A Collaborative Filter Approach to Adaptive Noise Cancellation

Michele Scarpiniti, Danilo Comminiello, Raffaele Parisi, and Aurelio Uncini

Department of Information Engineering, Electronics and Telecommunications (DIET),
"Sapienza" University of Rome, Via Eudossiana 18, 00184, Rome
{michele.scarpiniti,danilo.comminiello,
raffaele.parisi}@uniroma1.it, aurel@ieee.org

Abstract. In this paper we propose a filter combination for the adaptive noise cancellation (ANC) problem in nonlinear environment. The architecture consists in a convex combination of two adaptive filters: a classical filter and a nonlinear filter based on Functional Links. While the convergence of the linear filter is very fast, the convergence of the nonlinear one might be slower, even if it provides a more accurate solution. The convex combination of both filters allows to reach good performances in terms of convergence and speed. In addition a variable step size is used in order to obtain better performance. Several experimental results, in different reverberant conditions, demonstrate the effectiveness of the proposed approach.

Keywords: Noise Cancellation, Adaptive Filters, Functional Link Network, Convex Combination, Variable Stepsize.

Introduction

Adaptive Noise Cancellation (ANC) is an evergreen application useful in a wide range of scenarios [1], such as conference rooms. Although ANC architectures are well known in literature and applications, some issues are still open. In fact, in real-time applications two main contrasting characteristics are very important: high convergence speed and good convergence performance. These characteristics can be controlled by the stepsize used in adaptive algorithms. Unfortunately a small value of this parameter provides good convergence performance, i.e. small error for the obtained clean signal, but results in a slow convergence speed. In contrast a large value for the stepsize provides a fast convergence but poor quality. In order to meet these conflicting requirements, the stepsize needs to be controlled, for example, by sequentially scaling its value [2,3]. In addition, signals are often deteriorated by nonlinear distortions, due to low-cost audio equipment, such as amplifiers and loudspeakers, or to vibrating structures and chests. These distortions usually neutralize the work of linear ANC systems, making them useless in most applications.

Recently a convex combination filter approach was proposed [4,5,6]. This approach can simply solve the ambiguity on the choice of the stepsize. In fact, the adaptive filter results in a combination of two filters, the first with a large stepsize and the latter with

B. Apolloni et al. (Eds.): *Neural Nets and Surroundings*, SIST 19, pp. 101–109.
DOI: 10.1007/978-3-642-35467-0_11

a small one. This solution overcomes the dichotomy between speed and convergence: while the second filter provides good convergence performances, the first one speeds up the convergence of the overall filter [5]. More recently such a combination was successfully applied to nonlinear acoustic echo cancellation problems [7,8].

In this paper we propose the use of this convex filter combination to implement an ANC architecture that can exploit both positive improvement of such an architecture: achieving good performance in a small amount of time and recovering the nonlinear distortion. In addition a variable stepsize is used in order to enhance the performance obtained by the algorithm [3]. These two adaptive filters collaborate toward the convergence: in this way we name this architecture as *collaborative* ANC (CANC).

Several experimental results conducted in an office environment varying the reverberation time T_{60} in a wide range, demonstrate the effectiveness of the proposed approach.

This paper is organized as follows: section 1 describes the proposed architecture, while section 2 shows some experimental results. Finally section 3 draws our conclusions.

1 The Proposed Architecture

The standard ANC architecture consists in an adaptive filter trained by a a desired signal $d[n] = s_f[n] + r[n]$, where $s_f[n] = s[n] * h_s[n]$, $s[n]$ is the desired clean speech signal, $h_s[n]$ is the room impulse response (RIR) between the speech source and the microphone, and $r[n]$ is the background noise captured by the microphone. The filter input signal $x[n]$ is a version of the background noise not altered by the environmental effects: the microphone recording $x[n]$ must be placed close to the noise source $b[n]$, $x[n] = b[n]$. The contribution to the desired signal is $r[n] = b[n] * h_n[n]$, where $h_n[n]$ is the RIR between the noise source and the microphone. At convergence the adaptive filter coefficients must be an estimation of this impulse response. In addition we suppose that the signal received by the microphones is distorted, due to a non ideal behavior of the transducer and its electronic equipment.

