PROPORTIONATE ALGORITHMS FOR TWO-MICROPHONE ACTIVE FEEDBACK CANCELLATION

Felix Albu[1], *Renato Nakagawa*[2], *Sven Nordholm*[3]

[1]Faculty of Electronics, Valahia University of Targoviste, Targoviste, Romania
Email: felix.albu@valahia.ro
[2]Signal Processing Research Department, Starkey Hearing Technologies, USA
Email: Renato_Nakagawa@starkey.com
[3]Department of Electrical and Computer Engineering, Curtin University, Bentley, Australia
Email: Sven.nordholm@curtin.edu.au

ABSTRACT

In this paper we propose the use of the proportionate principle in order to improve the convergence characteristics of the two microphone method for acoustic feedback cancellation in hearing aids. The reason of using proportionate algorithms is to exploit the sparseness of the adaptive filter coefficients in the transform domain. The convergence improvement can be achieved for both speech and music signals at a moderate increase of the numerical complexity over that of a previous solution in the transform domain.

Index Terms— acoustic feedback cancellation; proportionate adaptive algorithms, two microphones approach, adaptive algorithms.

1. INTRODUCTION

Hearing aids suffer from acoustic feedback problems caused by the acoustic coupling between loudspeaker and microphone. The microphone(s) picks up the loudspeaker's signal and the created acoustic loop can cause potentially system instability. Therefore, the feedback problem limits the maximum stable gain (MSG) achievable and deteriorates the sound quality [1]. The preferred option to avoid this problem is to use acoustic feedback cancelers (AFC) [2]-[3]. The purpose of AFC is essentially to identify a model of the feedback path and to estimate the feedback signal. The feedback estimate is then subtracted from the microphone signal. Unfortunately, due to the closed loop signals, the biased canceler's coefficients lead to poor system performance [1], [3]. Different techniques have been proposed to reduce this correlation including phase modification, frequency shifting, decorrelating prefilters, adaptive filters in tandem, use of synthesized signals, inverse gain filters, and probe noise injection [2]–[10]. The signal correlation can be reduced by using orthogonal transforms; these have been shown to increase the convergence rates in stochastic gradient algorithms such as the least mean squares (LMS) algorithm [9], Normalized LMS (NLMS) [11]. The discrete cosine transform (DCT) is applied to the prediction error method (PEM) in [11] to boost the PEM performance for acoustic echo cancellation (AEC) and AFC. In [12], an additional microphone was employed to obtain an incoming signal estimate that is removed from the error signal prior to adapting the canceler. This method was called the two microphone (TM) acoustic feedback canceler. In [13] the orthogonal transforms were used with the TM method. The discrete Fourier transform (DFT) and the discrete Cosine transform (DCT) were implemented to transform the adaptive filter signals. The transform was not applied only to the input signal of the canceler as in [11], but also to the error signal. Also a bank of adaptive filters was employed, each adapting to different portions of the spectrum. The full band filter coefficients were synthesized and used to provide the necessary signal estimates [13]. In [13] the NLMS algorithm was used for adapting the filter weights and only results for speech were reported. We've noticed the sparseness of the feedback path in [14] and investigated the use of simplified proportionate algorithms for TM. Thus the use of improved proportionate NLMS (IPNLMS) which is a variant of the proportionate NLMS (PNLMS) [15] and proportionate sign NLMS (PSNLMS) algorithms is proposed for different variations of the transformed domain two-microphone schemes for both speech and music signals. The results with music signals of two microphone methods have not been reported before.

This paper is structured as follows. First, the proposed proportionate transform domain with filtered error version of the TM method is presented followed by simulation results. At the end, the conclusions of the paper and future work are presented.

