ADAPTATION OF ACOUSTIC ECHO CANCELLERS EXPLOITING SPECTRAL CHARACTERISTICS OF ROOM IMPULSE RESPONSES

Christiane Antweiler

Institute of Communication Systems and Data Processing,
Aachen University of Technology, Templergraben 55, 52056 Aachen, Germany
Tel.: +49 241 806983, Fax: +49 241 8888186, Email: antweiler@ind.rwth-aachen.de

ABSTRACT

Adaptive filters are used extensively in telecommunication e.g. for acoustic echo control, noise reduction or channel equalization. The so called *normalized least mean square* (NLMS) algorithm represents one of the most widely used gradient-based adaptation algorithm. One of the main problems of the NLMS algorithm in an echo control application is that the adaptation is driven by speech signals, i.e. spectrally colored signals, which reduce the convergence speed significantly. In this paper an alternative approach to several existing algorithms (e.g. [1,6,7]) is introduced, which is based on a priori knowledge of the spectral characteristics of the room impulse response in form of spectral weighting filters. This technique results in an accelerated convergence rate and benefits from a more robust behavior in the presence of background noise, such as occurring for example in a car. Besides the theoretical derivation of the proposed algorithm its realization aspects, complexity features, and simulation results are discussed in detail.

1. INTRODUCTION

Any hands-free operation of a telephone comprising a loudspeaker-microphone system instead of a hand-set has to deal with the problem that due to the coupling between the loudspeaker and the microphone the loudspeaker signal is echoed back to the microphone (see Fig. 1). A common technique to reduce the echo signal is the use of a compensation filter $\underline{c}(i)$, which models the acoustic echo path $\underline{g}(i)$. Processing the loudspeaker signal by the compensation filter results in a replica of the acoustic echo signal, which is subtracted from the microphone signal.

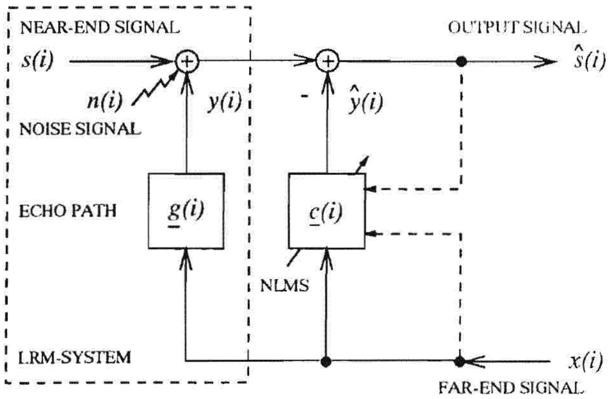

Figure 1: Conventional echo cancellation

Due to the long duration of the room impulse response the selection of an adaptation algorithm for the compensation filter coefficients is restricted. One approach, which fulfills the requirements of complexity and stability is the so called *normalized least mean square* (NLMS) algorithm:

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

with

$$\tilde{s}(i) = s(i) + \left[\underline{g}(i) - \underline{c}(i)\right]^T \underline{x}(i) \tag{2}$$

$$\underline{c}(i) = (c_0(i), c_1(i), ..., c_k(i), ..., c_{N-1}(i))^T \tag{3}$$

$$\underline{x}(i) = (x(i), x(i-1), ..., x(i-N+1))^T \tag{4}$$

$$||\underline{x}(i)||^2 = \underline{x}^T(i)\,\underline{x}(i)\,, \tag{5}$$

where i denotes the time instant, k the coefficient index and N the length of filter $\underline{c}(i)$.

One major drawback of the NLMS algorithm is that its convergence rate depends on the statistical characteristics of the input signal $x(i)$. Especially, in case of colored signal input such as speech signals the performance of the NLMS algorithm degrades. To overcome this problem different approaches applying linear prediction techniques have been proposed to decorrelate the speech signals (e.g. [1,6,7]). With these concepts a significant improvement of the identification process is achieved since the introduction of the linear prediction filters leads to a spectrally more uniform excitation of all frequency components of the room impulse response. Taking into consideration that only those frequency components of the room impulse response, which are excited by the input signal $x(i)$, can be identified, the benefits of the prediction filters becomes obvious.

