2nd European Conference on
Speech Communication and Technology
EUROSPEECH '91
Genova, Italy, September 24-26, 1991

Acoustic Echo Cancellation Using Prediction Residual Signals

C. Acker P. Vary H. Ostendarp

Institute for Communication Equipment and Data Processing
Technical University of Aachen
Templergraben 55, 5100 Aachen, Germany

Abstract

Adaptive echo cancellers are currently being studied for application to hands-free telephone sets with high speech quality. The major problem is the acoustic echo control from the loudspeaker to the microphone.
This proposal presents a new concept by using linear prediction techniques in combination with a conventional adaptive echo canceller. Computer simulations show that the introduced algorithm improves significantly the performance compared to conventional methods.
In this paper the new structure is described. A comparison of the results with the state of the art indicates that the complexity can be reduced by one third.

1. Introduction

The principle of adaptive echo cancellation is the suppression of the echo by linear compensation (see Fig.1). The acoustic echo path represented by the impulse response g_k is simulated by the adaptive filter $c_k(i)$ of length N.

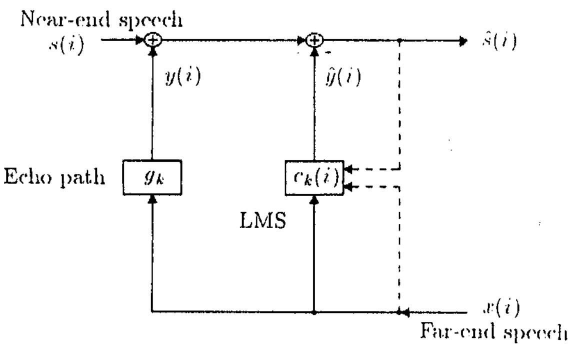


Fig. 1: Conventional echo canceller

The far-end input signal $x(i)$ is processed by an adaptive filter and subtracted from the signal on the return path, which includes the echo $y(i)$ and the near-end signal $s(i)$. The estimated value of $s(i)$ represented by

$$\begin{aligned} \hat{s}(i) &= x(i) * g_k - x(i) * c_k(i) + s(i) \\ &= x * g_k - x * c_k + s \end{aligned} \qquad (1)$$

where $0 \leq k < N$ and

$$\begin{aligned} \underline{g} &= (g_0, g_1, \dots g_k, \dots g_{N-1})^T \\ \underline{c}(i) &= (c_0(i), c_1(i), \dots c_k(i), \dots c_{N-1}(i))^T \\ \underline{x}(i) &= (x(i), x(i-1), \dots x(i-N+1))^T . \end{aligned}$$

The most common implementation of echo cancellers uses the LMS or least-mean-square algorithm.

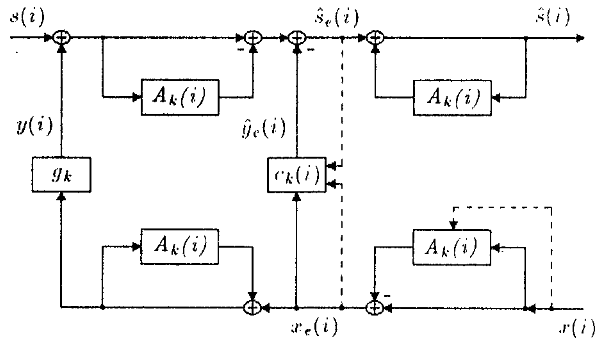


Fig. 2: First approach using prediction filters

One problem of echo compensation techniques (such as LMS or related algorithms) is that the convergence properties degrade with colored signal input such as speech signals [1]. To overcome this problem algorithms using linear prediction to decorrelate the speech signal have been proposed [2-4].

The basic idea was to perform the learning identification of the acoustic echo path with prediction residual signals instead of the speech signal itself. Fig. 2 illustrates the first approach of an echo canceller adapted with two residual signals $x_e(i)$ and

$$\hat{s}_e = x * g_k * a_k - x * a_k * c_k + s_e \qquad (2)$$

where $a_k(i)$ denotes the impulse response of the FIR prediction filter:

$$a_k(i) = \begin{cases} 1 & : \ k = 0 \\ -A_k(i) & : \ k > 0 \end{cases}$$

However, the usage of the prediction filters within the echo path increases the time dependency of the modified path for the compensation algorithm, because of the additional blockwise adaption of the prediction filters. The use of the residual signals causes an improved system distance

$$\frac{D(i)}{[dB]} = 10 \log \frac{||\underline{g} - \underline{c}(i)||^2}{||\underline{g}||^2} \qquad (3)$$

but the time dependency degrades the measure for the echo power cancellation – the echo return loss enhancement

$$\frac{ERLE(i)}{[dB]} = 10 \log \frac{E\{y^2(i)\}}{E\{(y(i) - \hat{y}(i))^2\}.} \qquad (4)$$

The developement of a new architecture has to refer to this problem.

2. Theoretical Limits

Proceeding from Fig. 2 a theoretical model can be constructed by removing the prediction filters out of the echo path. Instead of the speech signal $s(i)$ the residual signal $s_e(i)$ is used as near-end speech.

