A PROPORTIONATE ADAPTIVE ALGORITHM WITH VARIABLE PARTITIONED BLOCK LENGTH FOR ACOUSTIC ECHO CANCELLATION

Pradeep Loganathan, Emanuël A. P. Habets and Patrick A. Naylor

Department of Electrical and Electronic Engineering, Imperial College London, UK
Email: {pradeep.loganathan, e.habets, p.naylor}@imperial.ac.uk

ABSTRACT

Due to the nature of an acoustic enclosure, the early part (i.e., direct path and early reflections) of the acoustic echo path is often sparse while the late reverberant part of the acoustic path is normally dispersive. In order to account for this structure within the acoustic impulse response when performing acoustic echo cancellation, we propose an adaptive filter that consists of two time-domain partition blocks, with adaptive block partitioning, such that different adaptive algorithms can be used for each block. Specifically, the improved proportionate normalized least-mean-square (IPNLMS) algorithm is used. Simulation results show that the proposed variable length partitioned block IPNLMS (VLPB-IPNLMS) algorithm works well in both sparse and dispersive circumstances and in practical applications involving time-varying systems.

Index Terms— Sparse and dispersive impulse responses, Adaptive algorithms, Variable filter length, IPNLMS algorithm

1. INTRODUCTION

Time variation of the acoustic impulse response (AIR) arises in hands-free telephony, in which the user moves through significantly different acoustic environments during the call. For example, starting in an office, then moving into an elevator, then a lobby, out into the open air and finally into a car. During this call the level and nature of sparseness in the AIR to be identified both change significantly. Our main aim is to adaptively control the structure of the adaptive filter to deal effectively with such changes.

Initially, research on sparse adaptive filters [1] was aimed at network echo cancellation (NEC) while more recently [2, 3] these filters have been investigated in the context of acoustic echo cancellation (AEC). For both sparse and dispersive AIR, the early part of the echo path that consists of the direct path and a few early reflections is almost always sparse while the remaining late part is normally dispersive. We propose that normally it would be advantageous to use a sparse adaptive algorithm for the early part of the AIR and a non-sparse adaptive algorithm for the late part.

A single filter structure with two partitioned blocks was utilized in the partitioned block IPNLMS (PB-IPNLMS) [4] algorithm. It uses two IPNLMS algorithms each employing a different proportionate/non-proportionate factor α for the two corresponding time-domain partitioned blocks with a fixed size of L_1. As the AIR is time-varying, due to a change in temperature [5], pressure, changes in the acoustic environment and also changes in the loudspeaker microphone distance [3], the block size L_1 should be made time dependent for a robust convergence performance of PB-IPNLMS. A filter length control (FLC) algorithm was also addressed in [6]. However, the objective in [6] was to find the length of an adaptive filter that identifies a system with unknown and time-varying system memory.