In order to accelerate the convergence rate of the NLMS algorithm in this contribution an alternative concept is proposed, where the linear prediction filters are replaced by spectral weighting filters exploiting spectral characteristics of the room impulse response, see also [2]. Within the subsequent section the new algorithm is motivated and theoretically derived. Its introduction comprises the discussion of its performance in noisy environment, the realization, and complexity aspects. The benefits of the new concept are shown in form of simulation results for different situations.

2. NLMS WITH SPECTRAL WEIGHTING

In analogy to the structure proposed in [1,6] with Fig. 2 an alternative approach is presented. The comparison with Fig. 1 shows the difference to the conventional echo canceller.

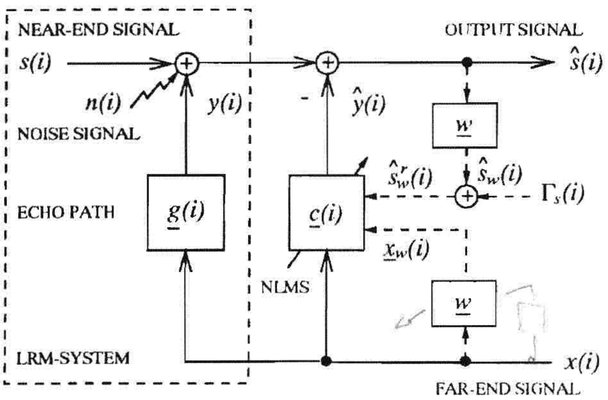

Figure 2: NLMS with spectral weighting filters

In order to accelerate the convergence rate spectral weighting filters $\underline{w}$ have been implemented in the adaptation paths of the compensation filter as depicted in Fig. 2. The basic idea of the new approach is to choose weighting filter $\underline{w}$ such that its transfer function represents the spectral envelope of the room impulse response (RIR) as shown in Fig. 3.

In the lower adaptation path filter $\underline{w}$ performs a spectral weighting of the excitation signal $x(i)$ in such a way that spectral components, which contribute more to the echo signal, are emphasized due to the spectral envelope of the RIR. The emphasized excitation of the "more

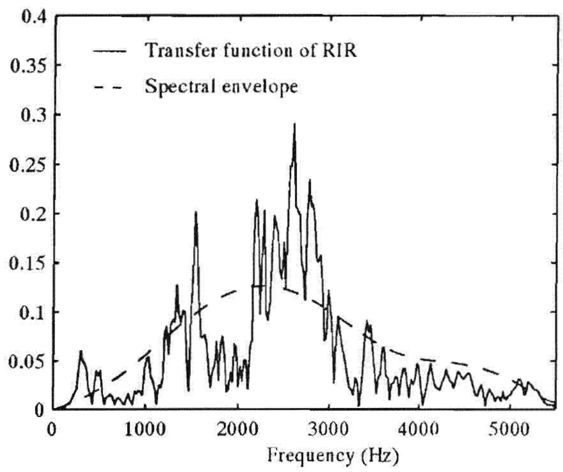

Figure 3: Transfer function and spectral envelope of the room impulse response (RIR)

important" spectral components leads to an improved tracking of these components within the NLMS-based identification process.

2.1. Algorithm

While filter $\underline{w}$ in the lower path performs the desired spectral weighting, the filter in the upper path and the correction term $\Gamma_s(i)$ guarantee that in analogy to the NLMS algorithm with the resulting gradient

$$\underline{c}(i+1) = \underline{c}(i) - \frac{1}{2}\,\frac{\alpha}{||\underline{x}_w(i)||^2}\,\underline{\nabla}(i) \qquad (6)$$

$$= \underline{c}(i) + \frac{\alpha}{||\underline{x}_w(i)||^2}\,\hat{s}^r_w(i)\,\underline{x}_w(i) \qquad (7)$$

the square of the instantaneous signal $\hat{s}^r_w(i)$ is minimized. For this purpose the correction term $\Gamma_s(i)$ (see also [6,1]) eliminates the dependence of $\hat{s}_w(i)$ from previous sets of coefficients $\underline{c}(i-1), \ldots, \underline{c}(i-N_w+1)$ according to