2. THE PROPOSED METHOD

In the two microphone method [13], the location of the microphones is shown in Fig. 1. One microphone is in the ear and the additional microphone is located behind the ear. A challenge with the TM approach in comparison with the classical one microphone approach is the pres-

The work of F. Albu was supported by a grant of the Romanian National Authority for Scientific Research, CNCS-UEFISCDI, project number PN-II-ID-PCE-2011-3-0097.

ence of a second feedback channel $G_2(q)$. It was shown in [12] that although $G_2(q)$ introduces a bias to the solution it can be assumed that $|G_2(q)|$ is weak in many cases of directions of the main incoming wave. The feedback path between the loudspeaker and the microphone is modeled by a finite impulse response (FIR) filter with the coefficient vector $\mathbf{g}_1 = \left[g_{1_0} \ldots g_{1_{L_g-1}} \right]^T$ with filter length L_g, which is represented by a polynomial transfer function $G_1(q)$ in q as $G_1(q) = \mathbf{g}_1^T \mathbf{q}$ with $\mathbf{q} = \left[1 \; q^{-1} \; \ldots \; q^{-L_g+1} \right]^T$ [13].

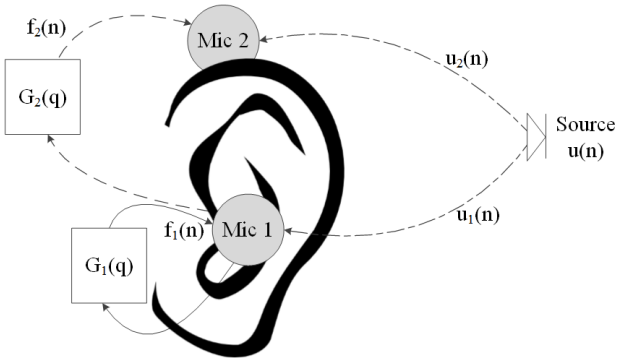

Fig. 1. Microphone arrangement [13]

The TM configuration is presented in Fig. 2. We have the incoming signals $u_1(n)$ and $u_2(n)$ and $H(q)$ is a FIR filter with length L_h [13]. The delay d_m in the first microphone signal path is needed to predict inputs signals no matter the direction of the incoming wave.

$\hat{\mathbf{H}}(q)$ is an adaptive FIR filter of length $L_{\hat{h}} > L_h + d_m$ which filters the second microphone signal $m_2(n)$ producing the incoming signal estimate $\hat{u}_1(n)$ which is subtracted from $\tilde{u}_1(n)$ to yield the error signal $e_1(n)$ [12]. We denote the scheme from Fig. 2 as TM Type I to refer to both adaptive filters $\hat{\mathbf{H}}(q)$ and $\hat{\mathbf{G}}_1(q)$ that filters $y(n)$ being in the transform domain.

In Fig. 3 we show the block diagram of the transformed domain implementations of the TM method used in this paper. The transform can be DFT or DCT [16], [17]. The scheme from Fig. 3 has not been previously reported (w(n) represent the white noise used as a probe signal). We use the DFT transform in this paper, but the results of using DCT are similar. The scheme from [13] differs from Fig. 3 by having a supplementary bank of filters for $\hat{\mathbf{H}}(q)$. Therefore we denote this scheme as the TM Type II (see Fig. 3) where only $\hat{\mathbf{G}}_1(q)$ is updated in the transform domain.

In [13] the NLMS algorithm was used for updating the filter coefficients. The complex NLMS algorithm [18] updates its coefficients as follows:

$$\hat{\mathbf{g}}_i(n) = \hat{\mathbf{g}}_i(n-1) + \frac{\mu \mathbf{y}_1^*(n) e_{M,i}(n)}{\delta + \left\| \mathbf{y}_1(n) \right\|^2} \tag{1}$$

where μ is the step-size parameter, $i = 1 \ldots M$, M is the number of subbands, $\hat{\mathbf{g}}_i(n)$ are the coefficients of ith subband filter, $e_{M,i}(n)$ is the ith element of filtered error $\mathbf{e}_M(n)$ vector, and δ is a regularization constant. The subbands are equally spaced. The filtered error vector $\mathbf{e}_M(n)$ is the FFT of the vector collecting previous $\mathbf{e}_1(n)$ values $\left[e_1(n), e_1(n-1), .., e_1(n-M) \right]^T$ divided by $\sqrt{M}$.

