ACOUSTIC ECHO CANCELLATION USING WAVELET TRANSFORM AND ADAPTIVE FILTERS

Bianca Alexandra FAGARAS, Cristian CONTAN, Marina Dana TOPA, Bases of Electronics Department, Technical University of Cluj-Napoca, Romania bianca.fagaras@gmail.com, {cristian.contan,marina.topa}@bel.utcluj.ro

<u>Abstract:</u> The paper proposes a new method for acoustic echo cancellation(AEC), in order to improve the performances of the normalized least-mean-square (NLMS) and non-parametric variable-step-size-NLMS (NPVSS-NLMS) algorithms. The acoustic system is modeled using an impulse response measured in a low reverberant enclosure. After using the two algorithms for identifying the mentioned acoustic system, one can observe that the NPVSS-NLMS provides better performances than the NLMS in terms of convergence rate at the same steady-state error. Simulations have been performed using both white Gaussian noise and a non-stationary audio signal as source sequences. For further convergence improvement, we propose the wavelet transform for the input signal decomposition. The two adaptive algorithms are applied to the wavelet structure obtained. The proposedwavelet based identification methods provide better convergence rates than each of the two conventional adaptive filters. Also one can notice that the wavelet NPVSS-NLMS(WNPVSS-NLMS) provides better performances than the wavelet NLMS (WNLMS). MATLAB simulations performed for white Gaussian noise and non-stationary audio signal used as input prove the previous statements.

Keywords: wavelet transform, adaptive filters, acoustic echo cancellation, convergence rate.

I. INTRODUCTION

The acoustic echo cancellation (AEC) [1, 2] setup is illustrated in Figure 1 with the aim of minimizing the residual error signal e(n) defined as:

$$e(n) = u(n) - y(n),$$
 (1)

where u(n) is called the desired signal, containing the output of the unknown system d(n) and the local signal from the acoustic enclosure (unknown system) z(n)(u(n) = d(n) + z(n)), while y(n) is the output of the adaptive filter. By minimizing the estimated error at each iteration, the AEC setup has the role to identify the coefficients of the unknown system in the form of the weights of the adaptive filter. Eventually, in the ideal AEC scenario, the error signal will resemble the local signal from the unknown acoustic system $(\lim_{n\to\infty} e(n) = z(n)) [3].$

The signals involved in equation (1) are defined as:

$$\begin{cases} d(n) = \mathbf{h}^T \cdot \mathbf{x}(n) \\ y(n) = \widehat{\mathbf{h}}^T(n) \cdot \mathbf{x}(n), \end{cases}$$
(2)

where the following vector definitions are involved:

$$\boldsymbol{h} = [h_0, h_1, \dots, h_{M-1}]^T, \tag{3}$$

$$\widehat{h}(n) = \left[\widehat{h}_{0}(n), \widehat{h}_{1}(n), \dots, \widehat{h}_{M-1}(n)\right]^{T}, \quad (4)$$



Figure 1. Basic schematic diagram of an adaptive filter

$$\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n-M+1)]^{T}.$$
 (5)

The vector **h** contains the coefficients of the unknown system (room impulse response), while $\hat{h}(n)$ includes the tap weights of the adaptive filter, which should resemble, in the steady-state phase of the adaptive filter, the coefficients of **h**. For simplicity, both vectors have the same length, *M*. The vector $\mathbf{x}(n)$ is an*M*length mobile window which contains the last*M* samples of the input signal starting from the *k*-th element.

Usually, in the literature [4] the adaptive filters use adaptive least-mean-square (LMS) methods to update the tap weights of the filter.

The LMS algorithm is known due to its simplicity in design and implementation. The method, on which this

algorithm is based, is the steepest descent recursion [5].

The major disadvantage of the LMS algorithm is its slow convergence rate. In order to improve its performances, a series of new algorithms have been developed, derived from the LMS algorithm.

A. The normalized LMS algorithm (NLMS)

A special implementation of the LMS algorithm is the NLMS [6]. This algorithm takes into account the signal variation at the filter's input and selects a normalized step-size parameter that is a more stable and faster converging adaptation algorithm than the LMS [7].