$$\Gamma_s(i) = \hat{s}^r_w(i) - \hat{s}_w(i) \qquad (8)$$
$$= \sum_{j=0}^{N_w-1} w_j\,\left(\underline{c}^T(i-j) - \underline{c}^T(i)\right)\,\underline{x}(i-j)$$

and it can be formulated for all $k = 0, \ldots, N-1$:

$$\nabla_k(i) = -2\,\hat{s}^r_w(i)\,x_w(i-k)$$
$$= -2\,\hat{s}^r_w(i)\sum_{j=0}^{N_w-1} w_j\;x(i-k-j)$$
$$= 2\,\hat{s}^r_w(i)\sum_{j=0}^{N_w-1} w_j\,\frac{\partial\,\hat{s}^r(i-j)}{\partial\,c_k(i)}$$
$$= 2\,\hat{s}^r_w(i)\,\frac{\partial\,\hat{s}^r_w(i)}{\partial\,c_k(i)}$$
$$= \frac{\partial\,\hat{s}^{r2}_w(i)}{\partial\,c_k(i)}\,, \qquad (9)$$

where N_w denotes the length of weighting filter $\underline{w}$. Like in the conventional NLMS algorithm within the proposed algorithm the compensation is still performed with the original signals $x(i)$ and $s(i)$. However, the adaptation process now applies with $x_w(i)$ and $\hat{s}^r_w(i)$ spectrally shaped signals minimizing the instantaneous gradient of the square of the shaped signal $\hat{s}^r_w(i)$. Furthermore, filter $\underline{w}$ in the upper path leads to an additional positive effect, i.e. the reduction of the interfering influence of the background noise signal $n(i)$.

2.2. Influence of Background Noise

The signal $s(i)$ of the near-end talker affects the adaptation process of the echo canceller. The use of a time variant step size parameter – serving as double talk detector and slowing down the adaptation process in the presence of signal $s(i)$ – represents one approach to overcome this problem (see e.g. [5]). Moreover, especially in mobile transmission systems the algorithm has to deal with impairing acoustic background noise $n(i)$. As an example, Fig. 4 shows the short-time spectrum of a noise signal $n(i)$ recorded in a moving car. Typically the noise spectrum comprises particularly low frequency components, which interfere the adaptation process. The introduction of weighting filter $\underline{w}$ in the upper path now leads to an attenuation of the lower frequency components due to the bandpass characteristic of the spectral envelope. As a result, the influence of the noise signal $n(i)$ for the adaptation process is reduced. So, besides the accelerated convergence rate the proposed algorithm also benefits from a more robust behavior in the presence of background noise, such as occurring for example in a car.

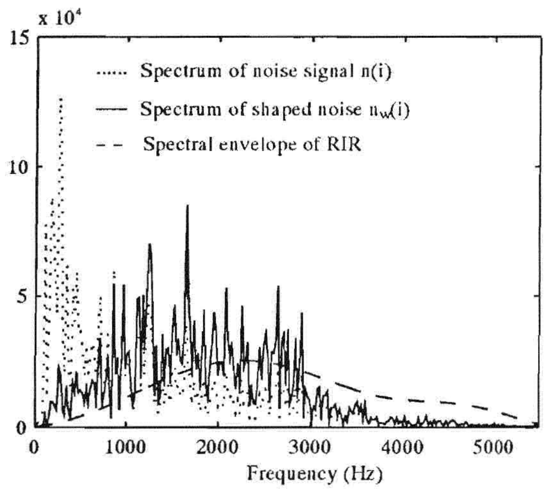

Figure 4: Spectral weighting of the background noise

2.3. Realization Aspects

The proposed concept is based on the fundamental property of the room impulse response, that changes in the room only influence the fine structure of the room transfer function, while its spectral envelope mainly remains unchanged. Investigations documented in [5] and the example in Fig. 5 underline, that in practice this assumption is sufficiently met.

