SIGNAL PROCESSING VII: Theories and Applications
M. Holt, C. Cowan, P. Grant, W. Sandham (Eds.)

Acoustic Echo Control Combined with Two Orthogonalizing Techniques

Christiane ANTWEILER* and Andreas SCHMITZ*

** Institute for Communication Systems and Data Processing,*
Aachen University of Technology, Templergraben 55, D-52056 Aachen, Germany,
Tel/Fax: +49 (241) 806983 / 8888186, E-Mail: acker@ind.rwth-aachen.de

Abstract. One of the main problems of echo cancellation techniques – such as the normalized least mean square algorithm (NLMS algorithm) [1] – is that the convergence properties degrade for colored signal input such as speech signals. The introduction of linear prediction filters into the structure of the conventional echo canceller represents one approach to decorrelate the speech signal. Various attempts have been proposed, which achieve an acceleration of convergence speed by this prewhitening technique [2-5]. An alternative algorithm, named *excited* LMS or ELMS algorithm [6], performs an orthogonalization of the input signal by applying perfect sequences [7] as additional excitation signal, which can provide perfect i.e. maximum speed of convergence. This contribution presents a modified structure of the echo canceller, which combines the ELMS-algorithm with linear prediction techniques. The combination of both algorithms also superposes the improvements achieved for each approach.

1. Introduction

Adaptive echo cancellers are currently being studied for applications such as audio teleconference systems or hands-free telephone sets with high speech quality [1]. The purpose of the echo control is to eliminate the acoustic feedback from the loudspeaker to the microphone.

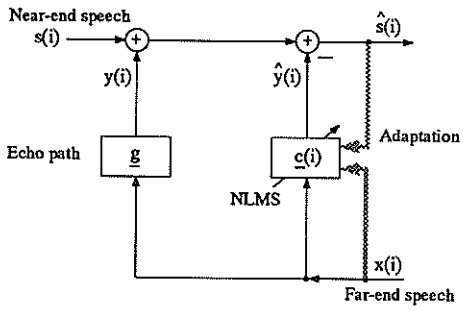


Figure 1: Conventional echo canceller

According to Fig. 1 the acoustic echo path from the loudspeaker to the microphone is represented by the impulse response $\underline{g}$. The compensation filter $\underline{c}(i)$ of length N, where i denotes the time, simulates the acoustic echo path to synthesize a replica $\hat{y}(i)$ of the echo. Then the replica is subtracted from the superposition of the echo $y(i)$ and the near-end speech signal $s(i)$.

The NLMS algorithm, or normalized least-mean-square algorithm [1], represents the most common implementation to adapt the compensation filter:

$$\begin{aligned} \hat{s}(i) &= s(i) + \left(\underline{g} - \underline{c}(i)\right)^T \underline{x}(i) \\ \underline{c}(i+1) &= \underline{c}(i) + \alpha \frac{\hat{s}(i)\,\underline{x}(i)}{||\underline{x}(i)||^2} \end{aligned} \tag{1}$$

with $\underline{c}(i) = (c_0(i), c_1(i), \ldots c_k(i), \ldots c_{N-1}(i))^T$

$\underline{x}(i) = (x(i), x(i-1), \ldots x(i-k), \ldots x(i-N+1))^T$

$||\underline{x}(i)||^2 = \underline{x}^T(i)\,\underline{x}(i)$, where α denotes the stepsize.

One of the main problems of the NLMS algorithm is that the adaptation is driven by speech signals, i.e. colored signals, which reduces the convergence speed significantly. To overcome this drawback different approaches have been proposed to orthogonalize the adaptation signals e.g. [2-6]. In this contribution two of these algorithms are combined into one structure.

- The *excited* LMS or ELMS algorithm [6] represents one method to accelerate convergence speed by orthogonalizing techniques. It is characterized by its modified input signal, where a perfect sequence $\tilde{p}(i)$ with a certain level K is periodically added to the original far-end signal $x(i)$.
- In an alternative approach linear prediction filters are introduced in the adaptation process [3]. The improvement of this algorithm is due to the decorrelation of the adaptation signals.

The following paragraphs 2 and 3 outline the principle concepts of the two approaches. Subsequently, paragraph 4 contains the introduction of the combined structure. Based on computer simulations the results of the new concept are finally analyzed and discussed.

