Proceedings of 20th International Congress on Acoustics, ICA 2010

23–27 August 2010, Sydney, Australia

Improved Acoustic Echo Cancellation for low SNR based on blockwise combination of filters

Luis A. Azpicueta-Ruiz (1), Marcus Zeller (2), Aníbal R. Figueiras-Vidal (1), Jerónimo Arenas-García (1) and Walter Kellermann (2)

(1) Dept. of Signal Theory and Communications, Universidad Carlos III de Madrid, 28911 Leganés, Madrid, Spain.
(2) Multimedia Communications and Signal Processing, University of Erlangen-Nuremberg, 91058 Erlangen, Germany.

PACS: 43.60.Mn

ABSTRACT

Acoustic echo cancellers (AEC) are becoming increasingly important because of the widespread use of hands-free devices. Due to their simplicity, most of the cancellers rely on NLMS-type adaptive filters to model and track the time-varying echo path. Recently, adaptive combinations of filters are gaining increasing popularity as a flexible and versatile approach to overcome compromises inherent to adaptive filters, thus enhancing the overall performance. Regarding AEC scenarios, such filter combinations have already been proposed for, e.g., improving the trade-off between convergence speed and steady-state error or for reducing the dependency on varying ratios of linear and nonlinear distortions.

In this paper, we present a new AEC approach, showing improved performance for unknown or time-varying signal-to-noise ratios (SNR). The proposed scheme exploits the fact that the coefficient energy of a typical echo path is not uniformly distributed, but decays exponentially. Under this condition, an NLMS filter will introduce significant estimation errors for less significant filter taps due to gradient noise. Since the number of affected coefficients strongly depends on the SNR and hence on the implied noise floor, the cancellation performance may degrade considerably for low SNRs. In order to relieve the coefficient noise, the adaptive impulse response is split into a number of non-overlapping blocks, each of which is combined with a virtual 'zero-block', having fixed zero coefficients, with time-varying relative weights for both the nonzero and the zero-block. In practice, this results in a possibly biased estimation of some of the filter coefficients. However, it has been shown that such estimates can yield advantages in terms of mean-square error, especially for low SNRs. The combination of each block is implemented by a convex mixing, where the control parameter is updated according to a stochastic gradient descent method so as to minimize the global error of the AEC. For moderate block numbers, the increment in computational cost over a conventional NLMS canceller is negligible. In particular, this contribution investigates the operation of the blockwise combined filter for low SNR conditions, comparing its performance with standard NLMS- and PNLMS-type filters.

The robustness and benefits of the proposed approach are thereby experimentally verified for noise and speech inputs. Moreover, the influence of the number of blocks and the mixing parameters is also studied and indications on future work (as e.g. accounting for impulsive noise or an extension to nonlinear filters) are given.

INTRODUCTION

In many of today's communication scenarios, the desire for a natural, untethered movement of all participating speakers results in the need for hands-free communication systems. Consequently, due to the unconstrained sound propagation from the loudspeaker to microphone, such setups usually require an acoustic echo cancellation (AEC) mechanism in order to prevent the echo component in the local room from being fed back to the far-end speaker (Hänsler and Schmidt 2004). Moreover, such cancellers become also important for the widespread use of mobile devices, since the light housings of these units inevitably yield a relatively free sound propagation even in the normal (i.e. not hands-free) use mode or in video telephony situations. Fig. 1 illustrates the basic AEC task, where the echo $y(k)$ contained in the microphone signal $d(k)$ is to be removed by the replica $\widehat{y}_c(k)$ generated by the adaptive cancellation filter.

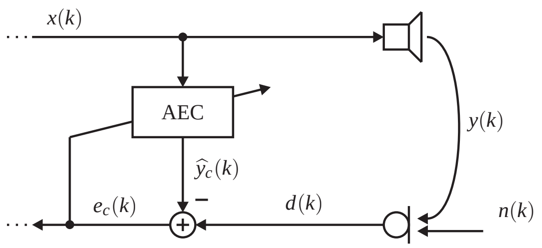


Figure 1: Acoustic echo cancellation scenario with adaptive filter (AEC). No double-talk is considered in this scheme.

In order to model and track the time-varying echo path, mainly corresponding to the room impulse response (RIR), most echo cancellers rely on adaptive transversal filters due to their simplicity. In that sense, AEC is a classical and prime example of the system identification problem, involving considerably large filter sizes and input signals with challenging statistical properties (e.g. nonstationary speech). To date, most practical algorithms often employ rules based on the least mean square (LMS) scheme for the adjustment of the filter coefficients for the sake of computational simplicity and numerical robustness.

However, there is usually the well-justified *a priori* assumption that RIR exhibits an exponential decay curve and, therefore, the energy reflected by the adaptive model coefficients will decrease in the same way.