Table 1 illustrates the steps of the adaptation process of the NLMS algorithm in order to minimize the residual error.

Process	Equation			
Initialization	$\widehat{\boldsymbol{h}}(0) = \boldsymbol{0}$			
Setting up	$0 < \alpha < 2, \alpha = ct.$			
Parameters	$\delta = 0.1\sigma_x^2$, for $x(n) \in [-1,1]$			
Error	$e(n) = u(n) - \widehat{\mathbf{h}}^T(n)\mathbf{x}(n)$			
Update	$\mu(n) = \frac{\alpha}{\alpha}$			
	$\mu(n) = x^T(n)x(n) + \delta$			
	$\widehat{\boldsymbol{h}}(n+1) = \widehat{\boldsymbol{h}}(n) + \mu(n)\boldsymbol{x}(n)\boldsymbol{e}(n)$			

Table 1. The steps of the NLMS algorithm

The NLMS algorithm converges faster than the LMS, at little extra cost in computing the norm of the input signal.

The step-size parameter is the one that governs the stability of the NLMS algorithm. The tradeoff between fast convergence and low misadjustment is reflected through the choice of this parameter. This tradeoff appears due to the fact that the step-size parameter is chosen constant in the range (0, 2), regardless of the filter stages (convergence or steady-state). In order to minimize this tradeoff even more, a set of variable step-size algorithms were used as in [8, 9]. In this works the step-size is chosen adaptively, depending on the values of the residual error.

B. The non-parametric variable step-size NLMS (NPVSS-NLMS)

The NPVSS-NLMS has been developed in the context of increasing the performances of the NLMS algorithm and in order to allow the choice of the step-size parameter adaptively and not as a constant value set by the user like in the NLMS algorithm.Table 2 summarizes the NPVSS-NLMS algorithm as in [10].One of the most important aspects of the NPVSS-NLMS is the step-size parameter. It is determined based on the variation of the input sequence and also on the variance of the local noise. The step-size is known as a positive factor, which controls the system stability, convergence speed and system inadequacy [11]. In the AECsetup, the two adaptive filter's performances are evaluated in terms of the attenuated amount of echo. The most common performance measures used are the mean-square-error (MSE) and the echo-return loss enhancement (ERLE). The ERLE is defined as the ratio of the power of the desired signal and the power of the residual error signal. It is a smoothed measure (in dB) of the amount of the echo that has been attenuated.

	Process	Equation				
	Initialization	$\widehat{\boldsymbol{h}}(0) = \boldsymbol{0}$				
		$\widehat{\sigma_e^2}(0) = 0$				
	Setting up	$\lambda = 1 - \frac{1}{KL}$, exponential window with				
	1 urumeters	$K \geq 2$				
		$L = \frac{\boldsymbol{x}^{T}(n)\boldsymbol{x}(n)}{2}$				
		σ_x^2				
		σ_z^2 , local signal variance				
		$\delta = 0.1\sigma_x^2, x(n) \in [-1,1]$				
		$\epsilon > 0$, very small number to avoid				
	Error	$\widehat{\mathbf{UVIding}}$ by zero				
	EII0I	$e(n) = u(n) - \mathbf{h}^{T}(n)\mathbf{x}(n)$				
	Update	$\widehat{\sigma_e^2}(n) = \lambda \widehat{\sigma_e^2}(n-1) + (1-\lambda)$				
J		$\cdot e^2(n)$				
		$\alpha(n) = \frac{\left[1 - \frac{\sigma_z}{\epsilon + \widehat{\sigma_e}(n)}\right]}{\left[\delta + \boldsymbol{x}^T(n)\boldsymbol{x}(n)\right]}$				
		$\mu_{\text{NPVSS}}(n) = \begin{cases} \alpha(n), & \text{if } \widehat{\sigma_e}(n) > \sigma_z \\ 0, & \text{otherwise} \end{cases}$				
		$\widehat{\boldsymbol{h}}(n+1)$				
		$= \widehat{h}(n) + \mu_{\text{NPVSS}}(n) x(n) e(n)$				

Table 2. The steps of the NPVSS-NLMS algorithm

ERLE
$$(n) = 10 \log_{10} \frac{E\{u^2(n)\}}{E\{e^2(n)\}} [dB],$$
 (6)

where $E\{\cdot\}$ denotes statistical expectation.