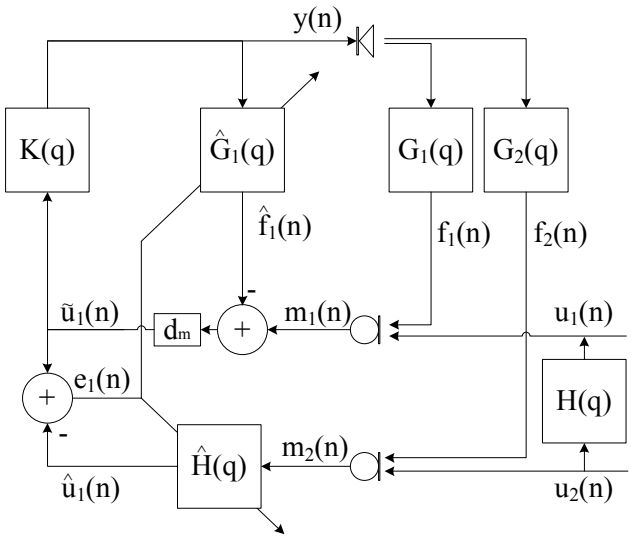

Fig. 2. TM Type I block diagram [13].

We have noticed the sparseness of the adaptive filter coefficients in the M subbands and one known way to exploit this sparseness is to use the proportionality principle for updating the filter coefficients.

A proportionate-type NLMS algorithm [19] updates its coefficients according to:

$$\hat{\mathbf{g}}_i(n) = \hat{\mathbf{g}}_i(n-1) + \frac{\mu \mathbf{C}(n-1) \mathbf{y}_1^*(n) e_{M,i}(n)}{\delta + \mathbf{y}_1(n) \mathbf{C}(n-1) \mathbf{y}_1^*(n)} \tag{2}$$

where $\mathbf{C}(n-1)$ is a diagonal matrix which assigns an individual step-size to each filter coefficient

$$\mathbf{C}(n-1) = diag\left\{ c_0(n-1), ..., c_{L_g-1}(n-1) \right\}. \tag{3}$$

In case of the IPNLMS algorithm [19], the diagonal elements of $\mathbf{C}(n-1)$, $c_l(n-1)$, with $0 \le l \le L_g - 1$, are evaluated as:

$$c_l(n-1)=\frac{1-\beta}{2L_g}+\frac{|g_l(n-1)|(1+\beta)}{2\sum_{i=0}^{L_g-1}|g_i(n-1)|+\xi}, \quad (4)$$

where $-1\le\beta<1$, $g_i(n-1)$ is the ith coefficient of real part of $\sum_{i=0}^{M-1}|\hat{g}_i(n-1)|$ and the small positive constant ξ avoids division by zero.

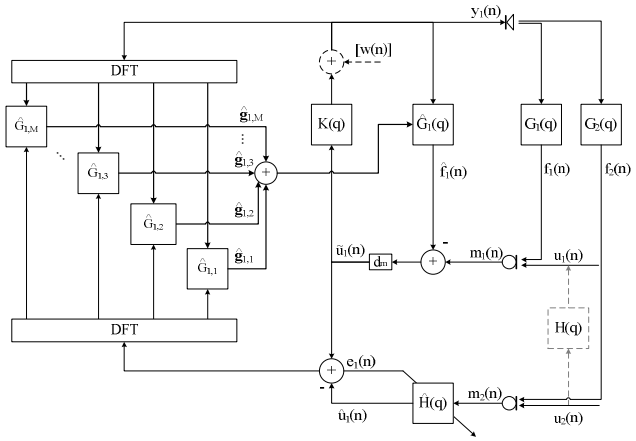

Fig. 3. TM Type II block diagram without subband decomposition of $\hat{\mathbf{H}}(q)$.

One variant is to compute a proportionate matrix $\mathbf{C}_i(n-1)$ for each subband using $\hat{g}_i(n-1)$ instead of $g_i(n-1)$. An average of less than 0.5 dB improvement has been obtained in our simulations and for each subband an additional $2ML_g$ multiplications are needed (in our simulations M = 8 was used). This important additional numerical complexity does not justify the relatively small performance improvement that is obtained.