In (Makino et al. 1993), this observation has been used in order to weight the filter tap updates according to their position of the coefficient in the filter, thereby increasing the speed of convergence over a normalized LMS (NLMS) algorithm. However, the effective application of this approach requires some knowledge regarding the bulk delay and weight-decay constant, which is usually not available. In (Duttweiler 2000), this modification has been generalized to the notion of 'proportionate' updating and has mainly been motivated by seeking to alleviate the drawback of the uniform LMS-type updates in presence of quite sparse impulse responses. This scheme does not assume other *a priori* knowledge about the echo channel but its sparsity. The extended version in (Benesty and Gay 2002), named improved PNLMS (IPNLMS), has proposed a control parameter that can be used to switch between the pure NLMS and the weighted updates such that the performance of the adaptation is more robust to scenarios with dispersive response (i.e. PNLMS). Both PNLMS and IPNLMS offer additional advantages with respect to the NLMS method, mainly in sparse enviroments, requiring, however, an increased computational cost. In addition, PNLMS can obtain a worse performance with respect to that of a single NLMS adaptive filter, if the RIR of the echo cancellation scenario is not so sparse, but rather dispersive.

Furthermore, adaptive combinations of adaptive filters (Arenas-García et al. 2006) have recently gained increased popularity as a flexible and versatile approach to overcome the various compromises due to the specific properties of each adaptation algorithm. In the context of echo cancellation, these schemes have been proposed for, e.g., improving the trade-off between convergence speed and residual error (Arenas-García et al. 2006, Azpicueta-Ruiz et al. 2008), reducing the dependency on unknown and varying ratios of linear and nonlinear distortions (Azpicueta-Ruiz et al., 2010) and for estimating the optimum model length of the adaptive filter (Zeller et al. 2009).

This paper proposes a novel AEC approach, that mainly focuses on improving the performance for unknown and/or time-varying signal-to-noise ratios (SNR), especially for low SNRs. Like the methods in (Makino et al. 1993) and (Benesty and Gay 2002), the proposed combination scheme exploits the fact that the coefficient energy of a typical echo path is not uniformly distributed, but decays exponentially. Under this condition, an NLMS filter will introduce significant estimation errors for less significant filter taps due to gradient noise. Since the number of affected coefficients strongly depends on the present SNR and hence the implied noise floor, the cancellation performance may degrade considerably for low SNRs.

Splitting the adaptive impulse response into a number of non-overlapping blocks, the impact of quite low SNR on the coefficient noise can be relieved, since each block is combined with a virtual 'zero-block', having fixed zero coefficients. The combinations between each block with actual, non-zero coefficients and the 'zero-block' is implemented by a convex mixing, where the control parameter is updated according to a stochastic gradient descent method so as to minimize the global error of the AEC. In practice, this results in a possibly biased estimation of some of the filter coefficients. However, it is well-known that such estimates can yield advantages in terms of mean-square error, especially for low SNRs (Kay and Eldar 2008). In addition, for moderate block numbers, the increment in computational cost over a conventional canceller is negligible.

In particular, we will study the performance of the new approach based on standard NLMS-type filters and the IPNLMS from (Benesty and Gay 2002). The robustness and benefits of the proposed approach are thereby experimentally verified for stationary noise and nonstationary speech inputs.

The rest of this contribution is structured as follows: the second section presents a detailed explanation of the proposed echo canceller. The behavior of our scheme is shown in the third section for different input signals by means of several experiments, where the influence of different parameters has been studied. Finally, the conclusions of our work and suggested further research are provided in the last section.

PROPOSED ECHO CANCELLER

Echo cancellation constitutes a particular case of plant identification problems, with the objective to estimate an acoustic echo path in order to eliminate the replicas of the input signal $u(k)$ present in the microphone signal $d(k)$, that can be expressed adopting vector notation as

$$d(k) = \mathbf{h}^T(k)\mathbf{u}(k) + n(k) = y(k) + n(k). \quad (1)$$

Column vector $\mathbf{h}(k)$ represents the impulse response of the echo path, $\mathbf{u}(k)$ is the input vector containing the last N input signal samples, i.e., $\mathbf{u}(k) = [u(k), u(k-1), ..., u(k-N+1)]^T$, where N denotes the length of the plant $\mathbf{h}(k)$, and $n(k)$ is the uncorrelated noise that is also received by the microphone. Note that double-talk is not considered in (1).

Different kinds of adaptive filters are normally employed to model $\mathbf{h}(k)$ as part of an AEC, among others, the NLMS and IPNLMS adaptive filters. The objective of our proposal is to enhance the behavior of acoustic echo cancellers based on single adaptive filters under low SNR conditions, while still maintaining a good performance if the actual SNR present in the cancellation scenario is high.

The proposed scheme is based on a block-wise decomposition of an adaptive filter with length N (Arenas-García and Figueiras-Vidal 2009), whose impulse response reads $\mathbf{w}^T(k) = [w_1(k), w_2(k), ..., w_N(k)]$. This filter produces an output $\widehat{y}(k)$, that can be expressed as the superposition of M partial block outputs $\widehat{y}_m(k)$, i.e.,

$$\widehat{y}(k) = \mathbf{w}^T(k)\mathbf{u}(k) = \sum_{m=1}^{M} \mathbf{w}_m^T(k)\mathbf{u}_m(k) = \sum_{m=1}^{M} \widehat{y}_m(k). \quad (2)$$