The MSE is defined as:

$$MSE(n) = E\{[u(n) - y(n)]^2\} = E\{[e(n)]^2\}.$$
 (7)

C. Performance comparison in AEC

The analysis was performed using two input signals with different probability density functions (pdfs): white Gaussian noise (WGN) and a non-stationary audio signal. Also, additive white Gaussian noise (AWGN) was usedas local signal. The unknown system used here is the impulse response of the acoustical enclosure, having 320 coefficients in length. In order to implement these two algorithms, we have to follow the steps of their adaptation processes mentioned in Tables 1 and 2.

Electronics and Telecommunications

The signal-to-noise ratio (SNR) value is set to 25dB in all conducted simulations and this is why in all simulations the ERLE, computed using the average on a mobile window of 800 samples, should stabilize around this value [12]. For the exponential window, K is chosen 700.

As observed in Figure 2, the NPVSS-NLMS algorithm converges faster than the NLMS algorithm and stabilizes approximately to the required SNR value of 25dB. For the MSE, there are similar results, meaning that the NPVSS-NLMS MSE is smaller than the MSE of the NLMS algorithm, which leads to the fact that the NPVSS-NLMS converges faster to an optimum value. This aspect is shown in Figure 3.Similar observations regarding the two involved adaptive filters can be made for a non-stationary audio input signal as shown in Figures 4 and 5.



Figure 2. ERLE evolution for the NLMS and NPVSS-NLMS algorithms for a WGN as input signal (M=320, $\alpha=0.04$)



Figure 3. MSE evolution for the NLMS and NPVSS-NLMS algorithms for WGN as input signal (M=320, α =0.04)



Figure 4. ERLE evolution for the NLMS and NPVSS-NLMS algorithms for a non-stationary audio input signal $(M=320, \alpha=0.04)$



Figure 5. MSE evolution for the NLMS and NPVSS-NLMS algorithms for a non-stationary audio input signal $(M=320, \alpha=0.04)$

II. PROPOSED AEC METHOD USING WAVELET TRANSFORM AND ADAPTIVE FILTERS

A. Wavelet Transform. Theoretical Aspects

Wavelet may be seen as a complement to classical Fourier decomposition method. Unlike short time Fourier transform (STFT), wavelet analysis uses a windowing technique with variable-sized regions and in this way we have short windows at high frequencies and long windows at low frequencies in order to obtain precise information as one can observe in Figure 6 [13].

The wavelet analysis does not use time-frequency domain, but time-scale domain.

One major advantage offered by wavelets is the possibility of performing local analysis in order to examine a localized area from a larger signal.

Fourier analysis consists of separating a signal into sinusoids of different frequencies. In a similar way, wavelet analysis is the separation of a signal into scaled and shifted versions of the original wavelet, called mother wavelet denoted by Ψ .

In practice, we use only a subset of scales and positions, based on powers of two, called*dyadic* scales and positions. Therefore, our analysis will be much more efficient and just as accurate. This analysis is obtained from the Discrete Wavelet Transform (DWT) [14].

It is known that the low-frequency content is the most important part because it gives the signal its identity, while the high-frequency gives the nuance of the signal. This is why we have approximation and detail coefficients in wavelet analysis. Approximations are large-scale, low-frequency components of the signal and the details are the low-scale, high-frequency components.



Figure 6. Wavelet Transform



Figure 7. Signal analysis and signal synthesis using wavelet coefficients

Figure 7 depicts the decomposition and reconstruction processes of the original signal using wavelet coefficients.

The decomposition process consists of passing the discrete signal *x* through two complementary filters, and thus, resulting two output signals. If this operation is performed on a real digital signal, the resulting quantity of information will be doubled. To adjust this problem, we introduce the notion of *downsampling*, meaning that every second data point will be thrown away (without losing information). This introduces aliasing in the signal components. The decomposition of the signal in coefficients is called wavelet analysis and it involves filtering and downsampling.