The collaborative ANC, or CANC, consists in a convex combination of a linear adaptive filter (LAF) and a nonlinear filter, implemented by a Functional Link Network (FLN). In this way the proposed system is constituted by the architecture shown in Figure 1. A single FLN adaptive filter was already used for ANC in [9], obtaining promising results. In order to consider alway present the path due to the linear filter, the combination is performed by summing the output of the linear filter with a scaled version of the output of the nonlinear filter:

$$y[n] = y_{\mathrm{L}}[n] + y_{\mathrm{NL}}[n] = y_{\mathrm{L}}[n] + \lambda[n] y_{\mathrm{FL}}[n], \tag{1}$$

where $0 \leq \lambda \leq 1$, $y_{\mathrm{L}}[n]$ is the output of the linear filter, evaluated simply as

$$y_{\mathrm{L}}[n] = \mathbf{w}_n^T \cdot \mathbf{x}_n, \tag{2}$$

where $\mathbf{w}_n = [w_1, \ldots, w_L]^T$ is the vector of the L coefficients for the linear filter at instant n and $\mathbf{x}_n = [x[n], x[n-1], \ldots, x[n-L+1]]^T$.

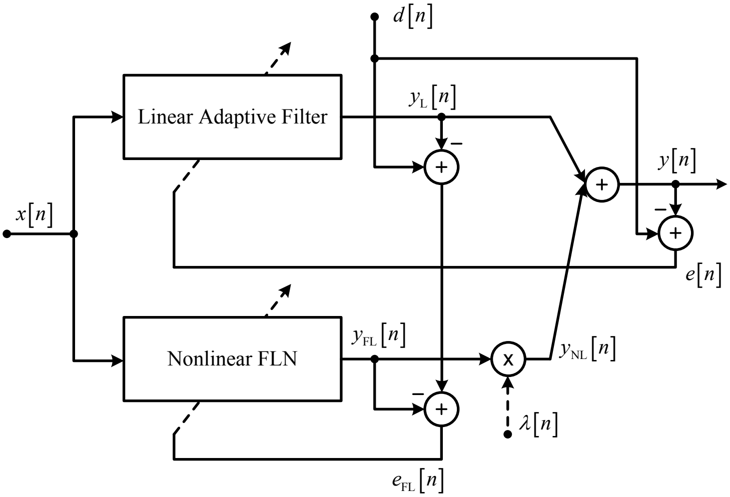

Fig. 1. The proposed collaborative combination of adaptive filters

In [10] the Functional Link Network (FLN) is described as an artificial neural network (ANN) with a single layer, that is able to process an input signal using a functional expansion block. This block consists of a series of linearly independent functions that accept the input signal as argument. The functional expansion projects the input in a higher dimension space.

The functional expansion used in [7] is a trigonometric series expansion composed of both linear and nonlinear elements. However, in order to implement a purely nonlinear kernel, we will consider that the enhanced pattern $\mathbf{z}[n]$ is composed just by some nonlinear transformations of the original variables:

$$\mathbf{z}[n] = \left[1, \mathbf{z}_0^T[n], \ldots, \mathbf{z}_{L_{in}-1}^T[n]\right]^T, \tag{3}$$

where $\mathbf{z}_i[n] = [z_{i,0}[n], \ldots, z_{i,Q-1}[n]]$, whose elements $z_{i,j}[n]$ assume the following values:

$$z_{i,j}[n] = \begin{cases} \sin(p\pi x[n-i]) & \text{for } j = 2p \\ \cos(p\pi x[n-i]) & \text{for } j = 2p+1 \end{cases} \tag{4}$$

in which $p = 0, 1, \ldots, P-1$, being P the expansion order and $x[n]$ represents the input sample at n-th instant, while $i = 0, 1, \ldots, L_{in} - 1$ is the input sample index and L_{in} is the length of the input signal block; finally $j = 0, 1, \ldots, Q-1$ is the enhanced pattern index and $Q = 2P$ is the number of functional links for the i-th input. Therefore, the enhanced pattern is composed of $L_{en} = 2QL_{in} + 1$ elements.

It is now possible to write the output of the nonlinear FLAF simply at the instant n as the output of a FIR filter:

$$y_{\mathrm{NL}}[n] = \mathbf{h}_n^T \cdot \mathbf{z}_n, \tag{5}$$

where $\mathbf{h}_n = [h_0, h_1, \ldots, h_{L_{en}-1}]^T$.