Therefore, whole filter $\mathbf{w}(k)$ is split into M non-overlapping blocks with length $P = N/M$, i.e.,

$$\mathbf{w}_m(k) = [w_{1+(m-1)P}(k), w_{2+(m-1)P}(k), ..., w_{mP}(k)]^T, \quad (3)$$

where $m = 1, ..., M$ denotes a specific block. Similarly, $\mathbf{u}_m(k)$ represents a column vector including the P input samples of signal $u(k)$ necessary to obtain $\widehat{y}_m(k)$, i.e.,

$$\mathbf{u}_m(k) = [u(k-(m-1)P), u(k-1-(m-1)P), ..., u(k-mP+1)]^T. \quad (4)$$

The proposed scheme can be implemented employing different kinds of adaptive filters $\mathbf{w}(k)$. Normally, the adaptation of filter coefficients depends on error signal $e(k) = d(k) - \widehat{y}(k)$, and we can assume that $\mathbf{w}(k+1) = \mathbf{f}[\mathbf{w}(k), \mathbf{u}(k), e(k)]$, where $\mathbf{f}$ refers to the function which characterizes each particular adaptation algorithm.

The performance of the adaptive filter can be improved adaptively weighting partial outputs $\widehat{y}_m(k)$, with $m = 1, ..., M$, which could bias the estimation of the unknown filter. Thus, output of the proposed scheme $\widehat{y}_c(k)$ is constructed by means of convex combinations of each block output $\widehat{y}_m(k)$ with the output of a *virtual* zeros-block, whose coefficients are always equal to zero, and hence, do not need to be updated. Therefore,

$$\widehat{y}_c(k) = \sum_{m=1}^{M} \{\widehat{y}_m(k)\lambda_m(k) + [1-\lambda_m(k)]\cdot 0\} = \sum_{m=1}^{M} \widehat{y}_m(k)\lambda_m(k) \tag{5}$$

where $\lambda_m(k)$ corresponds to the m-th mixing parameter. The multiplication included in (5) results in a biased estimation of the coefficients of unknown filter $\mathbf{h}(k)$ when $\lambda_m(k) \neq 1$. In (Lázaro-Gredilla et al., 2010) it has been shown that these biased estimates can in fact provide reductions in terms of mean-square error (MSE), mainly for low SNRs.

The aforementioned interpretation of multiplications in (5) as combinations with zero-blocks permits to employ well-known update schemes for adaptively learning $\lambda_m(k)$ with $m = 1,...,M$. All the mixing parameters are adapted in order to minimize power of the estimation error $e_c(k) = d(k) - \widehat{y}_c(k)$, by means of the normalized stochastic gradient algorithm published in (Azpicueta-Ruiz et al. 2008). In order to keep mixing parameters in range [0,1], we define

$$\lambda_m(k) = \mathrm{sgm}[a_m(k)] = \frac{1}{1+\exp[-a_m(k)]}, \tag{6}$$

and, instead of directly updating $\lambda_m(k)$, we adjust $a_m(k)$ as follows

$$\begin{aligned} a_m(k+1) &= a_m(k) - \frac{\mu_a}{p_m(k)}\frac{\partial e_c^2(k)}{\partial a_m(k)} \\ &= a(k) - \frac{\mu_a}{p_m(k)} e_c(k)\widehat{y}_m(k)\frac{\partial \lambda_m(k)}{\partial a_m(k)}, \end{aligned} \tag{7}$$

where μ_a is a step-size that manages the adaptation of $a_m(k)$, and $p_m(k) = \beta p_m(k-1) + (1-\beta)\widehat{y}_m^2(k)$, with β close to 1, is a low-pass filtered estimation of the power of $\widehat{y}_m(k)$.

Although other adaptation schemes can be used to adjust $\lambda_m(k)$ (Candido et al. 2008), the sigmoïdal activation in (6) reduces the gradient noise in (7) when $\lambda_m(k) \approx 1$ or $\lambda_m(k) \approx 0$. Note that this is especially important in order not to degrade algorithm's performance, if no bias is necessary, i.e., $\lambda_m(k) \approx 1$. See (Lázaro-Gredilla et al., 2010) for more details.

In order to avoid algorithm paralysis in (7), the range of values of $a(n)$ is tipically restricted to [-4,4]. However, this prevents $\lambda_m(k)$ from reaching the limit values of the interest range, i.e., 0 and 1. This drawback can be solved by slightly modifying (6),

$$\lambda_m(k) = \frac{\mathrm{sgm}[a_m(k)] - \mathrm{sgm}[-4]}{\mathrm{sgm}[4] - \mathrm{sgm}[-4]} \tag{8}$$

which allows $\lambda_m(k)$ to reach all the values inside [0,1].

Fig. 2 represents a block diagram of the proposed echo canceller, which can be seen as consisting of two different parts: An adaptive filter whose coefficients are updated to minimize its error signal, $e(k)$; and the structure necessary to obtain outputs of each block $\widehat{y}_m(k)$ and to adapt mixing parameters $\lambda_m(k)$ in order to minimize the power of cancellation error, $e_c(k)$. In addition, it should be noted that our algorithm can be applied employing different kinds of adaptation algorithms.

