Fig. 4**. Proposed Dual Microphone Echo Canceller.

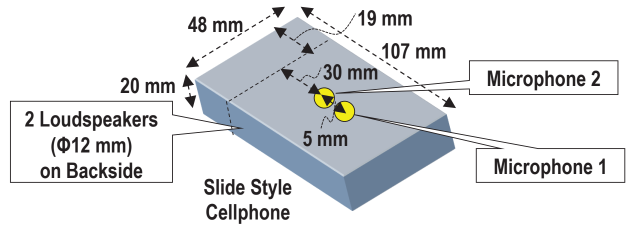


**Fig. 5**. Loudspeaker and Microphone Arrangment.

sion. In the proposed DMEC, the nulls by a new adaptive null former (ANF) concentrate on nonlinear echo suppression, and almost all the linear echo is cancelled by the later traditional echo canceller (AF1). Even when the nonlinear echo suppression by the ANF is not so high, concentration on the nonlinear echo suppression is still more efficient than the simple DMEC. For the echo suppressor, when the nonlinear echo is comparable to the nearend speech, complete suppression of the nonlinear echo results in terrible degradation of the nearend speech. Even a little nonlinear echo reduction by the adaptive null former is very useful to reduce the degradation of nearend speech.

Structure of the proposed DMEC is shown in Fig. 4. In order to make the adaptive null former concentrate on the nonlinear echo suppression, linear echo components in the microphone-2 signal is eliminated and the nonlinear echo components are extracted by adaptive filter 3 (AF3), which works like an ordinary linear echo canceller. The nulls formed by the ANF with AF2 are assigned to the nonlinear echo because the input signal of AF2 is mainly nonlinear echo. The ANF is basically an adaptive noise canceller in which the nonlinear echo components in the microphone-1 signal correlated to the nonlinear echo components extracted by AF3.

Control of adaptation is also an important factor with the proposed DMEC. If adaptation of AF2 is performed during speech-only period, nulls are formed to the target nearend speech, which means serious degradation. To assign the nulls to the nonlinear echo sources not to the target nearend speech, the adaptation of AF2 is performed only when the nonlinear echo is loud, which is detected based on the power of the linear echo estimated by AF3. Other conservative control techniques are used to avoid performance degradation in cascaded adaptive filters. Tracking speed of AF2 and AF3 is set to be slower than AF1.

The structure of proposed DMEC is similar to microphone array structure proposed by Kellermann for combining echo canceller and adaptive beamformer [11]. The difference from the proposed DMEC is that most of linear echo is already cancelled out before null forming in [11], while the linear echo is suppressed after null forming in the proposed DMEC.

### 4.1. Characteristic of Nulls

Characteristics of the nulls formed by the adaptive null former in the proposed DMEC contributes to stability of the nearend speech quality. When there is a dominant nonlinear echo and it is picked up by both of the microphones, nulls are assigned efficiently to suppress the nonlinear echo, which is obvious.

When the sources of nonlinearity are distributed, the nulls formed to minimize the nonlinear echo are distributed to various direction and frequencies. The nulls can not show explicit spatial nulls or directivity, which means no serious attenuation to any directions. Therefore, the quality of the nearend speech is more stable with the proposed DMEC than the simple DMEC as far as the control of adaptation is appropriate.

## 5. EVALUATIONS

The proposed DMEC was evaluated with speech data recorded using real cellphone bodies with additional microphones on the bodies. Loudness and distortion of the echo were set so that total amplitude of the echo is over 10 dB larger than the nearend speech, and that the nonlinear echo, which was measured as the residual echo after a linear echo canceller, was comparable to the nearend speech amplitude. For comparison, a single microphone nonlinear echo canceller (NLEC) [4] and the simple DMEC in Fig. 2 were also evaluated.

The sampling frequency was 8 kHz. The numbers of filter taps with adaptive filters were 256 for AF1, 128 for AF2, and 128 for AF3, and the number of delays in adaptive null formers was 10 samples. For appropriate control, all the adaptive filters employed noise robust structure based on slave filter[12]. Parameters for the echo suppressor were set so that the nonlinear echo is suppressed almost completely.

Over 10 microphone arrangements were tested. Of course, when there was a dominant nonlinear echo, the DMECs worked well. The proposed DMEC was superior when the echo is very loud. In other cases, improvement by the proposed DMEC was not significant. In this paper, one of the most difficult conditions shown in Fig. 5 is presented. Figure 6 shows the output waveforms. At microphone-1, echo was 12 dB larger than the nearend speech as shown in Fig 6(a). The nonlinear echo was large and comparable to the amplitude of the nearend speech. By using the echo suppressor, the nonlinear echo can be suppressed as in Fig. 6(b), however, the nearend speech, which is the target signal, was also suppressed and sometimes unintelligible.