1.1 Adaptation of the Proposed Architecture

Let us pose $e[n] = d[n] - y[n]$ the overall error of the proposed architecture. Then the adaptation of each filter is performed minimizing the overall square error $e^2[n]$ by using the classical *Normalized Least Mean Square* (NLMS) algorithm [11]:

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu_{\mathrm{L}}[n] \frac{e[n]\mathbf{x}_n}{\delta + \mathbf{x}_n^T \mathbf{x}_n}, \tag{6}$$

$$\mathbf{h}_{n+1} = \mathbf{h}_n + \mu_{\mathrm{NL}}[n] \frac{e[n]\mathbf{x}_n}{\delta + \mathbf{z}_n^T \mathbf{z}_n}, \tag{7}$$

where δ is a regularization parameter.

In order to take into account the under-modeling scenario we adopt a variable step-size (VSS) [3]. The variable stepsize parameter $\mu_j[n]$ in (6) and (7), after M iterations, is chosen according to [12], :

$$\mu[n] = \begin{cases} \mu_f, & n \leq M \\ \left| 1 - \frac{\sqrt{\left|\hat{\sigma}_d^2[n] - \hat{\sigma}_y^2[n]\right|}}{\hat{\sigma}_e^2[n] + \xi} \right|, & n > M \end{cases} \tag{8}$$

where ξ is a regularization parameter, and the general parameter $\hat{\sigma}_\theta^2[n]$ represents the power estimate of the sequence $\theta[n]$:

$$\hat{\sigma}_\theta^2[n] = \gamma \hat{\sigma}_\theta^2[n-1] + (1-\gamma)\theta^2[n], \tag{9}$$

where $\gamma = 1 - 1/(kM)$ is a weighted factor with $k > 1$. The initial value is $\hat{\sigma}_\theta^2[0]$ and is set to 1.

In order to reduce the gradient noise and to keep the mixing parameter in the range $(0,1)$, the adaptation of $\lambda[n]$ can be carried out through the adaptation of another parameter, $a[n]$, related to $\lambda[n]$ by the following equation:

$$\lambda[n] = \left(1 + e^{-a[n]}\right)^{-1}. \tag{10}$$

The update of $a[n]$ is then given by [8]:

$$a[n] = a[n-1] + \frac{\mu_a}{p[n]} \lambda[n](1-\lambda[n]) e[n] \varepsilon[n], \tag{11}$$

where $\varepsilon[n] = y_1[n] - y_2[n]$; the term $p[n] = \beta p[n-1] + (1-\beta)\varepsilon^2[n]$ is the estimated power of $\varepsilon[n]$, and β is a threshold close to one [13].

The proposed architecture is robust to any nonlinearity level, since when the convolutive path is merely linear $\lambda[n]$ will converge towards 0 and the whole scheme will behave like a purely linear filter, thus avoiding any gradient noise from the nonlinear filter. On the other hand, when the convolutive path presents nonlinearities the mixing parameter will approach 1 according to the nonlinearity level in the path. Moreover, after convergence, the value of $\lambda[n]$ represents a measure of the degree of nonlinearity present in the system.

2 Experimental Results

The experimental tests are conducted in an office room of standard dimension of $6 \times 4 \times 3$ m. The location of microphones and sources is depicted in Figure 2. Different reverberation times are simulated in the range $0 \div 350$ ms, in order to test the performance of the proposed architecture versus the reverberation time. The room impulse responses are simulated using the Matlab tool RoomSim[1] [14]. The performance of the proposed algorithm is measured in terms of output SNR, defined by the following equation:

$$\mathrm{SNR}_{out} = 10\log_{10} \frac{E\left\{s_f^2[n]\right\}}{E\left\{(r[n]-y[n])^2\right\}}. \tag{12}$$

Eq. (12) is justified by the fact that

$$e[n] = d[n] - y[n] = (s[n] * h_s[n] + x[n] * h_n[n]) - y[n] = s_f[n] + (r[n] - y[n]),$$

where the first term in the right side is the desired clean speech signal, while the terms in round brackets is the residual noise, that converges to zero.