2. *Excited* LMS Algorithm

Especially, the geometric interpretation of the NLMS algorithm [8] illustrates one of the main problems of the NLMS algorithm, i.e. the degradation of convergence speed for colored signal input such as speech signals. The convergence properties of the adaptation process depend on the angle of consecutive compensation filter vectors $\underline{x}(i)$, $\underline{x}(i-1)$. For $\alpha = 1$ perfect adaptation, i.e. $\underline{c}(i) = \underline{g}$, is achieved, if N consecutive vectors $\underline{x}(i)$, $\underline{x}(i-1)$, ..., $\underline{x}(i-N+1)$ are orthogonal in the N–dimensional vector space.

For this reason a white noise process, which is orthogonal in the infinite-vector space, represents an improved excitation signal compared to a speech signal. However, if we consider the N–dimensional vector space, only quasi orthogonality is achieved.

Perfect sequences [7] are characterized by their periodic autocorrelation function which vanishes for all out-of-phase values, i.e. all N phase shifted perfect sequences are orthogonal in the N–dimensional vector space.

→ Perfect sequences represent an optimal excitation signal for the NLMS algorithm.

Fig. 2 confirms these results by summarizing the system distance

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

for different input signals $x(i)$ (solid lines).

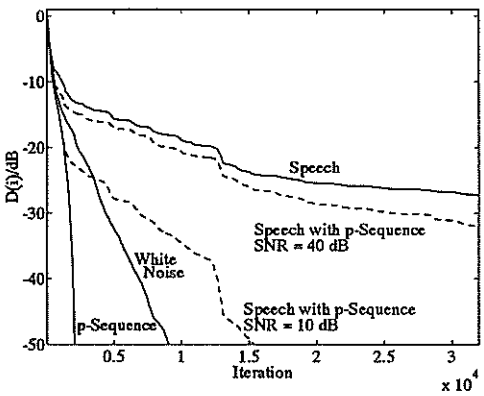


Figure 2. System distance for different input signals ($\alpha = 1$, $N = 1023$, $s(i) = 0$, see Fig. 1)
——— $x(i)$ - - - $x_p(i)$

The special properties of perfect sequences are exploited within the *excited* LMS algorithm or ELMS algorithm [6], which is characterized by its modified input signal

$$x_p(i) = K\tilde{p}(i) + x(i)$$

A perfect sequence $\tilde{p}(i)$ with a certain level K is periodically added to the original signal $x(i)$. Note, that periodic sequences are herein marked with a tilde to distinguish them from related aperiodic sequences or functions. Due to the superposition of $x(i)$ and $\tilde{p}(i)$ and its orthogonal characteristics the perfect sequence fills in the gaps of the spectrum of the speech signal $x(i)$, which can be interpreted as a prewhitening technique. The equations of the ELMS algorithm for compensation and adaptation are as follows:

$$\begin{aligned} \hat{s}(i) &= s(i) + \left(\underline{g} - \underline{c}(i)\right)^T \underline{x}_p(i) \\ \underline{c}(i+1) &= \underline{c}(i) + \alpha \frac{\hat{s}(i)\, \underline{x}_p(i)}{\|\underline{x}_p(i)\|^2} \end{aligned} \tag{4}$$

Fig. 2 depicts the performance of the ELMS algorithm for different power ratios $\frac{E\{x^2(i)\}}{E\{K^2 \tilde{p}^2(i)\}}$ (dashed lines). Obviously, with decreasing power ratio or with increasing loudness level of the perfect sequence the performance of the echo canceller is improved. On the other side the power of the perfect sequence must be limited to keep the subjective annoyance for the near-end listeners sufficiently small. The consideration of both requirements leads to a power ratio of 40 dB as an appropriate compromise.

For a practical implementation of the ELMS algorithm the additional computational complexity is negligeable, while N additional storage locations are required for the storage of the perfect sequence.

3. Decorrelation with Linear Prediction Techniques

In [3] a concept has been proposed improving the convergence behavior of gradient acoustic echo cancellation for hands-free telephone sets by linear prediction techniques.