With this scheme, it is possible to consider the exponential decay of the energy of echo path $\mathbf{h}(k)$, obtaining an echo estimate $\widehat{y}_c(k)$ that enhances that of an standard adaptive filter, mainly for low SNRs. The performance of the proposed scheme can be explained in the following way:

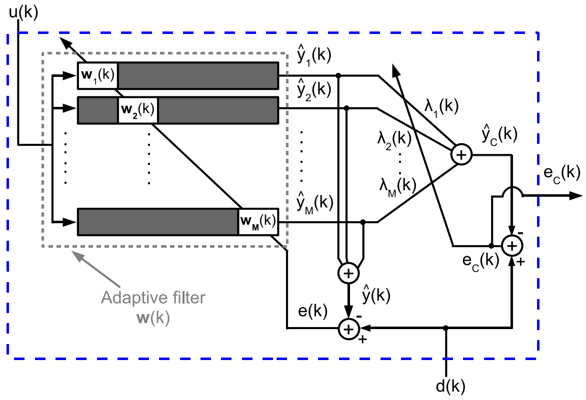


Figure 2: Scheme of the proposed canceller. For the sake of clarity, each block of the adaptive filter $\mathbf{w}_m(k)$ has been represented as a complete adaptive filter where the shadowed regions correspond with non-active coefficients.

- When the SNR is high, estimation $\widehat{y}(k)$ is normally not affected by gradient noise produced by the adaptation of whole adaptive filter $\mathbf{w}(k)$. In this case, $\lambda_m \approx 1$ for $m = 1,...,M$, and the overall scheme behaves as adaptive filter $\mathbf{w}(k)$, since $\widehat{y}_c(k) \approx \widehat{y}(k)$.
- If the SNR is low, estimation $\widehat{y}(k)$ is degraded due to the gradient noise associated with the adaptation of filter $\mathbf{w}(k)$. However, because of the exponential decay of the energy of a typical RIR, the number of affected coefficients strongly depends on both the SNR and the specific shape of $\mathbf{h}(k)$. This suggests that further advantages could be obtained if we split the adaptive filter in different blocks, considering different regions of the impulse response. Thus, the estimation of each block can be biased by means of a mixing parameter, yielding $\lambda_m(k)\widehat{y}_m(k)$, improving the identification of the coefficients in the blocks that are affected by gradient noise, in case of $\lambda_m(k) < 1$. The multiplication by $\lambda_m(k)$ reduces the MSE of the estimation, giving rise to an enhanced overall cancellation performance.

Due to the time-variant adaptation of the mixing parameters (7), this echo canceller is able to achieve a suitable performance with respect to unknown or possibly time-varying SNR and/or $\mathbf{h}(k)$, without requiring any *a priori* knowledge about the cancellation scenario.

Obviously, the proposed scheme involves an increment in terms of computational cost with respect to the operation of a single adaptive filter. This increase mainly depends on the adaptation of mixing parameters according to (7), and therefore, it depends on number of blocks M. However, for moderate number of blocks, the computational cost associated with the adaptation of $\lambda_m(k)$ with $m = 1,...,M$ can be considered negligible in comparison to the cost required by the update of the adaptive filter. In this case, the computational cost of the proposed scheme corresponds, approximately, to that of a standard adaptive filter, $\mathbf{w}(n)$.

EXPERIMENTS

In this section, we present several experiments in order to show the advantages of the proposed scheme with respect to a canceller based on a standard adaptive filter $\mathbf{w}(k)$. Simulations have been carried out considering an unknown echo path $\mathbf{h}(k)$ truncated to $N = 512$ taps, whose impulse response and energy distribution are depicted in Fig. 3 using a sampling rate of 8 kHz. We will show results using two kinds of input signals: White noise with unit variance and real speech. Uncorrelated

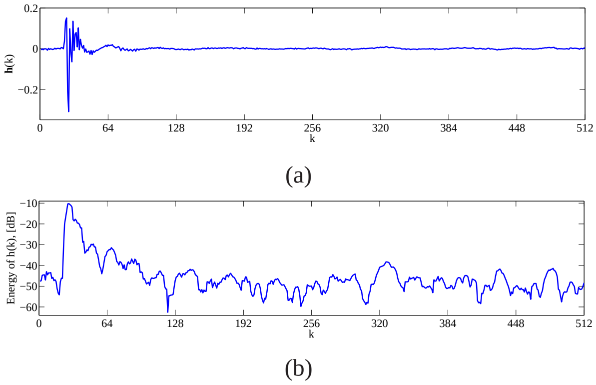


Figure 3: (a) Echo path employed in the experiments. (b) A smoothed representation of the energy distribution of $\mathbf{h}(k)$.

noise, $n(k)$, is added to the microphone signal, such that different SNR settings can be controlled.

Although our scheme can be implemented employing different kinds of adaptive algorithms, in this section we consider an adaptive IPNLMS filter whose length matches that of $\mathbf{h}(k)$. In this way, we will refer in the following to the proposed scheme as *blockwise biased IPNLMS* (BB-IPNLMS).

Each coefficient of the IPNLMS filter is adapted according to:

$$w_n(k+1) = w_n(k) + \frac{\mu g_n(k)}{\delta + \sum_{r=1}^{R} g_r(k) x_r^2(k)} e(k) x_n(k) \qquad (9)$$

where μ is the step size of the filter and R represents its length. In all experiments, μ has been set to 1. In (9), factors $g_n(k)$ define the adaptation gain for each weight, calculated by:

$$g_n(k) = (1-\kappa)\frac{1}{2N} + (1+\kappa)\frac{|w_n(k)|}{\varepsilon + 2\sum_{r=1}^{R} |w_r(k)|}. \qquad (10)$$

Although δ has been set to 0 in (9), ε in (10) is a small positive constant, that is equal to 10^{-6} for all of the simulations.