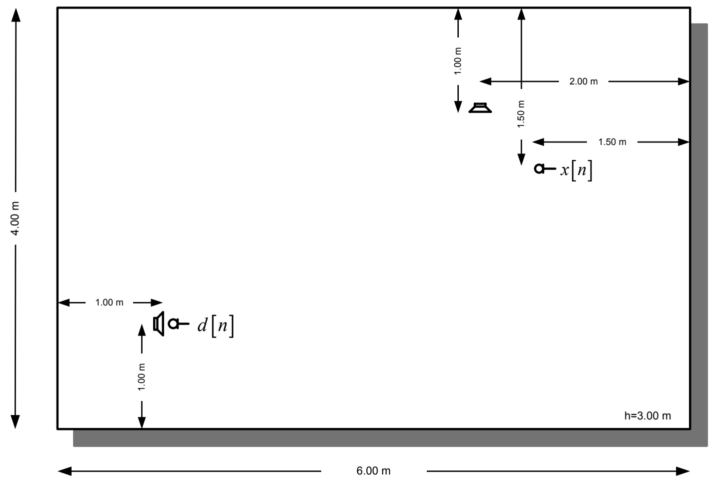

Fig. 2. The proposed experimental set-up

The filter lengths L depend on the reverberation time used, and it is listed in Table 1, while $L_{en} = 256$. The expansion order is set to $P = 2$. The signal test is a female speech, sampled at 8 kHz.

[1] Roomsim is a MATLAB simulation of shoe-box room acoustics for use in teaching and research. Roomsim is available from
`http://media.paisley.ac.uk/~campbell/Roomsim/`

Table 1. Reverberation time T_{60} and filter length L used in simulations

T_{60} [ms]	L
Anechoic	128
50	512
100	720
200	1024
350	2048

Table 2. Summary of output SNR [dB] for the proposed experimental tests with a female speech signal at different T_{60} and input SNR, using the NLMS algorithm and a mild distortion

SNR_{IN} [dB] \ T_{60} [ms]	Anechoic	50	100	200	350
10	25.33	20.13	18.93	16.49	14.16
5	23.78	20.27	17.42	15.45	12.33
0	20.88	17.72	16.36	13.99	10.91
-5	19.15	17.01	14.49	11.28	8.76
-10	18.13	15.32	11.41	7.81	5.89

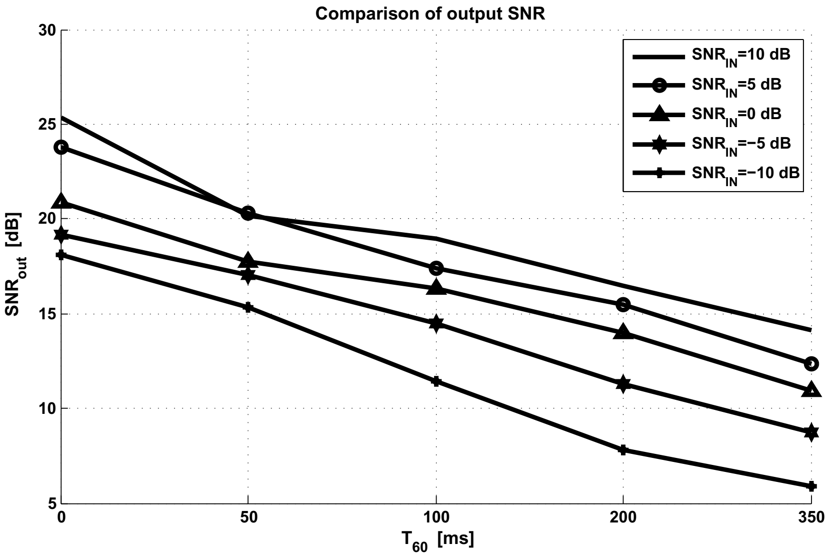

Fig. 3. Comparison of output SNR versus T_{60} and input SNR for the NLMS algorithm in the case of a mild distortion

The first experimental test is conducted using the NLMS algorithm (6). The other parameters used are $\mu_{\mathrm{L}}[0] = \mu_{\mathrm{N}}[0] = 0.2$, $\mu_a = 5 * 10^{-2}$, $\delta = 10^{-6}$, $\beta = 0.9$, $\gamma = 0.9$, while the parameter $a[n]$ is initialized as $a[0] = 1$ and the vectors of filter parameters $\mathbf{w}$

Table 3. Summary of output SNR [dB] for the proposed experimental tests with a female speech signal at different T_{60} and input SNR, using the NLMS algorithm and a strong distortion

SNR_{IN} [dB] \ T_{60} [ms]	Anechoic	50	100	200	350
10	24.30	20.09	18.63	15.51	12.88
5	22.73	19.88	17.22	14.82	11.43
0	19.79	17.27	15.35	13.19	10.87
-5	17.58	15.95	13.45	10.16	8.24
-10	16.49	13.69	10.15	7.66	4.18