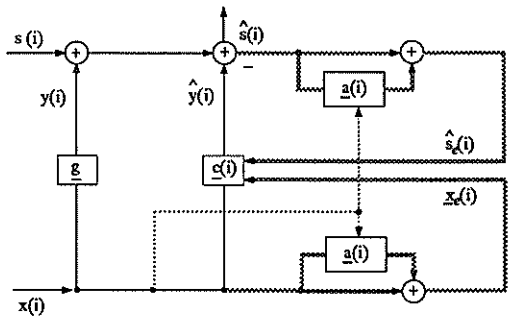


Figure 3. Echo canceller with linear prediction techniques

As illustrated in Fig. 3 the basic structure of the conventional echo canceller was maintained; only in the adaptation paths of the compensator linear prediction filters are inserted. Consequently, the compensation is performed as usual, while the adaptation process is driven by prediction residuals:

$$\underline{c}(i+1) = \underline{c}(i) + \alpha \frac{\hat{s}_e'(i)\, \underline{x}_e'(i)}{\|\underline{x}_e'(i)\|^2} \tag{5}$$

The P coefficients of the prediction filter $\underline{a}(i)$ are blockwise calculated, where $a_m(i)$ represents the m coefficient at time instant i. Due to a special refiltering operation some additional time dependencies are eliminated by using the *refiltered* signals $\underline{x}_e^r(i)$ and $s_e^r(i)$ according to:

$$x_e(i-k) = \sum_{m=0}^{P} a_m(i-k)\, x(i-k-m)$$

$$\rightarrow\ x_e'(i-k) = \sum_{m=0}^{P} a_m(i)\, x(i-k-m)$$

$$\hat{s}_e(i) = \sum_{m=0}^{P} a_m(i)\Big(s(i-m) + \left[\underline{g} - \underline{c}(i-m)\right]^T \underline{x}(i-m)\Big)$$

$$\rightarrow\ \hat{s}_e'(i) = \sum_{m=0}^{P} a_m(i)\Big(s(i-m) + \left[\underline{g} - \underline{c}(i)\right]^T \underline{x}(i-m)\Big)$$

In this approach the adaptation of the compensation filter $\underline{c}(i)$ is performed in the residual signal domain, i.e. with decorrelated signals, which results in an improved performance of the echo canceller.

Applying the mathematical derivation proposed in [9] the computational complexity can be reduced to an order of $2N$, which is mainly the same complexity as for the NLMS algorithm. However, in this concept vector $\underline{x}_e'(i)$ has to be stored, so that N additional storage locations are required.

4. Combined Approach

In this section a modified structure is introduced, which combines the two orthogonalizing techniques described above. As depicted in Fig. 4 the corresponding block diagram includes the addition of a periodic perfect sequence $K\tilde{p}(i)$ as well as linear prediction techniques.

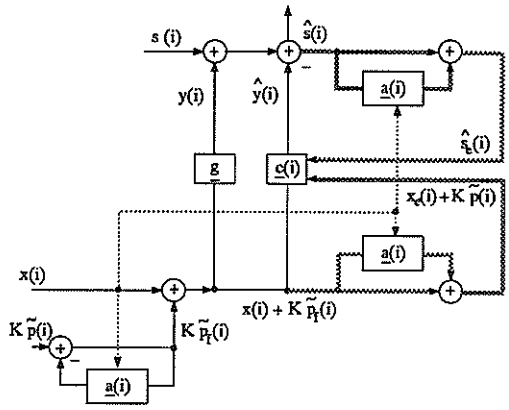


Figure 4. Combined algorithm

In order to recover the perfect sequence after the linear prediction filter for the adaptation process, $K\tilde{p}(i)$ is first inverse filtered by the LPC-synthesis filter. So besides the gain due to the superposition of both algorithms, the perfect sequence $\tilde{p}(i)$ is spectral shaped according to the speech signal $x(i)$, see Fig. 5. The shaping of the perfect sequence reduces the audible distortion for the near-end listeners. Consequently, the power ratio $\frac{E\{x^2(i)\}}{E\{K^2\tilde{p}_f^2(i)\}}$, which is responsible for the subjective annoyance, can be decreased to further improve the performance of the echo canceller.

The equations for compensation and adaptation result in:

$$\begin{aligned} \hat{s}(i) &= s(i) + \left(\underline{g} - \underline{c}(i)\right)^T \left(\underline{x}(i) + K\,\underline{\tilde{p}}_f(i)\right) \\ \underline{c}(i+1) &= \underline{c}(i) + \alpha\,\frac{\hat{s}_e'(i)\left(\underline{x}_e'(i) + K\,\underline{\tilde{p}}(i)\right)}{\left\|\underline{x}_e'(i) + K\,\underline{\tilde{p}}(i)\right\|^2} \end{aligned} \tag{7}$$

Due to the small additional effort for the two orthogonalizing techniques the computational complexity of the combined approach is mainly determined by the operations of the NLMS algorithm. However, the storage of the perfect sequence as well as vector $\underline{x}_e'(i)$ requires $2N$ additional storage locations.