In order to determine the spectral envelope, i.e. the filter coefficients $\underline{w}$, in a real application, a periodic perfect sequence [3] is briefly applied as excitation signal onto the system given in Fig. 1. Due to the special correlation characteristics of perfect sequences the resulting set of coefficients $\underline{c}$ obtained after the initialization phase provides an appropriate approximation of the room impulse response. Subsequently, an LPC analysis is performed with c_k as input signal and a high predictor order. The transfer function of the resulting FIR filter rebuilds the spectral envelope of the inverse transfer function of the

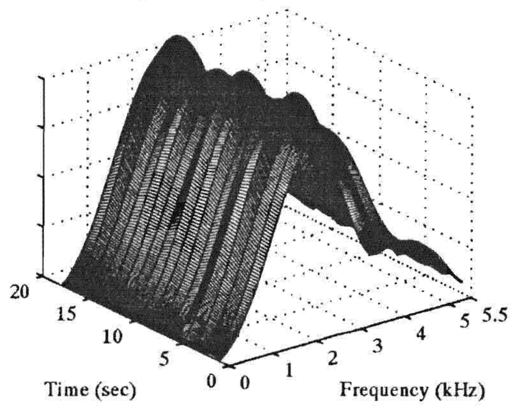

Figure 5: Spectral envelope of RIR for different time instants

room. A second LPC analysis of the impulse response of the just determined high resolution predictor finally leads to the spectral envelope of the room impulse response. A first optimization has indicated, that already a rough model of the spectral envelope with few coefficients $\underline{w}$ (e.g. $N_w = 4$ or 5) is sufficient, where the number of coefficients N_w represents the predictor order of the second LPC analysis. Due to the fact, that these computations only have to be performed once in the beginning of the hands-free operation, they do not significantly influence the complexity of the concept.

2.4. Complexity

The numerical complexity is increased by the additional filtering of the signals $x(i)$ and $s(i)$ and by the calculation of the correction term $\Gamma_s(i)$. The number of multiplications needed for the filtering procedure amounts to $2N_w$. Also the complexity for the determination of the correction term $\Gamma_s(i)$ can be kept relatively small with $(N_w - 1)(N_w + 3)$ multiplications using an efficient recursive algorithm according to [6].

3. SIMULATION RESULTS

The NLMS algorithm applying spectral weighting filters $\underline{w}$ has been introduced in order to accelerate the convergence rate. Furthermore, as stated in Section 2.2, the influence of an occurring background noise signal $n(i)$ for the adaptation process is reduced. Both effects can be visualized by simulation results. Fig. 6 illustrates the system distance according to

$$\frac{D(i)}{\mathrm{dB}} = 10 \ \lg \ \frac{\left|\left|\underline{g}(i) - \underline{c}(i)\right|\right|^2}{\left|\left|\underline{g}(i)\right|\right|^2} \qquad (10)$$

for three algorithms in two different situations.

Obviously, in comparison to the conventional NLMS algorithm the convergence rate of the NLMS algorithm with linear prediction filters [1] and with spectral weighting increases especially in the transition from a time