In [20] and [21] several proportionate affine projection sign algorithms were proposed. It was shown that they are robust in impulsive environments. In the following lines the proportionate sign NLMS (PSNLMS) algorithm is derived as a particularization for a projection order of one for the memory improved proportionate affine projection sign algorithm (MIP-APSA) [21]. If we note by

$$\mathbf{x}_{gs}(n)=\mathbf{C}(n-1)\mathbf{y}_1^*(n)\,\mathrm{sgn}\left(e_{M,i}(n)\right), \quad (5)$$

where $\mathrm{sgn}(\cdot)$ is the sign function, the complex PSNLMS algorithm updates its coefficients as follows:

$$\hat{\mathbf{g}}_i(n)=\hat{\mathbf{g}}_i(n-1)+\frac{\mu\mathbf{x}_{gs}(n)}{\sqrt{\delta+\left\|\mathbf{x}_{gs}(n)\right\|^2}} \quad (6)$$

Therefore the proposed PNLMS based algorithm is called TM-P-DFT-II, the proportionate sign based algorithm is called TM-PS-DFT-II. The original algorithm from [13] is called TM-DFT-I in this paper.

The added complexity of TM-P-DFT-II over TM-DFT-I is $2L_{\hat{h}}+(M+1)\cdot L_g$ multiplications. The complexity of the sign based TM-PS-DFT-II algorithm is slightly higher than that of the TM-DFT-I algorithm, i.e. it has $L_{\hat{h}}+L_g$ more multiplications.

3. SIMULATION RESULTS

Simulation signals were obtained from a recording studio using a Brüel & Kjær (B&K) head and torso simulator type 4128C. In our simulations the path change is given by placing a flat surfaced object very close to the ear. It was shown in [13] that the measured second feedback path's magnitude response is much weaker than the first feedback path. The input sequence used for the speech signals was real speech segments from NOIZEUS database which contains 30 IEEE sentences spoken by 3 male and 3 female speakers [22]. In order to assess the performance of the algorithm, the misalignment between the true and estimated feedback path and the maximum stable gain (MSG) measures are used. The misalignment is used to represent the accuracy of the feedback path estimation and is defined as

$$\text{Misaligment} = 20\log_{10}\frac{\int_0^{\pi}\left\|G(\omega)-\hat{G}(\omega)\right\|_2 d\omega}{\int_0^{\pi}\left\|G(\omega)\right\|_2 d\omega}. \quad (7)$$

MSG is defined as

$$\text{MSG} = 20\log_{10}\left[\min_{\omega}\frac{1}{\left|G(\omega)-\hat{G}(\omega)\right|}\right] \quad (8)$$

In the simulations the following parameters were used. The length of the actual feedback path is $L_g = 32$ samples. The simulation run lasts for 40 seconds with instantaneous change of feedback path occurring at time 20 seconds. Speech or music is used as the incoming signal and the angle of arrival is zero. The step size for all the algorithms is $\mu = 0.0001$ and $\delta = \xi = 10^{-8}$. The filter length for $\hat{\mathbf{H}}(q)$ is $L_{\hat{h}} = 8$, $M = 8$ and $d_m = 3$. The sampling frequency is 16 kHz, and the forward path gain K = 30 dB with a forward path delay of $d_k = 32$ samples [13]. We compare the TM-DFT-I [13] with the proposed TM-P-DFT-II and TM-PS-DFT-II.

Figures 4 and 5 show the misalignment and MSG performances of TM-DFT-I, TM-P-DFT-II and TM-PS-DFT-II in the two microphone situation in case of feedback path variation at 20 seconds for the speech case and music case, respectively. The step sizes for the TM-PS-DFT-II are chosen in order to have the same convergence for the initial iterations as TM-P-DFT-II for the speech signal. It can be noticed from Fig. 4 that TM-P-DFT-II algorithm achieves a faster convergence rate and better MSG values than TM-DFT-I and TM-PS-DFT-II algorithms for the first feedback path. Most of the time the TM-PS-DFT-II algorithm has worse performance than the original TM-DFT-I algorithm. The superior performance of TM-P-DFT-II over TM-PS-DFT-II and TM-DFT-I algorithms is obvious for the music input case shown in Fig. 5.

Figures 6 and 7 shows the performances of the investigated algorithms in case of probe injection and feedback path variation at 20 seconds for the speech case and music case respectively.