5. Performance

In order to verify the performance of the proposed echo canceller all concepts were examined by computer simulation for two test conditions. The results of the four concepts – the conventional, the two separate orthogonalizing techniques and the combined approach – are summarized in Fig. 6 a) and b), where each curve represents an average out of eight simulations of 4 seconds duration.

In addition the NLMS algorithm and the concept applying linear prediction filters were performed with a second modified initialization procedure. In these algorithms a perfect sequence was applied for the first N iterations, to reduce the effect of the initialization phase and to make the comparison more objective.

Fig. 6 outlines that with respect to the conventional NLMS algorithm both approaches applying orthogonalizing techniques show an improved performance. The combination of both techniques further improves the performance of the digital echo canceller. In comparison to the NLMS algorithm the system distance of the combined concept is decreased by 10.3 dB for the single talk condition and 8.7 dB for double talk. If the special initialization phase is applied the improvements of the system distance still range in an order of 5.5 dB and 3.9 dB, respectively.

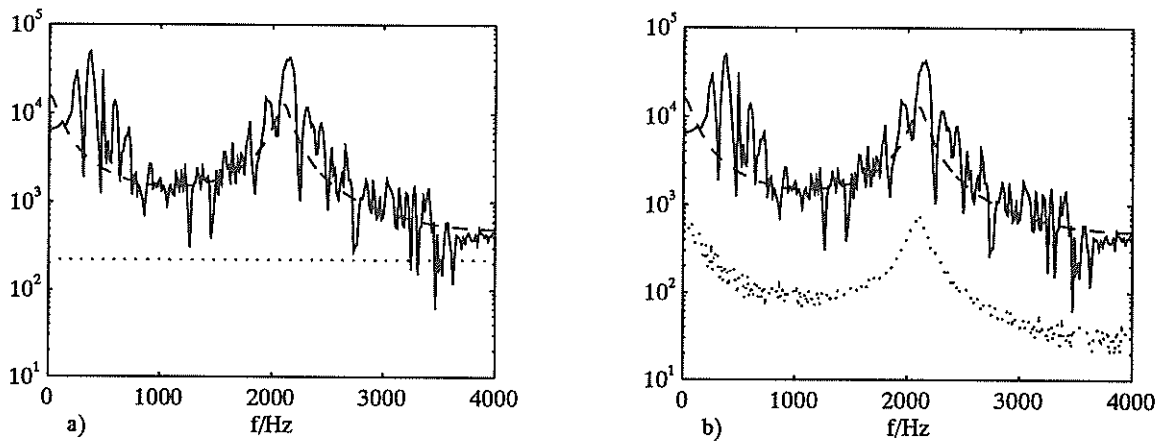
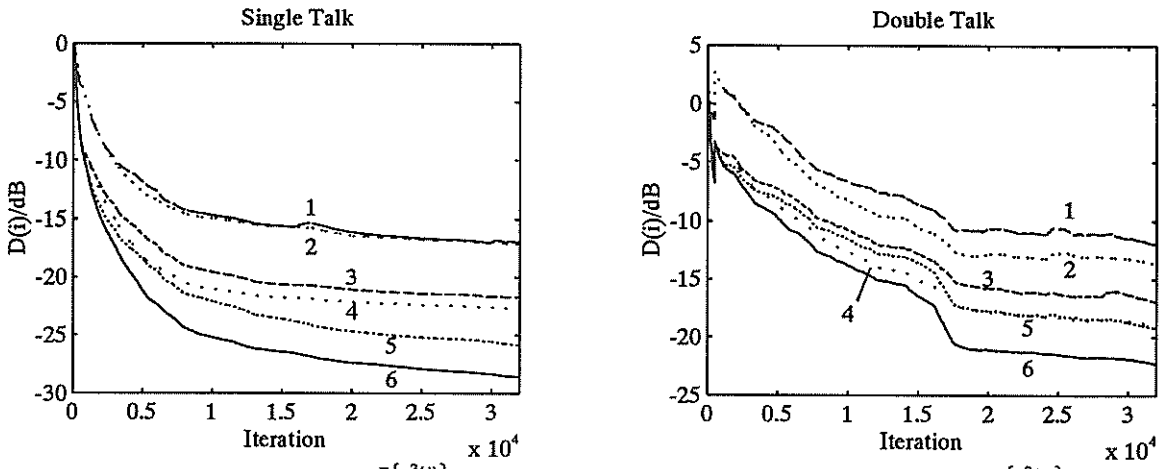