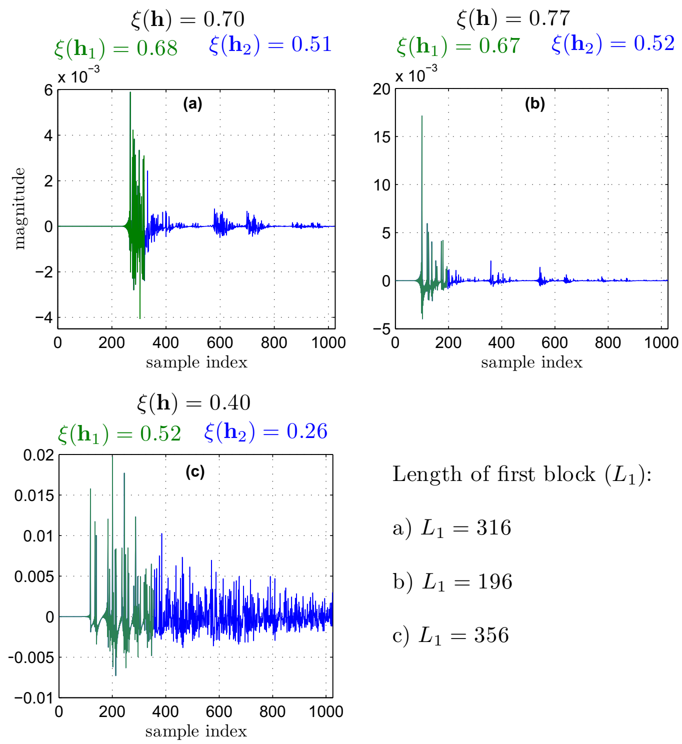

Fig. 1. Acoustic impulse responses obtained using the method of images [7]. $\xi(\mathbf{h}), \xi(\mathbf{h}_1)$ and $\xi(\mathbf{h}_2)$ respectively denote the sparseness measures [8] of the full impulse response, the first block with size L_1 and the second block.

In this work, our objective is to control the partition block length (L_1) by exploiting an efficient adaptive mechanism into PB-IPNLMS. We assume that the length of the system is known. The PB-IPNLMS algorithm is first reviewed in Section 2. We then show, in Section 3, how the sparseness of AIRs varies when we partition the echo path into two blocks with different sizes. Incorporating the findings, the proposed variable length partitioned block IPNLMS (VLPB-IPNLMS) algorithm is developed. Simulation results shown in Section 4, compare the tracking performance for both sparse and dispersive AIRs in the context of AEC.

2. REVIEW OF PB-IPNLMS

Defining the tap-input signal as $\mathbf{x}(n) = [x(n)\ x(n-1)\ \ldots\ x(n-L+1)]^T$, with length L, the unknown impulse response as $\mathbf{h}(n) = [h_0(n)\ h_1(n)\ \ldots\ h_{L-1}(n)]^T$ and the additive noise as $\nu(n)$, the output signal is expressed as

$$y(n) = \mathbf{h}^T(n)\mathbf{x}(n) + \nu(n). \tag{1}$$

The general computations of the adaptive filter coefficients $\widehat{\mathbf{h}}(n) = [\widehat{h}_0(n)\ \widehat{h}_1(n) \ldots \widehat{h}_{L-1}(n)]^T$ for many adaptive algorithms can be described by the following equations:

$$e(n) = y(n) - \widehat{\mathbf{h}}^T(n-1)\mathbf{x}(n), \tag{2}$$

$$\widehat{\mathbf{h}}(n) = \widehat{\mathbf{h}}(n-1) + \frac{\mu\mathbf{Q}(n-1)\mathbf{x}(n)e(n)}{\mathbf{x}^T(n)\mathbf{Q}(n-1)\mathbf{x}(n) + \delta}, \tag{3}$$

$$\mathbf{Q}(n-1) = \mathrm{diag}\{q_0(n-1)\ \ldots\ q_{L-1}(n-1)\}, \tag{4}$$

where μ is a step-size and δ is the regularization parameter. The diagonal step-size control matrix $\mathbf{Q}(n)$ determines the step-size of each filter coefficient and is dependent on the specific algorithm.

2.1. The IPNLMS algorithm

The IPNLMS algorithm [9] employs a combination of non-proportionate (i.e., NLMS with $\mathbf{Q}(n) = \mathbf{I}_{L\times L}$) and proportionate (i.e., PNLMS [10] with $q_l(n) \propto |\widehat{h}_l(n)|$) adaptation, with the relative significance of each controlled by a factor α such that the diagonal elements of $\mathbf{Q}(n)$ are given by

$$q_l(n) = \frac{1-\alpha}{2L} + \frac{(1+\alpha)|\widehat{h}_l(n)|}{2\|\widehat{\mathbf{h}}(n)\|_1 + \delta_{\mathrm{IP}}}, \qquad 0 \le l \le L-1. \tag{5}$$

where $\|\cdot\|_1$ is the ℓ_1-norm. It can be seen that IPNLMS reduces to NLMS when $\alpha = -1$ and PNLMS when $\alpha = 1$. For most AEC applications, $\alpha = 0$, -0.5 or -0.75 are favorable choices [9]. To avoid any problems during regularization of the algorithm, the diagonal elements of $\mathbf{Q}$ are normalized such that $\mathrm{tr}\{\mathbf{Q}\} \approx 1$ [2]. It has been shown that, regardless of the impulse response nature, the IPNLMS algorithm has faster convergence than NLMS and PNLMS for the same level of asymptotic convergence with the above choices of α [9]. However, we note from our simulations that the choice of α influences the tracking performance of IPNLMS for sparse and dispersive AIRs.