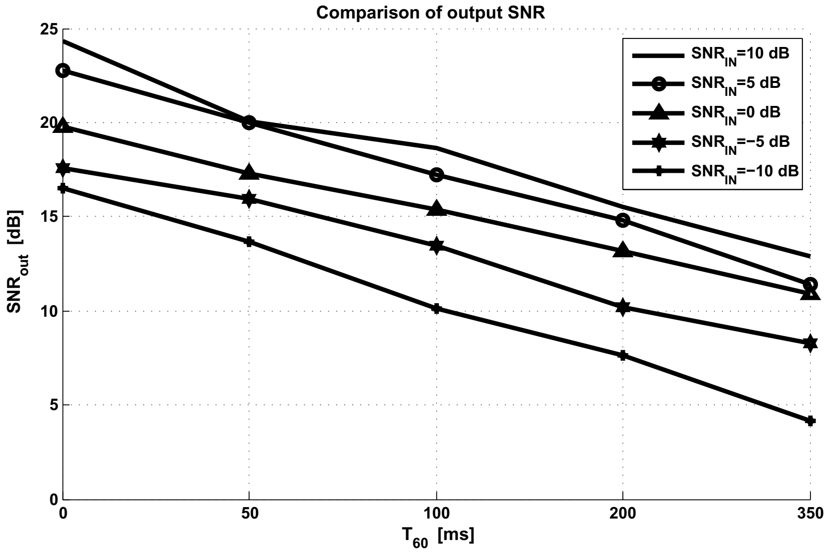

Fig. 4. Comparison of output SNR versus T_{60} and input SNR for the NLMS algorithm in the case of a strong distortion

and **h** are initialized to $\delta[n]$. A mild loudspeaker distortion is also considered, simulated as $d_{NL}[n] = \tanh(\beta d[n])$ and using $\beta = 0.5$. As background noise we use a white Gaussian noise (WGN). The power of the WGN is adjusted in order to have five different input SNR, precisely 10, 5, 0, -5 and -10 dB. Results for the considered test in all reverberant environments are summarized in Table 2. These results are averaged over 100 trials. Results presented in Table 2 are also depicted in Figure 3 for a more evident interpretation. As we can see from Figure 3 the performance in terms of output SNR is decreasing by increasing the reverberation time, due to the longer length of adaptive filters. This effect is more evident for high input SNR.

A second experimental test is conducted in the same simulated environment as the previous test, but using a stronger distortion, setting $\beta = 4$. The architecture parameters are set as in the previous case. Results for this simulations, averaged over 100 trials, are summarized in Table 3. These results are also depicted in the Figure 4. This figure

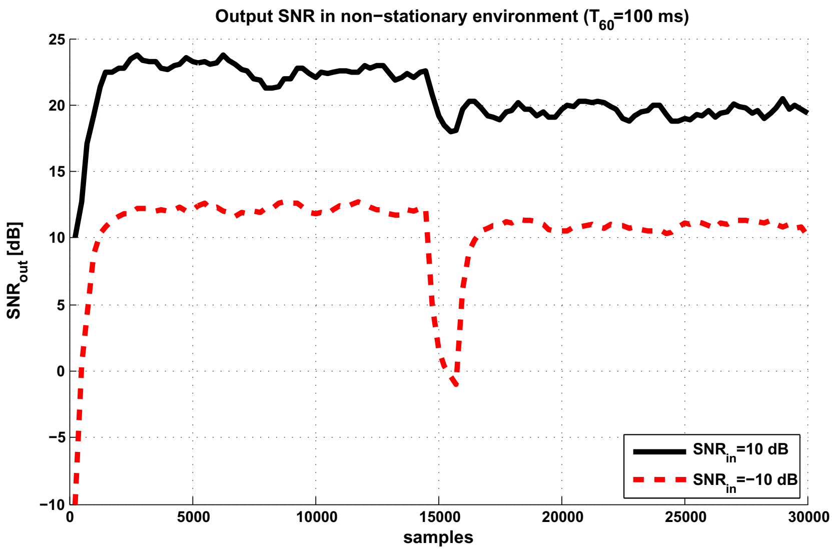

Fig. 5. Comparison of output SNR for two different input SNR at $T_{60} = 100$ ms in a changing environment

confirms that performances are decreasing by increasing the reverberation time and this trend is constant by decreasing the input SNR.