The IPNLMS filter has become a very popular solution during the last years; however, its behavior strongly depends on parameter κ. If $\kappa = -1$, (9) simply reduces to the adaptation rule of an NLMS filter, while $\kappa = 1$ turns the IPNLMS filter into a pure PNLMS. Intermediate values of κ give rise to mixed behaviors.

The decomposition into blocks (as outlined in previous section) establishes a compromise involving cancellation performance and additional complexity. Therefore, one of the main goals of this section is to present some insight about how to choose the appropriate number of blocks that allows a suitable echo cancellation, with a moderate increment in the number of ope-rations. Independently of the number of blocks M, we have chosen $\mu_a = 0.1$ for the adaptation of mixing parameters $\lambda_m(k)$.

White noise as input signal

Using white stationary noise as the input signal, we will first corroborate the advantages of our scheme for different values of SNR, κ, and number of blocks M. We will use the excess mean-square-error $\mathrm{EMSE}_c(k) = \mathrm{E}\{[e_c(k) - n(k)]^2\}$ as figure of merit, averaged over 100 independent realizations. In addition, we will employ $\mathrm{EMSE}(k) = \mathrm{E}\{[e(k) - n(k)]^2\}$ as a reference measurement corresponding to a standard filter with the same settings as the adaptive filter employed in our scheme.

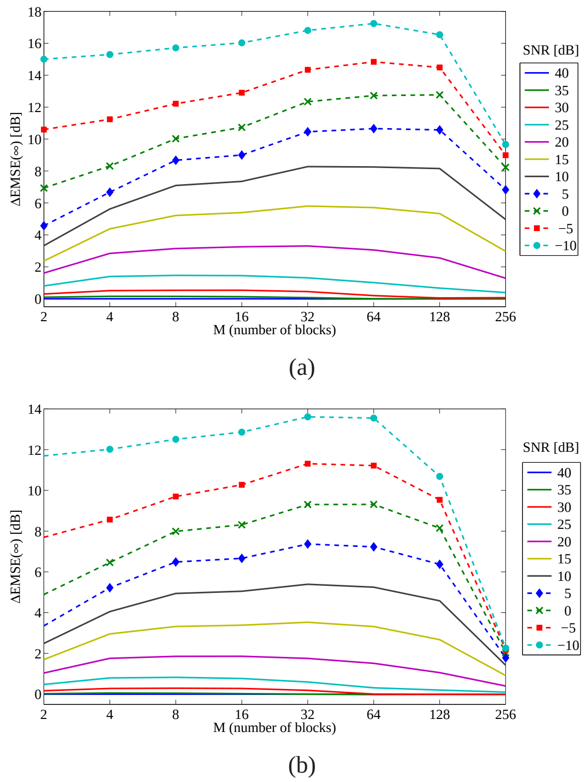


Figure 4: Echo cancellation gain reached by the BB-IPNLMS canceller with respect to a canceller based on a standard IPNLMS filter, in terms of $\Delta\mathrm{EMSE}(\infty) = \mathrm{EMSE}(\infty) - \mathrm{EMSE}_c(\infty)$. (a) $\kappa = -1$ (NLMS). (b) $\kappa = -0.5$.

Choosing number of blocks M

The steady-state performance of the BB-IPNLMS scheme has been studied for different numbers of blocks. To that end, Fig. 4 shows the gain of the proposed scheme with respect to a basic canceller based on a standard IPNLMS filter, i.e., $\Delta\mathrm{EMSE}(\infty) = \mathrm{EMSE}(\infty) - \mathrm{EMSE}_c(\infty)$, calculated over 25000 iterations after convergence of the algorithms. Moreover, we present results for two different values of κ, namely $\kappa = -0,5$ (lower plot) and $\kappa = -1$, that corresponds to an NLMS filter (upper plot).

According to the figure, it can be seen that our scheme offers important gains for low SNRs. For instance, for SNR = 5 dB, the BB-IPNLMS filter reaches a gain of around 6.5 dB and 9 dB, for $M = 16$ and $\kappa = -0.5$ and $\kappa = -1$, respectively. It is important to remark that, although the gain offered when SNR is high is negligible, no degradation of the behavior of $\mathbf{w}(k)$ is observed in this situation. For the sake of completeness, we have also included results for SNR < 0 dB , showing an increasing gain as compared to the standard IPNLMS. Negative SNRs can often be found, e.g., in in-car communication systems.

Regarding the selection of number of blocks M, similar conclusions can be drawn from both panels of Fig 4: The expressive capability of the scheme grows with the number of blocks, giving rise to incremental gains with respect to the non-divided case, for $M = 1$. However, if the length of a block is too small (see $M = 256$ corresponding to $P = 2$), this improvement is reduced due to noticeable gradient noise associated to the adaptation of $\lambda_m(k)$ with $m = 1, \ldots, M$. Furthermore, choosing very small block lengths would also yield important increment in the computational cost. For these reasons, $M = 16$ seems to be a reasonable selection since it offers interesting gains with negligible increment in terms of computational burden. Therefore, $M = 16$ will be employed for the rest of experiments.