Figure 5. Spectral shaping of the perfect sequence according to the speech signal
——— Short time spectrum of $x(i)$ - - - - Transferfunction of the inverse LPC-filter, $P = 2$
..... a) Short time spectrum of $K\tilde{p}(i)$ ..... b) Short time spectrum of $K\tilde{p}_f(i)$



a) $x(i)$: speech, $s(i)$: white noise, $\frac{E\{s^2(i)\}}{E\{y^2(i)\}} = -50$ dB

b) $x(i)$: speech, $s(i)$: speech, $\frac{E\{s^2(i)\}}{E\{y^2(i)\}} = 0$ dB

Abb. 6. Comparison of the performance for different concepts
($\alpha(i)$ adaptive [10], $N = 511$, $\frac{E\{x^2(i)\}}{E\{K^2\,\tilde{p}^2(i)\}} = 40$ dB, or respectively $\frac{E\{x^2(i)\}}{E\{K^2\,\tilde{p}_j^2(i)\}} = 30$ dB, P=2)

1 — — — NLMS algorithm
2 . . . . . Concept with linear prediction techniques
3 - - - - - - NLMS algorithm, addition of $K\tilde{p}(i)$ during the initialization phase
4 . . . . . Concept with linear prediction techniques, addition of $K\tilde{p}(i)$ during the initialization phase
5 ..... ELMS algorithm
6 ———— Combined approach

6. Conclusions

In this paper two individual orthogonalizing techniques were presented and discussed. In comparison to the conventional echo canceller both techniques lead to an improved adaptation process, while the computational complexity is mainly maintained.

The combination of the two orthogonalizing techniques into one algorithm also superposes the improvements achieved for each approach, see Fig. 6. In addition the perfect sequence excitation is masked by spectral shaping techniques. Consequently, the combined algorithm complements the characteristics of the individual approaches in such a way that the audible distortion for the near-end listeners is additionally reduced.

For a practical implementation it is of special interest that these gains can be achieved by only a small increase of computational complexity. However, as an additional effort $2N$ storage locations are required, which is warrantable due to the improved performance.

References

1. **E. Hänsler:** "The hands-free telephone problem – An annotated bibliography – An Update", *3rd International Workshop on Acoustic Echo Control*, Plestin les Greves, September 1993, pp. 5–18
2. **C. Acker (Antweiler), P. Vary, H. Ostendarp:** "Acoustic Echo Cancellation Using Prediction Residual Signals": *Proceedings of Eurospeech 91, Genova*, pp. 1297–1300, September 1991
3. **C. Acker (Antweiler), P. Vary:** "Combined Implementation of Predictive Speech Coding and Acoustic Echo Cancellation", *Proceedings of EUSIPCO-92, Brussels*, pp. 1641–1644, August 1992
4. **M. Mboup, M. Bonnet, O. Macchi:** "A New Adaptive Prewhitening Filter for Acoustic Echo Cancellation", *2nd Workshop on Acoustic Echo Cancellation*, Genova, September 1991
5. **U. Schultheiß:** "Über die Adaption eines Kompensators für akustische Echos", *VDI-Fortschritt-Berichte*, Reihe 10, Nr. 90, 1988
6. **C. Antweiler, M. Dörbecker:** "Perfect Sequence Excitation of the NLMS Algorithm and its Application to Acoustic Echo Control", to be published in *Annales des Telecommunications*, July-August 1994
7. **H. D. Lüke:** "Korrelationssignale", *Springer-Verlag Berlin Heidelberg New York*, (ISBN 3–540–54579–4), 1992
8. **T. A. C. M. Claasen, W. F. G. Mecklenbräuker:** "Comparison of the Convergence of Two Algorithms for Adaptive FIR Digital Filters", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-29, No. 3, June 1981, pp. 670–678
9. **R. D. Poltmann:** "A New Method for the NLMS Algorithm with Time-Variant Decorrelation Filters", submitted to *IEEE Transactions on Circuits and Systems II*
10. **S. Yamamoto, S. Kitayama:** "An Adaptive Echo Canceller with Variable Step Gain Method", *Trans. IECE Japan*, Vol. E65, No. 1, pp. 1–8, January 1982.