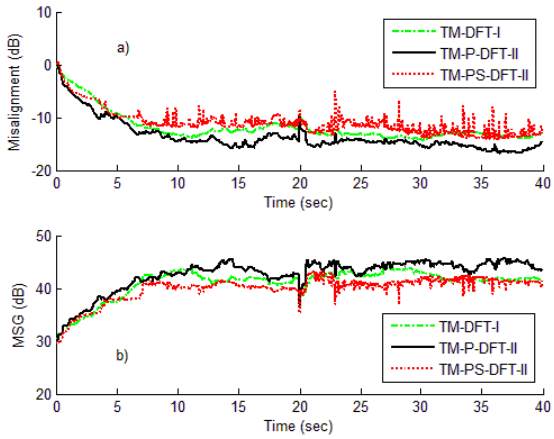

Fig. 4. (a) Misalignment and (b) MSG curves with speech input, feedback path variation at 20 seconds.

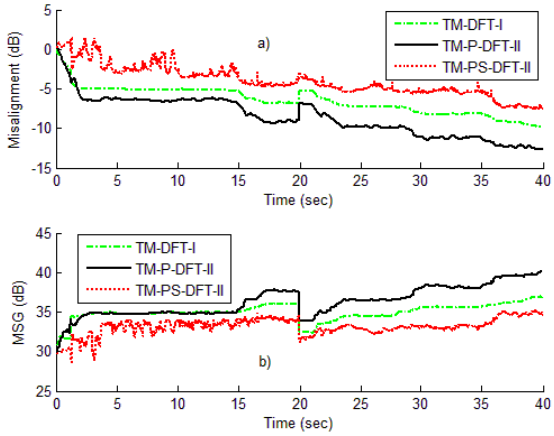

Fig. 5. (a) Misalignment and (b) MSG curves with music input, feedback path variation at 20 seconds.

The amplified signal to injected noise ratio is SINR = 20 dB. It can be noticed that the performance improvement of the proportionate algorithm differs from that obtained in the case without noise injection. In some cases (see Fig. 7) the performance of TM-P-DFT-II in terms of misalignment and MSG is better for the first feedback path but can be worse in some portions than that of TM-DFT-I for the second feedback path in case of music input. For the investigated cases the execution times of the TM-P-DFT-II is almost double than that of TM-DFT-I and about 30% higher than that of TM-PS-DFT-II. There is dependence of the steady-state performance of investigated algorithms on the input signal and gain (bias). Also, the noise probe injection affects the performance of the sign based algorithm, especially for music signals. The superior performance of PNLMS over the sign version is confirmed for other applications [15], [19] except those of impulsive environments [20].

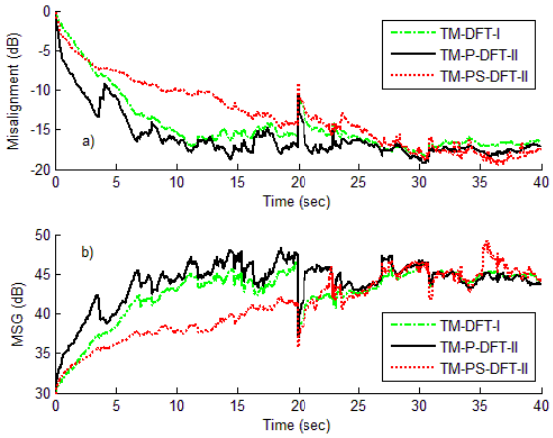

Fig. 6. (a) Misalignment and (b) MSG curves with speech input, probe injection and feedback path variation at 20 seconds.

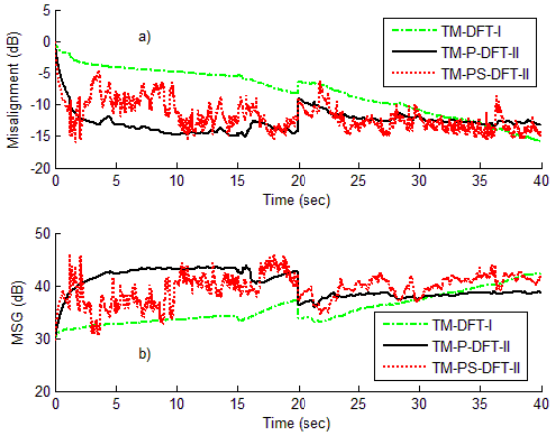

Fig. 7. (a) Misalignment and (b) MSG curves with music input, probe injection and feedback path variation at 20 seconds.