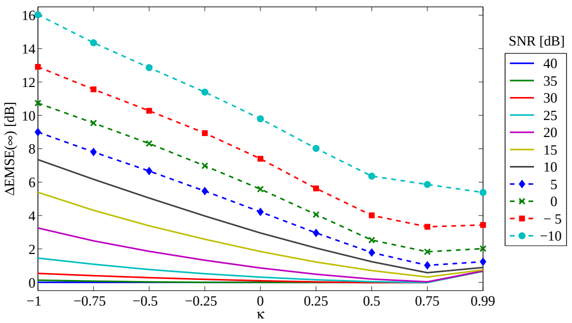


Figure 5: Echo cancellation gain reached by the BB-IPNLMS canceller ($M = 16$ blocks) with respect to a canceller based on a standard IPNLMS, in terms of $\Delta\text{EMSE}(\infty) = \text{EMSE}(\infty) - \text{EMSE}_c(\infty)$.

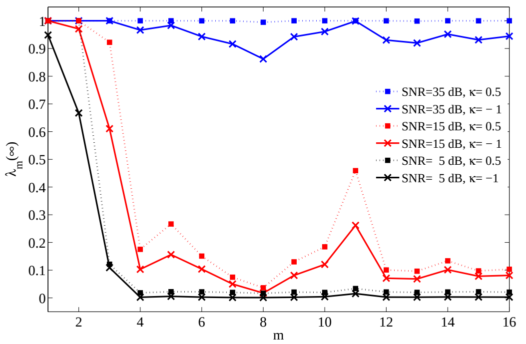


Figure 6: Steady-state value of the 16 mixing parameters, considering $\kappa = -1$ and 0.5; and for SNR = 35, 15 and 5 dB.

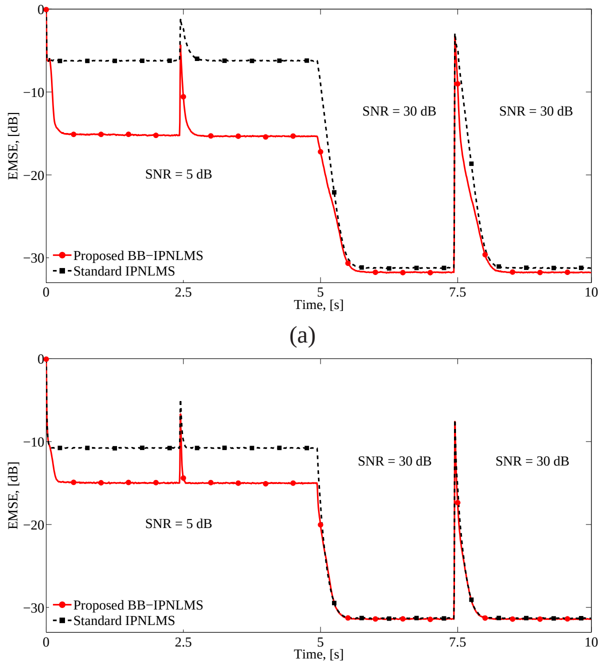
Influence of κ

Parameter κ establishes a compromise in the operation of the IPNLMS filter, making the adaptive filter behave closer to an NLMS filter ($\kappa = -1$) or to a PNLMS filter ($\kappa = 1$). In order to study this tradeoff, Fig. 5 shows the steady-state gain obtained for a canceller based on the BB-IPNLMS filter in terms of $\Delta\text{EMSE}(\infty)$, calculated exactly as in the previous subsection, as a function of κ and for different SNRs.

As it can be seen from Fig. 5, the proposed scheme improves the performance of the standard IPNLMS for all values of κ. However, the *amount* of the improvement depends on the selection of κ:

- If $\kappa \to -1$ (close to an NLMS filter), the BB-IPNLMS improves the performance of the non-divided filter $\mathbf{w}(k)$ noticeably. In this case, biasing the output of each block is convenient, especially for low SNRs, reducing the EMSE (see, for instance, $\kappa = -0.5$).
- However, for large κ, the adaptation speed for each tap of the IPNLMS filter $\mathbf{w}(k)$ is already independently adjusted to minimize the error, as can be seen in (9). In this case, splitting the whole filter offers less important gains (see, for instance, $\kappa = 0.75$).

In order to get a better understanding of the proposed scheme, Fig. 6 shows the steady-state value of the 16 mixing parameters considering $\kappa = -1$ and $\kappa = 0.5$, and for SNR = 35, 15 and 5 dB. As it can be seen, if the SNR is high, e.g. SNR = 35 dB, mixing parameters converge to values close to 1 since biasing is completely unnecessary and the whole scheme behaves as the adaptive IPNLMS (with $\kappa = -1$ or $\kappa = 0.5$). However, when the SNR decreases, the proposed scheme obtains an additional gain biasing output of some blocks $y_m(k)$. The specific value that each $\lambda_m(k)$ reaches directly depends on the energy distribution of unknown system $\mathbf{h}(k)$, that is represented in Fig. 3 (b), due to the constant power of the noise floor.