In a third experimental test, we use a changing environment. In the first half of the experiment the scenario is purely linear, then a strong distortion (with $\beta = 4$) is introduced. The reverberation time of this simulation is set to $T_{60} = 100$ ms and the architecture is run for an input SNR of 10 dB and -10 dB, respectively.

The output SNRs are evaluated averaging results over 250 samples. These results are depicted in the Figure 5. The profiles of graphics in Figure 5 show the same trend as the previous simulations, but evidence that the performance falls down when the scenario is abrupt changed, while it quickly returns to steady-state level. This fact confirms that we can take a great advantage from using a combination of a linear and nonlinear algorithms in the proposed collaborative architecture.

These simulations show that the proposed architecture is able to reach a good noise suppression in terms of SNR, even in the case of a strong nonlinear distortion and a changing environment.

3 Conclusions

In this paper an adaptive filter combination is proposed for the solution of Adaptive Noise Cancellation in nonlinear environment. The architecture described in this work is able to solve this hard problem in reverberant environment, as demonstrated by several experimental results. In particular the filter adaptation is conducted by the NLMS algorithm. It is shown that we can take advantage from the convex combination in the proposed collaborative architecture, reaching good performance in terms of both convergence and speed, and nonlinear distortion compensation.

References

1. Widrow, B., Glover, J.R., McCool, J.M., Kaunitz, J., Williams, C.S., Hearn, R.H., Zeidler, J.R., Dong, E.J., Goodlin, R.C.: Adaptive noise cancellation: Principles and applications. Proceedings of the IEEE 63(12), 1692–1716 (1975)
2. Ang, W.P., Farhang-Boroujeny, B.: A new class of gradient adaptive step-size LMS algorithms. IEEE Transactions on Signal Processing 49(4), 805–810 (2001)
3. Hwang, J.K., Li, Y.P.: Variable step-size LMS algorithm with a gradient-based weighted average. IEEE Transactions on Signal Processing 16(12), 1043–1046 (2009)
4. Arenas-García, J., Figueiras-Vidal, A.R.: Adaptive combination of normalised filters for robust system identification. Electronics Letters 41(15), 874–875 (2005)
5. Arenas-García, J., Figueiras-Vidal, A.R., Sayed, A.H.: Mean-square performance of a convex combination of two adaptive filters. IEEE Transactions on Signal Processing 54(3), 1078–1090 (2006)
6. Arenas-García, J., Gómez-Verdejo, V., Figueiras-Vidal, A.R.: New algorithms for improved adaptive convex combination of LMS transversal filters. IEEE Transactions on Instrumentation and Measurement 54(6), 2239–2249 (2005)
7. Comminiello, D., Scarpiniti, M., Parisi, R., Uncini, A.: A functional link based nonlinear echo canceller exploiting sparsity. In: Proc. of International Workshop on Acoustic Echo and Noise Control (IWAENC 2010), Tel Aviv, Israel, August 30-September 2 (2010)
8. Comminiello, D., Azpicueta-Ruiz, L.A., Scarpiniti, M., Uncini, A., Arenas-García, J.: Functional link based architectures for nonlinear acoustic echo cancellation. In: Proc. of Hands-free Speech Communication and Microphone Arrays (HSCMA 2011), Edinburgh, UK, May 30-June 1, pp. 180–184 (2011)
9. Sicuranza, G.L., Carini, A.: A generalized FLANN filter for nonlinear active noise control. IEEE Transactions on Audio, Speech, and Language Processing 19(8), 2412–2417 (2011)
10. Pao, Y.H.: Adaptive Pattern Recognition and Neural Networks. Addison-Wesley (1989)
11. Haykin, S.: Adaptive Filter Theory, 4th edn. Prentice-Hall (2001)
12. Paleologu, C., Ciochină, S., Benesty, J.: Variable step-size NMLS algorithm for under-modeling acoustic echo cancellation. IEEE Signal Processing Letters 15, 5–8 (2008)
13. Azpicueta-Ruiz, L.A., Zeller, M., Figueiras-Vidal, A.R., Arenas-García, J., Kellermann, W.: Adaptive combination of volterra kernels and its application to nonlinear acoustic echo cancellation. IEEE Transactions on Audio, Speech, and Language Processing 19(1), 97–110 (2011)
14. Campbell, D., Palomaki, K., Brown, G.: A Matlab simulation of shoebox room acoustics for use in research and teaching. Computing and Information Systems 9(3), 48 (2005)