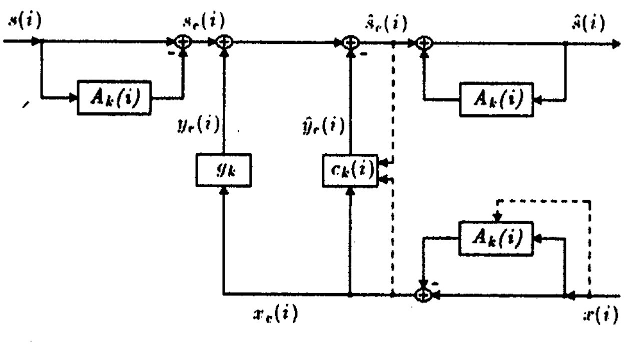


Fig. 3: Theoretical model

With these measures the echo canceller takes the decorrelated signals $x_e(i)$ and

$$\hat{s}_e = x * a_k * g_k - x * a_k * c_k + s_e \qquad (5)$$

for the identification of the echo path, while the time dependency of the path is only determined by the impulse response g_k.

Combining both properties in one algorithm provides an estimation of the theoretical limits concerning the convergence characteristics of acoustic echo cancellation using prediction residual signals. The ambition of the following steps is to approach the limits of this reference model as close as possible.

3. The New Echo Canceller

A comparison between equations (2) and (5) reflects to the main difference of the algorithms.
One solution was derived by [2,3] and [4]. Keeping the prediction coefficients constant for the length N of the compensation filter simulates a time invariant prediction filter. Under this condition the impulse responses g_k and $a_k(i)$ can be exchanged in equation (2).

Another possibility is to find a new structure where the adaptive variables $a_k(i)$ in equation (2) have a comparable neglectable influence to the time dependency as in equation (5). This concept is presented in the following.

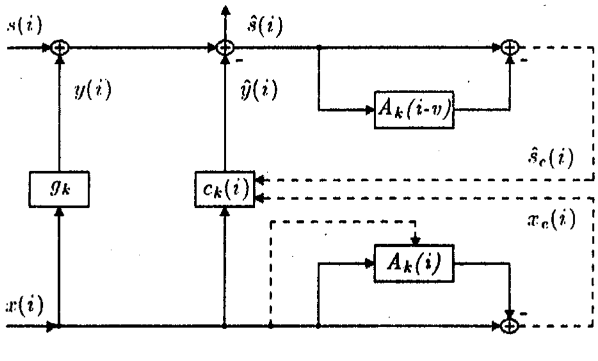


Fig. 4: Concept of the new echo canceller

The proposed structure of a new echo canceller was originally developed out of Fig. 2 by adding two prediction filters into the path of $c_k(i)$. Since preliminary experiments show that the architecture of Fig. 2 is capable of effectively increasing the system distance, the adaption algorithm with residual signals was maintained. With this modification the additional time dependency is internally considered within the compensation path and substracted from the signal on the return path $s(i) + y(i)$. By combining the prediction filters from the echo and compensation path a new structure is derived as shown in Fig. 4.

The new concept indicates directly that the time dependency of the modified algorithm is not increased. The adaptation of the echo compensator is performed using the decorrelated signals $x_e(i)$ and $\hat{s}_e(i)$:

$$\underline{c}(i+1) = \underline{c}(i) + \alpha(i)\hat{s}_e(i)\frac{\underline{x}_e(i)}{||\underline{x}_e(i)||^2} \qquad (6)$$

where x_e equals to the corresponding signals in Fig. 2 and Fig. 3. The signal $\hat{s}_e(i)$ is now determined by

$$\hat{s}_e = x * g_k * a_k - x * c_k * a_k + s_e. \qquad (7)$$

The echo replica $\hat{y}(i)$ is generated by processing the far-end input $x(i)$ through the echo path model $c_k(i)$:

$$\hat{y}(i) = \underline{c}(i)\underline{x}(i) \qquad (8)$$

To compensate the retardation caused by the cancellation filter $c_k(i)$ and the impulse response g_k, respectively, a delay v for the prediction filter of the transmitter side was introduced. Computational experiments show that an optimized delay v improves the performance of the echo canceller.

In contrast to this concept the identification of the echo path $c_k(i)$ and the generation of the echo replica $\hat{y}(i)$ are performed independently in the algorithm of Yamamoto [2,3] and Schultheiß [4]. So their algorithm requires two serparate filters $c_k(i)$ of length N, which increases significantly the complexity.

Furthermore, their algorithm has to fulfill the requirements of the time dependency. For that reason the prediction coefficients have to remain constant for N input samples $x(i)$. This refers to long blocklengths BL for the update of the prediction coefficients (corresponding to a low prediction gain) and to a refiltering of speech samples located outside of the current block (causing an immense computational peak load).