(a)

(b)

Figure 7: Cancellation performance of the proposed scheme (BB-IPNLMS) in terms of EMSE(k). (a) $\kappa = -1$ (NLMS filter). (b) $\kappa = 0$

By close inspection, another influence of κ can also be found: If κ grows, allowing a more independent adaptation of each tap in the m-th block, $\lambda_m(\infty)$ increases as a consequence of a smaller advantage from a biased $\hat{y}_m(k)$.

It is well known that IPNLMS filters with intermediate values of κ clearly outperform NLMS filters, especially with sparse unknown plants. However, it should be considered that, depending on the degree of sparseness which is *a priori* unknown, values of κ close to 1 may imply a degraded performance with respect to that of the NLMS filter. For these reasons, $\kappa = -0.5$ is a typical setting of IPNLMS (see Benesty and Gay 2002), and it is also a good selection for our proposed algorithm. Furthermore, using NLMS filters can also be an interesting choice, to benefit from the much smaller computational complexity of NLMS with respect to IPNLMS adaptive filter.

The convergence properties of the BB-IPNLMS canceller have also been studied for different values of κ, averaging the results of 1000 independent runs of the algorithm. Fig. 7 shows the EMSE evolution in a scenario where the SNR abruptly varies at $t = 5$ seconds from an initial value of 5 dB to 30 dB, in order to simulate an environment with an *a priori* unknown and time-varying SNR. In addition, in order to evaluate the reconvergence ability of the proposed scheme, the unknown echo path $\mathbf{h}(k)$ suddenly changes at $t = 2.5$ and $t = 7.5$ s.

As it can be seen in Fig. 7, the algorithm adapts to time-varying SNRs without requiring any *a priori* information. The experiment shows that our scheme obtains a better performance than a canceller based on a standard adaptive filter with identical κ, especially for low SNRs, not only under steady-state conditions but also when it reconverges (see panel (a) when $\kappa = -1$, i.e., NLMS filter).

Speech as input signal

We have also studied the performance of our proposal with respect to a canceller based on a standard IPNLMS filter using 13 seconds of real male speech as input signal. In this case the Echo Return Loss Enhancement (ERLE), defined as

$$\mathrm{ERLE}_{\mathrm{c}}(k) := 10\log_{10}\frac{\mathrm{E}\{[d(k)-n(k)]^2\}}{\mathrm{E}\{[e_c(k)-n(k)]^2\}}, \qquad (11)$$

is employed as figure of merit.

Fig. 8 shows the behavior of the algorithm for two different configurations of the proposed canceller using an IPNLMS filter with $\kappa = -0.5$ and an NLMS filter respectively. As we have explained before, for low SNRs the proposed canceller obtains better results, i.e., a higher ERLE(k), than a canceller based on a standard IPNLMS filter, see respectively panel (d) and (e). Furthermore, there is no degradation in the algorithm performance for high SNR, as it can be seen from both panel (b) and (c).

Conclusions and future work

In this paper we have presented a novel acoustic echo canceller especially convenient for low SNRs scenarios. Our scheme is based on a decomposition of an adaptive filter (for instance NLMS or IPNLMS) into non-overlapping blocks, whose partial outputs are adaptively combined with a *virtual* block of zeros. Following this approach, each output is adaptively biased reducing the residual echo when compared to the error that would be obtained by a canceller based on the same non-divided adaptive filter. Experimental results show that important enhancements can be reached with a moderate number of blocks, and, hence, with a negligible increment in terms of computational cost.

In contrast to other schemes, no *a priori* knowledge is necessary for the correct performance of the proposed canceller, and, furthermore, no degradation is introduced with high SNRs. In addition, the scheme holds for different kinds of adaptive filters, although using an NLMS or IPNLMS with small κ is a good option regarding performance and reliability.

Future work includes the extension of the block-wise decomposition to nonlinear filters, and the evaluation of the proposed scheme under impulse noise.

ACKNOWLEDGEMENT

This work has been partly supported by Spain Government under grant TEC2008-02473/TEC, and by the Deutsche Forschungsgemeinschaft (DFG) under contract number KE 890/5-1.

The authors thank Dr. J. Benesty for kindly supplying the real impulse response.

REFERENCES

J. Arenas-García and A. R. Figueiras-Vidal. Adaptive combination of proportionate filters for sparse echo cancellation. *IEEE Trans. on Speech and Audio Processing*, 17(6):1087–1098, August 2009.

J. Arenas-García, A. R. Figueiras-Vidal, and A. H. Sayed. Mean square performance of a convex combination of two adaptive filters. *IEEE Trans. on Signal Processing*, 54(3): 1078–1090, March 2006.

L. A. Azpicueta-Ruiz, M. Zeller, A. R. Figueiras-Vidal, J. Arenas-García, and W. Kellermann. Adaptive combination of volterra kernels and its application to nonlinear acoustic echo cancellation. *IEEE Trans. on Speech and Audio Processing*, in press. Available in http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5428855&isnumber=4358086.

L. A. Azpicueta-Ruiz, A. R. Figueiras-Vidal, and J. Arenas-García. A normalized adaptation scheme for the convex combination of two adaptive filters. In *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 3301–3304, Las Vegas, April 2008.