2.2. Block partitioned structure

As shown in Section 1 and in [4], the echo path $\mathbf{h}(n)$ can advantageously be partitioned into two blocks, such that the first block ($\mathbf{h}_1(n)$) with length L_1 includes the direct path and a few early reflections, which is more sparse than the second block ($\mathbf{h}_2(n)$) that includes all other reflections. Hence giving

$$\mathbf{h}(n) = [\mathbf{h}_1^T(n)\ \ \mathbf{h}_2^T(n)]^T, \tag{6}$$

with

$$\begin{aligned} \mathbf{h}_1(n) &= [h_0(n)\ \ldots\ h_{L_1-1}(n)]^T, & (7)\\ \mathbf{h}_2(n) &= [h_{L_1}(n)\ \ldots\ h_{L-1}(n)]^T, & (8) \end{aligned}$$

so that any sparse adaptive algorithm can be employed to estimate $\mathbf{h}_1$ and a non-sparse algorithm can be used to estimate $\mathbf{h}_2$.

2.3. The PB-IPNLMS algorithm

The partitioned block IPNLMS (PB-IPNLMS) algorithm [4] exploits IPNLMS with the mixing parameter α_1 close to 1 as the sparse algorithm for the first block of length L_1. And for the second block, as it is more dispersive compared to the first block, it uses IPNLMS with the mixing parameter α_2 ($\alpha_2 < \alpha_1$) close to -1. The specific equations to compute the diagonal elements for the two blocks are

$$\begin{aligned} q_l(n) &= \frac{(1-\alpha_1)}{2L_1} + \frac{(1+\alpha_1)|\widehat{h}_l(n)|}{2\|\widehat{\mathbf{h}_1}(n)\|_1 + \delta_{\mathrm{IP}}}, \\ &\qquad\qquad 0 \le l \le L_1 - 1, & (9)\\ \mathbf{Q}_1(n-1) &= \mathrm{diag}\{q_0(n-1), \ldots, q_{L_1-1}(n-1)\}, & (10)\\ q_l(n) &= \frac{(1-\alpha_2)}{2(L-L_1)} + \frac{(1+\alpha_2)|\widehat{h}_l(n)|}{2\|\widehat{\mathbf{h}_2}(n)\|_1 + \delta_{\mathrm{IP}}}, \\ &\qquad\qquad L_1 \le l \le L-1, & (11)\\ \mathbf{Q}_2(n-1) &= \mathrm{diag}\{q_{L_1}(n-1), \ldots, q_{L-1}(n-1)\}. & (12) \end{aligned}$$

As for the case of conventional IPNLMS, the constraint on $\mathrm{tr}\{\mathbf{Q}(n)\}$ of PB-IPNLMS, which is composed of $\mathbf{Q}_1(n)$ and $\mathbf{Q}_2(n)$, still needs to be 1 for very small values of δ_{IP}. As $\mathrm{tr}\{\mathbf{Q}_1\} \approx \mathrm{tr}\{\mathbf{Q}_2\} \approx 1$, the composed step-size control matrix $\mathbf{Q}(n-1)$ can be computed as

$$\mathbf{Q}(n-1) = \begin{bmatrix} 0.5\ \mathbf{Q}_1(n-1) & \mathbf{0}_{L_1\times(L-L_1)} \\ \mathbf{0}_{(L-L_1)\times L_1} & 0.5\ \mathbf{Q}_2(n-1) \end{bmatrix}. \tag{13}$$

It is important to note that the early studies on this block partitioned approach in [4] proposed to keep L_1 constant throughout the simulation. The next section significantly extends the early study to the important and more practical case in which the echo path is time varying, such as in a hands-free telephone call. The technique presented for adaptive block partitioning improves the robustness of convergence of PB-IPNLMS in a situation requiring identification of time-varying echo path.