The simple DMEC cancelled mainly the linear echo with a nearend speech attenuation of as much as 16 dB. After am-

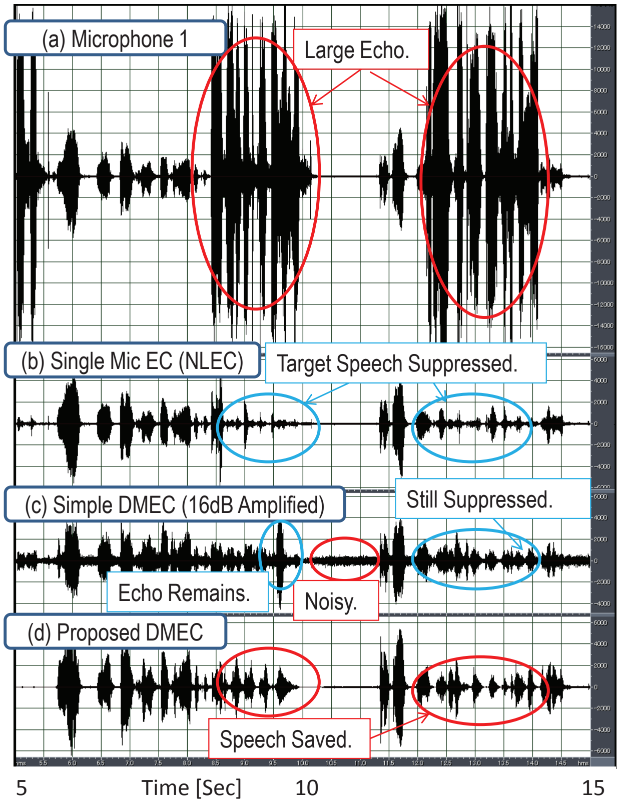


**Fig. 6**. Output Waveforms of Microphone-1, NLEC, Simple DMEC, and Proposed DMEC.

plification to compensate for the attenuation, the signal was quite noisy as illustrated in Fig. 6(c). The adaptive null former in the simple DMEC cancelled the linear echo by over 10 dB, but not the nonlinear echo. Therefore, the echo suppressor have to strongly suppress the nonlinear echo, resulting in serious nearend signal degradation. For different microphone arrangements, the nearend speech attenuation varied from 0 dB to -20 dB.

Output waveform of the proposed DMEC in Fig. 6(d) saves the target nearend speech thanks to the new adaptive null former. Actually, nonlinear echo reduction by the adaptive null former was just 2 or 4 dB because the sources of nonlinearity were distributed, however, its contribution to the nearend speech quality is quite evident.

For other microphone arrangements, as far as the microphones and the loudspeaker locate linearly forming an endfire array, quality of the nearend speech at the final output was stable. The attenuation of the nearend speech was -3 dB in the worst case.

Influence by the nearend talker position on the nearend speech attenuation was also evaluated. A cellphone was put on a desk and a signal source loudspeaker was moved around the cellphone body at a distance of 40 cm. Variation of the nearend speech attenuation was 10 dB for the simple DMEC, and 3 dB for the proposed DMEC. These results indicate that the adaptive null former in the simple DMEC formed explicit directional nulls to suppress the nonlinear echo components, while the adaptive null former in the proposed DMEC formed no explicit directional null and avoided serious attenuation of the nearend speech.

## 6. CONCLUSIONS

This paper has proposed a new DMEC structure for suppressing loud nonlinear echo common with hands-free talk on small terminals. The proposed DMEC has an adaptive null former whose nulls by the dual microphones focus on cancelling the nonlinear echo which can not be cancelled out by ordinary linear echo canceller. To focus on the nonlinear echo, an adaptive filter extracts the nonlinear echo in the microphone-2 signal before the adaptive null former. Nearend speech degradation by the following echo suppressor is reduced thanks to the adaptive null former reducing the nonlinear echo. Evaluations with real cellphones have demonstrated that the proposed DMEC can suppress almost completely loud nonlinear echos with power comparable to the nearend speech, while avoiding serious degradation of the nearend speech.

## 7. REFERENCES


[1] M. E. Knappe, and R. A. Goubran, “Steady-State Performance Limitations of Full-Band Acoustic Echo Cancellers,” IEEE Proc. ICASSP’94, pp. IV-81–84, 1994.

[2] R. Martin, and S. Gustaffson, “The Echo Shaping Approach to Acoustic Echo Control," Speech Communication, pp. 181–190, 1996.

[3] C. Faller, and J. Chen, “Suppressing Acoustic Echo in a Spectral Envelope Space,” IEEE Trans. SAP, pp. 1048–1062, 2005.

[4] O. Hoshuyama, and A. Sugiyama, “Nonlinear Echo Canceller Based on Spectral Shaping,” in E. Hänsler, G. Schmidt eds, *Speech and Audio Processing in Adverse Environments*, Springer, 2008.

[5] 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.

[6] M. Berouti, R. Schwartz, and J. Makhoul, “Enhancement of Speech Corrupted by Acoustic Noise,” Proc. ICASSP’79, pp. 208–211, 1979.

[7] Y. Ephraim, and D. Malah, “Speech Enhancement Using a Minimum Mean-Square Error Short-Time Spectral Amplitude Estimator,” IEEE Trans. ASSP, No. 6, pp. 1109–1121, 1984.

[8] R. Martin, “Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics,” IEEE Trans. SAP, pp. 504–512, 2001.

[9] J.F. Chen, K. Phua, L. Shue, and H. Sun, “Performance Evaluation of Adaptive Dual Microphone Systems,” Speech Communication, pp. 1180–1193, 2009.

[10] K. Kim, R. H. Baran, and H. Ko, “ Extension of Two-channel Transfer Function Based Generalized Sidelobe Canceller for Dealing with Both Background and Point-Source Noise,” Speech Communication, pp. 521–533, 2009.

[11] W. Kellermann, “Acoustic Echo Cancellation for Beamforming Microphone Arrays," in M. S. Brandstein and D. Ward, eds. *Microphone Arrays*, 2001, Springer.

[12] O. Hoshuyama, “An Echo Canceller Using Smoothed-Coefficient Filter with Adaptive Time Constant Controlled by High-Pass Errors,” Proc. IWAENC2008, 2008.