The resulting components have to be assembled back into the original signal without losing information, process called reconstruction or synthesis, which involves*upsampling* and filtering. Upsampling is the process of increasing the signal by inserting zeros between samples.

By applying this method in the AEC scenario to the adaptive structure, we reduce the computational effort by using parallel processing of the input signal that has a smaller length than the original signal. Also, this procedure can increase the convergence rate of the conventional algorithms.

B. AEC using Wavelet Transform and Adaptive Filters. Theoretical Aspects

It is common knowledge that, when the number of input samples involved in the adaptation process is very large, the convergence of adaptive filtering algorithms becomes slow. One of the solutions adopted for this problem is the use of adaptive filters in sub-bands [15]. The first step of this solution is the decomposition of the input signal, using an analysis filter bank. After this, the signals from different sub-bands are processed using adaptive filters. Thus, the newly formed input signal presents a lower correlation degree than the original signal. In order to add flexibility to the adaptive system we use the wavelet packet transform instead of the fixed schemes [16].

This paper presents a comparison between the performances of the classical NLMS and NPVSS-NLMS algorithms and the performances of the wavelet NLMS and wavelet NPVSS-NLMS. The last two methods are proposed in order to improve the convergence rate of the conventional two algorithms.



Figure 8.Adaptive structure composed of a wavelet transform and adaptive filters

The adaptive wavelet structure used in this paper is presented in Figure 8, where: x(n) is the input signal, $H_i(z)$, $i = \overline{0, Q-1}$, represents the analysis bank (wavelet decomposition), Q is the number of sub-bands, $G_i(z)$ represents the adaptive filter structures, y(n) is the output signal, u(n) is the desired signal, and e(n) is the error signal. The development of the proposed algorithm can be carried out for any type of wavelet families which will be mentioned.

C. Simulations results

C.1. Wavelet NLMS (WNLMS)

After proving in section 2 that the NPVSS-NLMS algorithm has better performances regarding convergence rate than the NLMS algorithm, the next step is the implementation of the Wavelet NLMS and Wavelet NPVSS-NLMS methods to improve even more the convergence rate of the conventional adaptive algorithms. To begin with, we apply wavelet decomposition to the input signal, through which a single-level onedimensional decomposition is performed with respect to a particular wavelet(Daubechies, Haar, Coiflet. Biorthogonal, Symlet) [17]. In our case, we use Biorthogonal 4.4. wavelet, that decomposes the input signal in2 new input signals corresponding to the two subbands, containing the approximation and detail coefficients. Then, we follow the main steps of the adaptation processes of the NLMS and NPVSS-NLMS algorithms, for the new input signals.

Figure 9 illustrates the ERLE performances for the classical NLMS algorithm and the WNLMS algorithm. It can be observed that the WNLMS ERLE converges much faster than the ERLE computed for NLMS, while both ERLE characteristics stabilize to the imposed SNR value of 25 dB. In the same way, the mean-square-error of the WNLMS is smaller than the mean-square-error of the NLMS, leading to faster convergence towards the optimum value. The MSE performance is presented in Figure 10.The value of the step-size parameter used in the adaptation process is the same for both algorithms



Figure 9. ERLE evolution for the NLMS and WNLMS algorithms for a WGN as input (M=320, SNR=25dB)



Figure 10. MSE evolution for the NLMS and WNLMS algorithms for a WGN as input (M=320, SNR=25dB)



Figure 11. ERLE evolution for the NLMS and WNLMS algorithms for a non-stationary audio input signal (M=320, SNR=25dB)

 $(\alpha_{\text{NLMS}} = \alpha_{\text{WNLMS}} = 0.04)$. The same results are obtained for a non-stationary audio input signal, shown in Figures 11 and 12, by using the same step sizes: $\alpha_{\text{NLMS}} = \alpha_{\text{WNLMS}} = 0.04$.