3. THE VARIABLE LENGTH PARTITIONED BLOCK IPNLMS (VLPB-IPNLMS) ALGORITHM

We now show how the AIRs can be partitioned into two blocks adaptively, so that the first block, with the dominant direct path and early reflections, is more sparse than that of the second block. This serves as a motivation for us to develop a new algorithm which improves the robustness and the tracking performance compared to IPNLMS.

3.1. Motivation

As shown in Section 2.3, the composed step-size control matrix in (13) satisfies the constraint on $\mathrm{tr}\{\mathbf{Q}(n)\}$, even though this constraint can be satisfied in many other ways. To motivate this work we define

$$\lambda(n|L_1) = \frac{\|\mathbf{h}_1(n)\|_1}{\|\mathbf{h}(n)\|_1}, \tag{14}$$

and show PB-IPNLMS performance in Fig. 2 for different λ values, which is controlled by varying L_1, to illustrate the time taken to reach -20 dB normalized misalignment defined by

$$\eta[\mathbf{w}(n), \widehat{\mathbf{w}}(n)] = \frac{\|\mathbf{w}(n) - \widehat{\mathbf{w}}(n)\|_2^2}{\|\mathbf{w}(n)\|_2^2}. \tag{15}$$

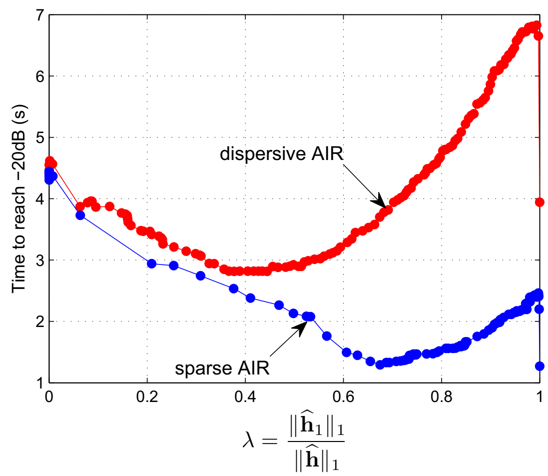

Fig. 2. Time to reach -20 dB normalized misalignment level for different values of λ in (14) for PB-IPNLMS using the input signal generated by (16). Impulse response in Fig. 1 (b) and (c) were used as sparse AIR and dispersive AIR respectively. [$\mu = 0.3$, SNR = 20 dB]

The input signal was generated using an AR(2) process given by [1]

$$x(n) = 0.4\, x(n-1) - 0.4\, x(n-2) + s(n), \qquad (16)$$

where $s(n)$ is a white Gaussian noise (WGN) with $\sigma_s^2 = 0.77$. Another WGN sequence was chosen for $\nu(n)$ such that an signal-to-noise ratio (SNR) of 20 dB was obtained. A step-size of $\mu = 0.3$ was used in this experiment. We see from the result that, the non-proportionate composition of $\mathbf{Q}$ in (13) works well only if

$$\lambda(n) \approx 0.5, \qquad (17)$$

for both sparse and dispersive AIRs. It is also interesting to note that, when the ratio is 0, PB-IPNLMS employs the non-sparse algorithm. If the ratio is 1, it employs the sparse algorithm. In addition, We note from the result that, for each case of λ, PB-IPNLMS has a higher rate of convergence for a sparse system compared to a dispersive system. This is due to the initialization choice of $\widehat{\mathbf{h}}(0) = \mathbf{0}_{L\times 1}$, where most filter coefficients are initialized close to their optimal values.