J. Benesty and S. L. Gay. An improved PNLMS algorithm. In *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 1881–1884, Orlando, May 2002.

R. Candido, M.T.M. Silva, and V.H. Nascimento. Affine combination of adaptive filters. In *Proc. 42nd Asilomar Conf. on Signal, Systems and Computers*, pages 236–240, Pacific Grove, October 2008.

D. L. Duttweiler. Proportionate normalized least mean square adaptation in echo cancelers. *IEEE Trans. on Speech and Audio Processing*, 8(5):508–518, September 2000.

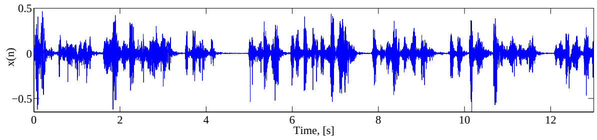
E. Hänsler and G. Schmidt. *Acoustic Echo and Noise Control: A Practical Approach*. Wiley, Hoboken, June 2004.

S. Kay and Y. C. Eldar. Rethinking biased estimation. *IEEE Signal Processing Magazine*, 25:133–136, 2008.

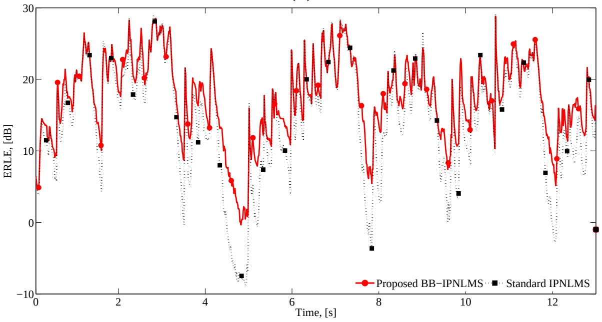
M. Lázaro-Gredilla, L. A. Azpicueta-Ruiz, A. R. Figueiras-Vidal, and J. Arenas-García. Adaptive biasing the weights of adaptive filters. *IEEE Trans. on Signal Processing*, in press. Available in http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5443561&isnumber=4359509.

S. Makino, Y. Kaneda, and N. Koizumi. Exponentially weighted step-size NLMS adaptive filter based on the statistics of a room impulse response. *IEEE Trans. on Speech and Audio Processing*, 1(1):101–108, January 1993.

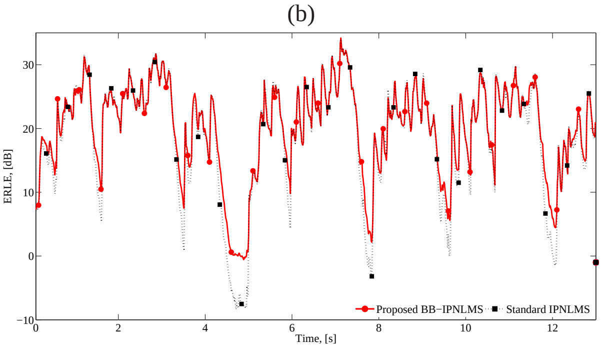
M. Zeller, L. A. Azpicueta-Ruiz, and W. Kellermann. Adaptive FIR filters with automatic length optimization by monitoring a normalized combination scheme. In *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 149–152, New Paltz (NY), October 2009.



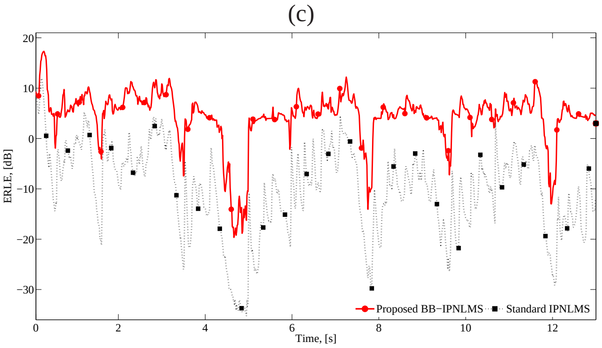
(a)



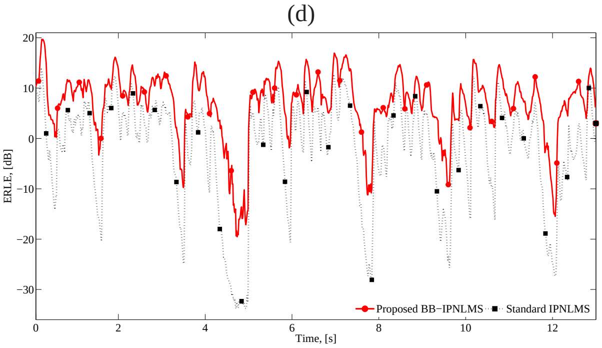
(b)



(c)



(d)



(e)

Figure 8: Performance of the proposed scheme (BB-IPNLMS) when using speech as input signal. (a) input signal. (b) ERLEs of a standard IPNLMS filter with $\kappa = -1$ (NLMS) and of the BB-IPNLMS canceller with same value κ when SNR = 35 dB. (c) Same than (a) but $\kappa = -0.5$. (d) Same than (b) but SNR = 5 dB. (e) Same than (c) but SNR = 5 dB.