C.2. Wavelet NPVSS-NLMS (WNPVSS-NLMS)

For the implementation of the wavelet NPVSS-NLMS algorithm we use the same wavelet decomposition (Biorthogonal 4.4) of the input signal. Therefore, we obtain the two input signals, used in the adaptation equations of the WNPVSS-NLMS algorithm.

Figure 12 depicts the ERLE characteristics of the NPVSS-NLMS and WNPVSS-NLMS structures. The wavelet based algorithm stabilizes faster than the NPVSS-NLMSaround the required SNR value of 25 dB. These

results are obtained for both WGN and non-stationary audio signal as input. The same hierarchy is held in the case of MSE as shown in Figure 13. The ERLE characteristics illustrated in Figure 14 are obtained for the same adaptive structures as used in Figure 12 but for a non-stationary audio signal applied as input. One can observe that in this case also, the WNPVSS-NLMS surpasses the NPVSS-NLMS in terms of convergence rate for almost the same steady-state error.

Figure 15depictsthe ERLE characteristics of all implemented adaptive structures (NLMS, WNLMS, NPVSS-NLMS, WNPVSS-NLMS). Choosing the NLMS algorithm as reference, one can observe that each of the remaining three adaptive structures provides better performances than the previous one. We observe that the algorithm with the best performances is the WNPVSS-NLMS, followed by the NPVSS-NLMS algorithm. Also, we managed to significantly improve the NLMS algorithm by using the wavelet transform.

Figure 16 represents the MSE characteristics of the four algorithms for WGN as input, while Figure 17 depicts all four ERLE curves for a non-stationary audio signal as input.

For more accuracy regarding the convergence time of each adaptive algorithm, measurements were performed in Figure 15. The results are summarized in Table 3.

Simulations were performed also for other types of wavelet families: Daubechies1 (db1), Daubechies4 (db4), Biorthogonal 2.4. (Bior2.4), Coiflet4 (coif4), and Symlet4 (sym4), in order to prove that we obtain similar results for each of them. Simulations were conducted for WGN as input and also as local signal. In the WNLMS case the value of the step-size parameter is set to 0.04 for all types of wavelets. Also the SNR value was maintained at 25 dB. In the WNPVSS-NLMS case, for the exponential window K was kept unchanged. The measured convergence time for each wavelet type is provided in Table 4.



Figure 12. ERLE evolution for the NPVSS-NLMS and WNPVSS-NLMS algorithms for WGN as input signal (M=320, SNR=25dB)



Figure 13. MSE evolution for the NPVSS-NLMS and WNPVSS-NLMS algorithms for WGN as input signal (M=320, SNR=25dB)

Implemented adaptive structure	Convergence Time (s)		
NLMS	4.0225		
WNLMS	1.87125		
NPVSS-NLMS	0.7086		
WNPVSS-NLMS	0.486		

 Table 3.Convergence time of the four structures from
 figure 15



Figure 14. ERLE evolution for the NPVSS-NLMS and WNPVSS-NLMS algorithms for non-stationary audio input signal (M=320, SNR=25dB)



Figure 15. ERLE evolution for NLMS, WNLMS, NPVSS-NLMS and WNPVSS-NLMSfor WGN as input signal



Figure 16. MSE evolution for NLMS, WNLMS, NPVSS-NLMS and WNPVSS-NLMSfor WGN as input signal



Figure 17. ERLE evolution for NLMS, WNLMS, NPVSS-NLMS and WNPVSS-NLMSfor a non-stationary audio signal as input

Convergence Time (s)									
	db1	db4	Bior2.4	coif4	sym4				
WNLMS	2,22	1,97	1,93	2,11	1,97				
WNPVSS- NLMS	0,45	0,43	0,48	0,44	0,56				

Table 4. Convergence time for each wavelet type

III. CONCLUSIONS

In this paper a novel adaptive acoustic echo cancellation (AEC) technique was proposed in order to minimize the tradeoff between convergence rate and steady-state misadjustment. This tradeoff is encountered usually in conventional least mean-square (LMS) based adaptive filters, due to the fact that the step-size parameter is chosen constant, regardless of the filter state (convergence or steady-state). In order to better understand this problem, in the first part of this paper, two algorithms were presented and implemented, the normalized leastmean-square (NLMS) and non-parametric variable stepnormalized least-mean-square (NPVSS-NLMS) size algorithms. From the obtained simulations, one can observe that the NPVSS-NLMS algorithm provides better convergence rates than the NLMS algorithm for the same signal to noise ratio (SNR) value. Therefore, the tradeoff between convergence rate and steady-state misadjustment minimized by choosing a variable step-size is parameter.Simulations were performed for two types of input signal: white Gaussian noise (WGN) and a nonstationary audio signal.