We can adopt the condition in (17) as a technique to partition the block, so that the first block contains the direct path and few early reflections. The AIRs in Fig. 1 were partitioned into two blocks using different sizes for the first block with length L_1, giving (a) $L_1 = 316$, (b) $L_1 = 196$ and (c) $L_1 = 356$ respectively. The three AIRs of length $L = 1024$ were simulated under different conditions using the method of images [7] in a loudspeaker-room-microphone system (LRMS) at a sampling frequency of 8 kHz. Figure 1(a) and (c) show the AIRs obtained when the loudspeaker-microphone distances, a, are 11.4 m and 5 m in a room of dimension $8 \times 10 \times 3$ m, with 0.2 and 0.53 as the reflection coefficients, respectively. Figure 1(b) illustrates the AIR attained when $a = 4.2$ m in a room of dimension $10 \times 15 \times 3$ m, with 0.2 as the reflection coefficient. The sparseness measures of these AIRs are computed using [8]

$$\xi(\mathbf{w}) = \frac{N}{N - \sqrt{N}} \left\{ 1 - \frac{\|\mathbf{w}\|_1}{\sqrt{N}\, \|\mathbf{w}\|_2} \right\}, \qquad (18)$$

where N is the length of the vector $\mathbf{w}$ and $\|\mathbf{w}\|_1$ and $\|\mathbf{w}\|_2$ represent ℓ_1 and ℓ_2-norms of $\mathbf{w}$. As can be seen from the sparseness measures of the first blocks ($\xi(\mathbf{h}_1)$) and second blocks ($\xi(\mathbf{h}_2)$), the first blocks in all cases are more sparse than the second blocks. As a consequence of this important finding, we propose a new adaptation technique for PB-IPNLMS with an efficient mechanism to make L_1 time-dependent, as described below.

3.2. Automatic control of the block length L_1

An example set of AIRs illustrated in Fig. 1(a)-(c) show that in order to satisfy the condition in (17), L_1 varies depending on the initial bulk delays and the overall sparseness levels of the AIRs. An efficient mechanism for the automatic control of $L_1(n)$ can be derived according to the ratio between $\|\widehat{\mathbf{h}}_1(n)\|_1$ and $\|\widehat{\mathbf{h}}(n)\|_1$ as follows

$$L_1(n) = \begin{cases} L/4, & n < L \\ L_1(n-1) + \triangle\ell, & \frac{\|\widehat{\mathbf{h}}_1(n)\|_1}{\|\widehat{\mathbf{h}}(n)\|_1} < \kappa_{\min} \\ L_1(n-1) - \triangle\ell, & \frac{\|\widehat{\mathbf{h}}_1(n)\|_1}{\|\widehat{\mathbf{h}}(n)\|_1} > \kappa_{\max} \\ L_1(n-1), & \text{otherwise} \end{cases} \qquad (19)$$

where $\triangle\ell$ denotes the number of coefficients by which $L_1(n)$ can be enlarged/reduced. The minimum and maximum threshold values $\kappa_{\min}$ ($0 \ll \kappa_{\min} < 0.5$) and $\kappa_{\max}$ ($0.5 < \kappa_{\max} \ll 1$) can be specified in order to define the region to satisfy (17). With the formulation in (19), L_1 initializes to $L/4$ and kept as constant for $n < L$. We enlarge the first block size L_1, when the ratio between $\|\widehat{\mathbf{h}}_1(n)\|_1$ and $\|\widehat{\mathbf{h}}(n)\|_1$ is less than a threshold value $\kappa_{\min}$, resulting adding some more early reflections into the first block. On the other hand, L_1 reduces to exclude some of the weaker reflections when the ratio is greater than $\kappa_{\max}$. Otherwise, L_1 stays the same. Thus, the proposed variable length partitioned block IPNLMS (VLPB-IPNLMS) algorithm is described by (2), (3), (9)-(13) and (19).