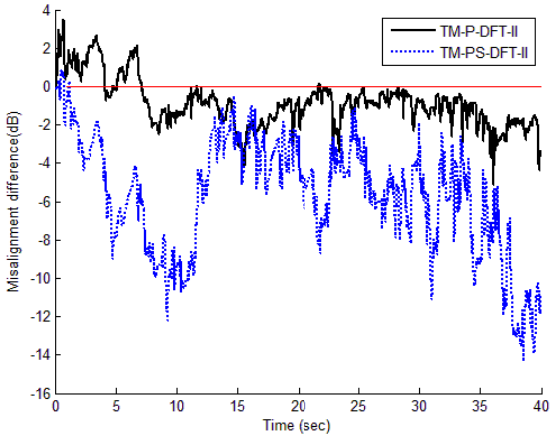

Fig. 8. Misalignment difference for 8 and 16 subbands respectively. Other conditions are those from Fig. 4.

Figure 8 show the misalignment difference between the results of the investigated algorithms for two values of the number of subbands (8 and 16) of TM-PS-DFT-II and TM-P-DFT-II when using the same signals and parameters from Fig. 4.The values above zero indicates that the convergence of the algorithms with 16 subbands is better than that obtained with 8 subbands. It can be noticed that while the initial convergence of TM-P-DFT-II with 16 subbands is better than that obtained with 8 subbands, the increase of the number of subbands worsen the performance of TM-PS-DFT-II. Therefore the number of subbands have an impact on the convergence performance but it should be correlated with the parameters of the proportionate algorithms. Other simulations (not shown here due to lack of space) have shown that the step size and regularization factor choices have also an impact on the performance of the proposed algorithms.

Table 1 shows the computed PESQ values for speech signals, with and without noise injection. It can be seen that TM-P-DFT-II algorithm achieves the best PESQ values among the considered algorithms. The TM-PS-DFT-II algorithm obtains better PESQ values than the TM-DFT-I algorithm for the music signal with probe injection. The PESQ scores for music signals were not reported since they can be inaccurate as pointed in [23].

	TM-DFT-I	TM-PS-DFT-II	TM-P-DFT-II
Speech / No Probe	4.36	4.27	**4.39**
Speech + Probe	3.99	4.03	**4.21**

Table 1. PESQ values

4. CONCLUSIONS

In this paper it is shown that a good compromise between numerical complexity and performance is obtained if proportionate algorithms are used for the two-microphone approach. Better performance are reported for investigated simulations using speech and music signals. Our future work will be focused on developing variable regularization and variable step size versions. Also, further studies are needed for performance optimization through subband range modification.

REFERENCES

[1] A. Spriet, S. Doclo, and M. Moonen, "Feedback control in hearing aids," Springer handbook of speech processing, 2008.

[2] M. Guo, S. H. Jensen, and J. Jensen, "Novel Acoustic Feedback Cancellation Approaches in Hearing Aid Applications Using Probe Noise and Probe Noise Enhancement," IEEE Trans. Audio, Speech Lang. Process., vol. 20, no. 9, pp. 2549–2563, 2012.

[3] T. van Waterschoot and M. Moonen, "Fifty years of acoustic feedback control: state of the art and future challenges," Proc.IEEE, no. 99, pp. 1–40, 2011.

[4] A. Spriet, I. Proudler, M. Moonen, and J. Wouters, "Adaptive feedback cancellation in hearing aids with linear prediction of the desired signal," IEEE Trans. Signal Process., vol. 53, pp. 3749–3763, Oct. 2005.

[5] C. R. C. Nakagawa, S. Nordholm, and W.-Y. Yan, "New Insights Into Optimal Acoustic Feedback Cancellation," IEEE Signal Process. Lett., vol. 20, pp. 869–872, Sept. 2013.