In order to obtain better performances and to minimize

even more this tradeoff, a new structure that employs wavelet transform to decompose the input signalthat employs adaptive filters was proposed. In this case also, for simulations, we used two types of input signals: WGN and a non-stationary audio signal. The simulations illustrate that the two conventional algorithms have been improved, as their convergence rates are better. Simulations were also performed for different types of wavelets, proving in this way that the same hierarchy regarding convergence rate is obtained for each of them.

Further work will be concentrated on performing the same simulations on multiple level wavelet decomposition.

REFERENCES

[1] E. Hänsler, G. Schmidt, *Acoustic Echo and Noise Control*, Wiley-Interscience, New Jersey, 2004.

[2] T. Ogunfunmi, Adaptive Nonlinear System Identification: The Volterra and Wiener Model Approaches, Springer, 2007.

[3] C. Contan, M. Topa, "Decoupled Second-Order Volterra Filter Structure Using the NLMF Adaptive Algorithm for Nonlinear Acoustic Echo Cancellation,"*Acta Technica Napocensis-Electronics and Telecommunications*, Vol. 53, No. 2, pp. 56-60, 2012.

[4] B. Farhang-Boroujeny, *Adaptive Filters Theory and Applications*, Wiley, 1998.

[5] C. Contan "Acoustic echo cancellation in nonlinear systems", *PhD Thesis*, Technical University of Cluj-Napoca, 2012.

[6] S. Haykin, *Adaptive Filter Theory*, Englewood Cliffs, NJ, USA, Prentice Hall, 1996.

[7] I. Dornean, *Contribution to the Design and Implementation of Adaptive Algorithms Using Multirate Signal Processing on FPGA*, PhD Thesis, Technical University of Cluj-Napoca, 2011.

[8] C. Rusu, C. F. N. Cowan, "The exponentiated convex variable step-size (ECVSS) algorithm", *Signal Processing*, Elsevier, vol.90, pp.2784-2791, 2010.

[9] S. Makino, Y. Kaneda, N. Koizumi, "Exponentially weighted Stepsize NLMS Adaptive Filter Based on the Statistics of a Room Impulse Response," *IEEE Transactions on Speech and Audio Processing*, vol. 1, no. 1, pp. 101–108, Jan. 1993.

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

[11] I.Homana, Acoustic Echo Cancellation using Adaptive Filters, PhD Thesis, Technical University of Cluj-Napoca, 2011. [12] F. Kuech, A. Mitnacht, W. Kellermann, "Nonlinear acoustic echo cancellation using adaptive orthogonalized power filters", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 3, pp.105-108, 2005.

[13] O. Rioul, M.Vettereli, "Wavelets and Signal Processing", *IEEE SP Maganzine*, October 1991.

[14] A.D.Poularikas, *Wavelet Transforms - The Handbook of formulas and tables for signal processing*, Ed. Alexander D. Poularikas, 1999.

[15] M. R. Petraglia, G.Barboza, "Improved PNLMS algorithm employing wavelet transform and sparse filters", *EUSIPCO*, 2008.

[16] N. Gonzalez Prelcic, F. Perez Gonzalez, M.E. Dominguez Jimenez, "Wavelet packet-based subband adaptive equalization", *Signal Processing*, Volume 81, Number 8, pp. 1641-1662(22), August 2001.

[17] Michel Misiti, Yves Misiti, G. Oppenheim, Jean-Michel Poggi, *Wavelet Toolbox- For use with MATLAB*, The MathWorks, Inc., Natick, Massachusetts 01760, April 2001.