4. PERFORMANCE EVALUATION

We present simulation results to evaluate the performance of the proposed VLPB-IPNLMS algorithm. Throughout our simulations, algorithms were tested using the input signal generated by (16) while $\nu(n)$ was a WGN sequence chosen such that an SNR of 20 dB was obtained. We assumed that the length of the adaptive filter $L = 1024$ is equivalent to that of the unknown system. Three receiving room impulse responses, simulated using the method of images [7], for AEC simulations have been used, with an echo path changed from that shown in Fig. 1 (a) to (b) and then to (c) and μ = 0.3. For PB-IPNLMS and VLPB-IPNLMS, $\alpha_1 = 0.9$ and $\alpha_2 = -1$ were used, while $\triangle\ell = 10$, $\kappa_{\min} = 0.45$ and $\kappa_{\max} = 0.65$ were employed specifically for the VLPB-IPNLMS algorithm. The threshold values were chosen from the results shown in Fig. 2.

Figure 3 (a) compares the overall performance of IPNLMS with $\alpha = -1$ and 0.9, PB-IPNLMS using the proportionate weight allocation technique [4] and VLPB-IPNLMS, in terms of normalized misalignment. As it can be seen that the proposed VLPB-IPNLMS achieves both a high rate of initial convergence similar to IPNLMS with $\alpha = 0.9$ for the identification of the first two AIRs and also approximately 3 dB better convergence performance for the identification of the third (dispersive) AIR. Moreover, VLPB-IPNLMS gives more than 7 dB better initial convergence performance and similar steady-state performance compared to IPNLMS with $\alpha = -1$ for

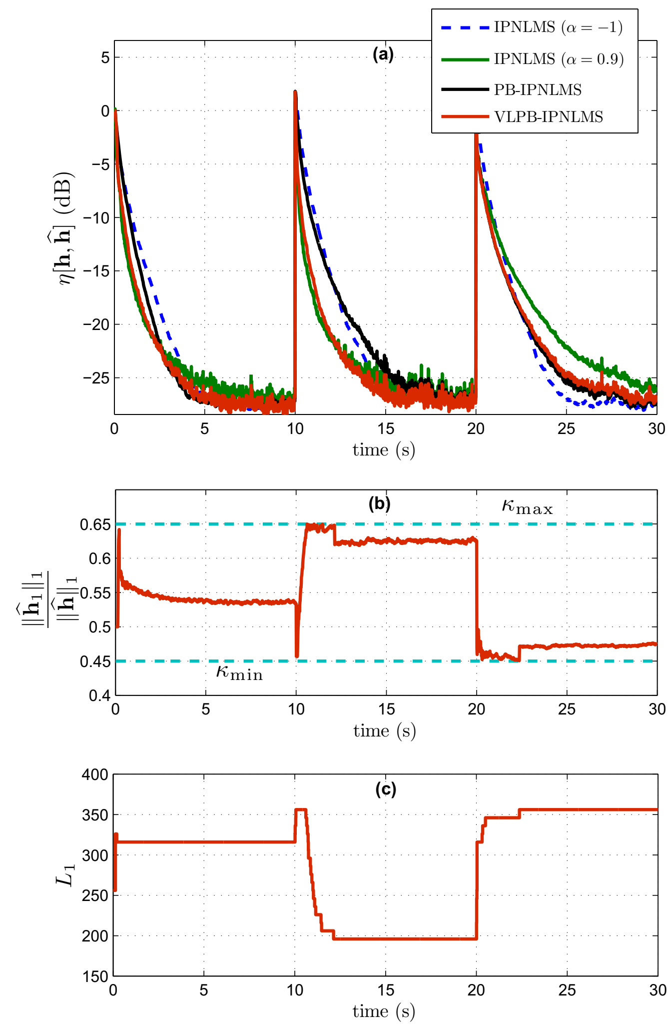

Fig. 3. **a)** Relative convergence of IPNLMS for $\alpha = -1$ and 0.9, PB-IPNLMS with proportionate weight allocation technique [4] and VLPB-IPNLMS; **b)** evolution of $\frac{\|\widehat{\mathbf{h}}_1(n)\|_1}{\|\widehat{\mathbf{h}}(n)\|_1}$; **c)** evolution of L_1; The input signal was generated using (16) with an echo path change at 10 s and 20 s, from that shown in Fig. 1 (a) to (b) and then to (c), μ = 0.3 and SNR = 20 dB.

the first two sparse AIRs and almost equivalent convergence performance for the dispersive AIR. Notably, VLPB-IPNLMS achieves more than 5 dB faster initial convergence compared to PB-IPNLMS for the sparse AIRs and similar performance for the dispersive AIR.