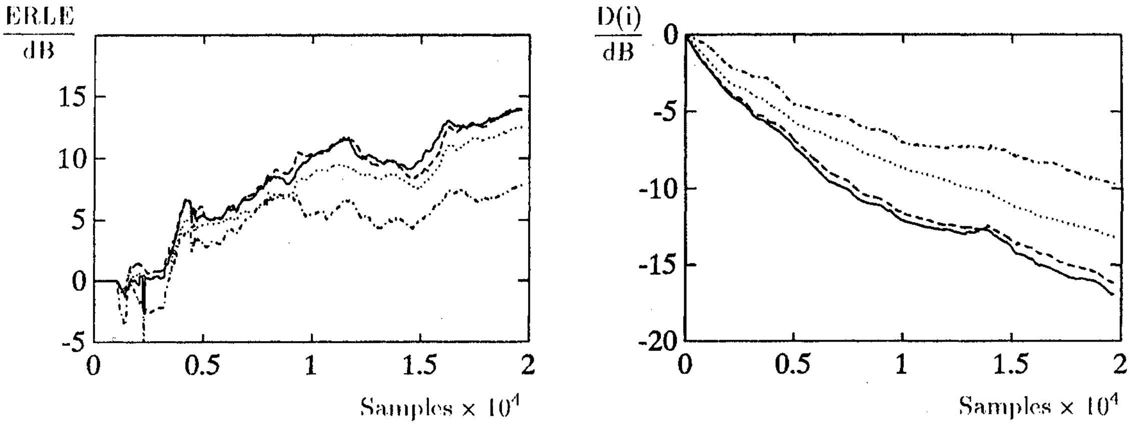
4. System Configuration

In order to verify the performance of the proposed echo canceller all concepts were examined by simulation under the following conditions:

- The LSI–system with a measured impulse response g_k of length $N = 512$ simulates the analog echo path.
- For the compensation a filter length N is used.
- A constant stepsize α represents a compromise between fast convergence and low steady state system distance. In order to improve both, a variable stepsize adaption according to the estimation of the system distance [4] was employed in all experiments.
- The predictor is of order $M = 8$, the coefficients are adapted every $BL = 160$ samples.
- The optimised delay is $v = 140$.
- Speech signals ($f_a = 8kHz$) were used as input signals, the levels of the echo and the near–end speaker adjusted to 0 dB.

5. Performance and Complexity

Computer simulations show that the Echo Return Loss Enhancement (ERLE) is increased by 4-5 dB in comparison to the conventional algorithm without linear prediction. Compared to the method in [2-4] only a slight degradation of the ERLE is detected (see Fig. 5).
Assuming a sufficient identification of the echo path at a certain time instant the new structure provides a convergence property which is close to the solutions proposed in [2-4], which is twice as fast as the conventional algorithm. However without this assumption the system distance of the new concept is still about 1.5 times faster than the conventional algorithm (see Fig. 5).
The main advantage of the new algorithm represents its low implementation complexity. Tab.1 reflects that the developed concept requires basically no major additional computation and memory compared to the conventional algorithm. In comparison to the solutions proposed in [2-4] a reduction of the complexity by one third is achieved.



—— Theoretical Model
- - - Model proposed in [2-4]
··· New Concept
-.-. Conventional Concept

Fig. 5: Echo return loss enhancement (ERLE) and system distance (D(i)) over an average of 7 speakers

Algorithm:	Operations per sample	Example	Memory	Example
Conv. Concept	$2N$	1024	$2N$	1024
Algorithm [2]	$3N + 4M + \frac{M^2-M}{N}$	1568	$3N + 3M$	1560
Algorithm [4]	$3N + 3M + \frac{M^2}{BL} + \frac{N-1}{BL}M$	1586	$3N + 3M$	1560
New Concept	$2N + 3M + \frac{M^2}{BL}$	1048	$2N + 4M$	1056

Tab. 1: Comparison of the complexity

6. Conclusions

A new algorithm is proposed for an acoustic echo canceller based on a linear prediction algorithm. Its performance and complexity are discussed in comparison with a conventional algorithm and the state of the art.
Computer simulations have shown that the ERLE can be increased by 4-5 dB compared to a conventional echo canceller. Furthermore the system identification is performed at least 1.5 times as fast, while the complexity does not increase. Thus comparing the new echo canceller with the model in [2-4] a slight degradation of the performance is detected but the complexity is reduced by one third.

Acknowledgements to Ms. C. Fraune who carried out parts of the simulations within her diploma thesis.

References

[1] B. Widrow et al., "Adaptive Noise Cancelling: Principles and Applications", *Proceedings of the IEEE*, Vol. 63, No. 12, Dec. 1975

[2] S. Yamamoto et al., "An Adaptive Echo Canceller with Linear Predictor", *Trans. of the IECE of Japan*, Vol. E 62, No. 12, Dec. 1979

[3] S. Yamamoto et al., "An Algorithm of Adaptive Echo Canceller with Linear Predictor Considering Effects of Circuit Noise", *Paper of Technical Group TGCS80 - 48, IECE Japan*, June 1980

[4] U. Schultheiß, "Über die Adaption eines Kompensators für akustische Echos", *VDI-Fortschritt-Berichte*, Reihe 10, Nr. 90,1988