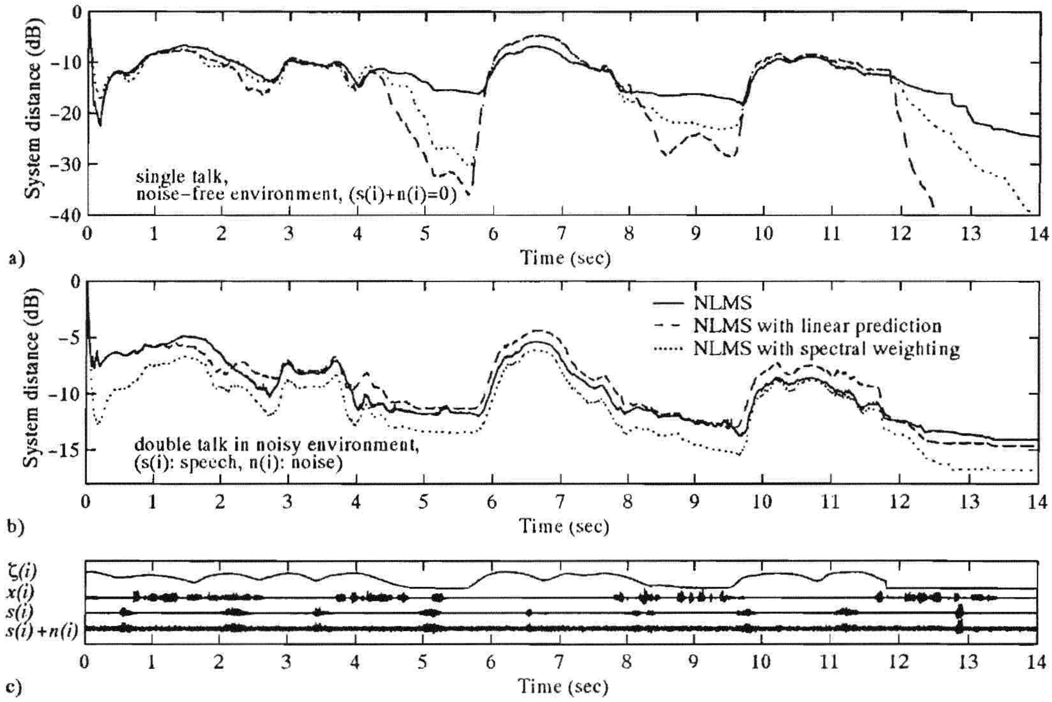

Figure 6: Comparison of the system distance for a time variant simulation (see [4]) ($\zeta(i)$ indicates the intensity of time variance of RIR)

variant ($\zeta(i)$ large) to a time invariant ($\zeta(i)$ small) room impulse response. While in the noise-free environment (Fig. 6–a) during phases of a time variant room impulse response all algorithms show comparable performance, in noisy environment (Fig. 6–b) the positive effect of the reduced interfering influence of the noise signal can be noted for the NLMS with spectral weighting filters resulting in a smaller system distance. With increasing level of the noise signal $n(i)$ this effect augments respectively.

4. CONCLUSIONS

The introduction of the spectral weighting filters exploiting a priori knowledge of spectral characteristics of the room impulse response results in an improved quality of speech transmission applying hands-free telephony due to its accelerated convergence rate – even in noisy conditions. In an application inside a car, where an echo canceller has been combined with an automatic speech recognition system, the recognition rates could be improved from 90.5% to 96% by the introduction of the weighting filters. It is of special interest, that these improvements are obtained without considerable additional computational effort (see [6]), which has been proven by a real-time implementation.

5. REFERENCES

[1] C. Antweiler, "Orthogonalisierende Algorithmen für die digitale Kompensation akustischer Echos", Dissertation, ABDN, Band 1, Verlag der Augustinus Buchhandlung, Aachen 1995

[2] C. Antweiler, "Kompensation von akustischen Echosignalen unter Verwendung von spektralen Gewichtungsfiltern", 9. Aachener Kolloquium "Signaltheorie", Bild- und Sprachsignale, Aachen, 18.-20. March 1997, pp. 229–232

[3] C. Antweiler, M. Antweiler, "System Identification with Perfect Sequences Based on the NLMS Algorithm", *AEÜ, International Journal of Electronics and Communications*, Vol. 49, No. 3, Mai 1995, S. 129–134

[4] C. Antweiler, H.-G. Symanzik, "Simulation of Time Variant Room Impulse Responses", *ICASSP-95*, Detroit, USA, 8.-12. Mai, 1995, S. 3031–3034

[5] R. Frenzel, "Freisprechen in gestörter Umgebung", *VDI-Fortschritt-Berichte,* Reihe 10, Nr. 228, VDI-Verlag, Düsseldorf, Deutschland, 1992

[6] R. Poltmann, "A Fast Method for the NLMS Algorithm with Time-Variant Decorrelation Filters", *Signal Processing*, Vol. 6, Nr. 2, März-April 1995, S. 203–206

[7] U. Schultheiß, "Über die Adaption eines Kompensators für akustische Echos", *VDI-Fortschritt-Berichte,* Vol. 10, No. 90, Düsseldorf, Deutschland, 1988

The author would like to thank the employees from Lernout & Hauspie Speech Products N.V. for kindly providing the software for the automatic speech recognition system.