[6] C. R. C. Nakagawa, S. Nordholm, and W.-Y. Yan, "Feedback Cancellation With Probe Shaping Compensation," IEEE Signal Process. Lett., vol. 21, pp. 365–369, Mar. 2014.

[7] M. Guo, S. H. Jensen, J. Jensen, and S. L. Grant, "On the Use of a Phase Modulation Method for Decorrelation in Acoustic Feedback Cancellation," in Eur. Signal Process. Conf., Bucharest, Romania, Aug. 2012.

[8] S. Narayan and A. Peterson, "Frequency domain least-mean square algorithm," Proc. IEEE, vol. 69, no. 1, pp. 124–126, 1981.

[9] B. Farhang-Boroujeny and S. Gazor, "Selection of orthonormal transforms for improving the performance of the transform domain normalised LMS algorithm," IEE Proc. F Radar Signal Process., vol. 139, no. 5, p. 327, 1992.

[10] M. Rotaru, F. Albu, H. Coanda, "A Variable Step Size Modified Decorrelated NLMS Algorithm for Adaptive Feedback Cancellation in Hearing Aids, " in Proc. of IEEE ISETC 2012, Timisoara, Romania, Nov. 2012, pp. 263-266.

[11] J. M. Gil-Cacho, T. van Waterschoot, and M. Moonen, "Transform Domain Prediction Error Method for Improved Acoustic Echo and Feedback Cancellation," in Eur. Signal Process. Conf., Bucharest, Romania, Aug. 2012, pp. 2422–2426.

[12] C. R. C. Nakagawa, S. Nordholm, and W.-Y. Yan, "Dual microphone solution for acoustic feedback cancellation for assistive listening," in 2012 IEEE Int. Conf. Acoust. Speech Signal Process., Apr. 2012, pp. 149–152.

[13] C. R. C. Nakagawa, S. Nordholm, F. Albu and W.-Y. Yan, "Closed-loop feedback cancellation utilizing two microphones and transform domain processing", in IEEE Proc. of ICASSP 2014, Florence, Italy, P5.1, May 2014, pp. 3673-3677.

[14] F. Albu, H. Coanda, D. Coltuc, & M. Rotaru, "Intermittently Updated Simplified Proportionate Affine Projection Algorithm", in Proc. of ADAPTIVE 2014, Venice, Italy, May 2014, pp. 42-47.

[15] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancellers," IEEE Trans. Speech Audio Process., vol. 8, no. 5, pp. 508–518, Sept. 2000.

[16] D. F. Elliott and K. R. Rao, Fast transforms: algorithms, analyses, applications. Academic press New York, 1982.

[17] P. Duhamel and M. Vetterli, "Fast Fourier transforms: A tutorial review and a state of the art," Signal Processing,vol. 19, pp. 259–299, Apr. 1990.

[18] A. Sayed, Adaptive filters, John Wiley & Sons, 2008.

[19] J. Benesty and S. L. Gay, "An improved PNLMS algorithm", in Proc. of ICASSP 2002, Orlando, USA, May 2002, vol. 2, May 2002, pp. 1881 –1884.

[20] F. Albu, H.K. Kwan, "Memory Improved Proportionate Affine Projection Sign Algorithm", IET Electronics Letters, vol. 48, issue 20, pp. 1279-1281, Oct. 2012.

[21] Z. Yang, Y. R. Zheng, and S. L. Grant, "Proportionate affine projection sign algorithms for network echo cancellation", IEEE Transactions on Audio, Speech, and Language Processing, vol. 19, no. 8, pp. 2273-2284, Nov. 2011.

[22] I.-T. R. P. 862, "Perceptual Evaluation of Speech Quality (PESQ), an Objective Method for End-to-end Speech Quality Assessment of Narrowband Telephone Networks and Speech Codecs," 2001.

[23] J. G. Beerends, A. P. Hekstra, A. W. Rix, and M. P. Hollier, Perceptual Evaluation of Speech Quality (PESQ), the new ITU standard for end-to-end speech quality assessment.Part II – Psychoacoustic model, Journal Audio Eng. Soc, 2002.