The VLPB-IPNLMS algorithm achieves this better performance by varying the block size L_1, as illustrated in Fig. 3 (c), such that the first block of the AIR with the dominant parts of the echo path gets larger individual step-sizes and therefore achieves faster initial convergence performance. Moreover, the second partitioned block is set to have almost equal step-sizes and as a result attains better steady-state performance. Here, the adaptation of the L_1 was automatically controlled by the evolution of the ratio shown in Fig. 3 (b).

5. CONCLUSION

In practical scenarios of acoustic echo cancellation for hands-free mobile telephony devices, the level of sparseness in the AIR can be highly variable. To deal with this issue, we developed a partitioned block IPNLMS algorithm with a control mechanism for the dynamic adjustment of the block size. The proposed algorithm achieves improved convergence compared to classical IPNLMS with fixed single proportional/non-proportionate factor α. For the proposed VLPB-IPNLMS, we incorporated a self-configuration method based on the ratio between the ℓ_1-norm of the first block's estimated filter coefficients and that of the overall filter coefficient into PB-IPNLMS for AEC to achieve faster convergence for both sparse and dispersive acoustic echo paths with variable bulk delay.

6. REFERENCES

[1] F. das Chagas de Souza, O. Tobias, R. Seara, and D. Morgan, "A PNLMS algorithm with individual activation factors," *IEEE Trans. Signal Processing*, vol. 58, pp. 2036–2047, 2010.

[2] C. Paleologu, J. Benesty, and S. Ciochina, *Sparse Adaptive Filters for Echo Cancellation*. Morgan & Claypool Publishers, 2010.

[3] P. Loganathan, A. W. H. Khong, and P. A. Naylor, "A class of sparseness-controlled algorithms for echo cancellation," *IEEE Trans. Audio, Speech, and Language Processing*, vol. 17, no. 8, pp. 1591 – 1601, Nov. 2009.

[4] P. Loganathan, E. A. P. Habets, and P. A. Naylor, "A partitioned block proportionate adaptive algorithm for acoustic echo cancellation," in *(to appear) Proceedings of APSIPA Annual Summit and Conference 2010*, Biopolis, Singapore, Dec 2010.

[5] G. W. Elko, E. Diethorn, and T. Gänsler, "Room impulse response variation due to thermal fluctuation and its impact on acoustic echo cancellation," in *Proc. Int. Workshop on Acoustic Echo and Noise Control*, 2003, pp. 67–70.

[6] M. Zeller, L. Azpicueta-Ruiz, and W. Kellermann, "Adaptive FIR filters with automatic length optimization by monitoring a normalized combination scheme," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. WASPAA '09*, Oct. 2009, pp. 149 –152.

[7] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Amer.*, vol. 65, no. 4, pp. 943–950, Apr. 1979.

[8] J. Benesty, Y. A. Huang, J. Chen, and P. A. Naylor, "Adaptive algorithms for the identification of sparse impulse responses," in *Selected methods for acoustic echo and noise control*, E. Hänsler and G. Schmidt, Eds. Springer, 2006, ch. 5, pp. 125–153.

[9] J. Benesty and S. L. Gay, "An improved PNLMS algorithm," in *Proc. IEEE Int. Conf. Acoustics Speech Signal Processing*, vol. 2, 2002, pp. 1881–1884.

[10] D. L. Duttweiler, "Proportionate normalized least mean square adaptation in echo cancellers," *IEEE Trans. Speech Audio Processing*, vol. 8, no. 5, pp. 508–518, Sep. 2000.