ADAPTIVE THRESHOLD NONLINEAR CORRELATION ALGORITHM FOR ROBUST FILTERING IN IMPULSIVE NOISE ENVIRONMENTS

Shin'ichi Koike

Consultant, Tokyo, Japan   E-mail: shin-koike@mub.biglobe.ne.jp

ABSTRACT

In this paper, we first present mathematical models for two types of impulse noise in adaptive filtering systems; one in additive observation noise and another at filter input. To combat such impulse noise, a new algorithm named Adaptive Threshold Nonlinear Correlation Algorithm (ATNCA) is proposed. Through analysis and experiment, we demonstrate effectiveness of the ATNCA in making adaptive filters highly robust in the presence of both types of impulse noise while realizing convergence as fast as the LMS algorithm. Fairly good agreement between simulated and theoretical convergence behavior in transient phase and steady state proves the validity of the analysis.

1. INTRODUCTION

Among many adaptation algorithms for adaptive filters, the LMS algorithm is most widely used in practical systems. While we find the LMS algorithm attractive in many aspects, one of its weaknesses is vulnerability to disturbances, *e.g.* impulse noise found in DSL systems using metallic subscriber lines or in indoor wireless systems [1],[2].

We identify two types of impulse noise that may intrude adaptive filtering systems, seriously degrading filter performance and often making filters unstable. Typically, impulse noise is found in additive observation noise as exemplified above. To combat such impulse noise, sign algorithm (SA) [3], adaptive threshold nonlinear algorithm (ATNA) proposed by the author [4], least mean M-estimate algorithms [5], etc. are proven effective.

Often we find another type of impulse noise that is present at the input to adaptive filters. An example of such impulse noise is the one that has large peaks in the background noise (often 20 dB greater) in active noise cancellers for use in hands-free mobile phones. To combat such impulse noise at the filter input, normalized LMS algorithm, signed regressor LMS (SR-LMS) algorithm, etc. are known to be effective. However, these algorithms have poor robustness against the former type of impulse noise, *i.e.* impulsive observation noise.

Therefore, when both types of impulse noise are present, in the observation noise as well as at the filter input, none of the above listed algorithms is satisfactory. One possible solution is use of sign-sign (SS) algorithm that takes the signum of both the error and the regressor [6]. The SS algorithm exhibits sufficient robustness against both types of impulse noise, but a serious drawback is its considerably slow convergence speed.

In this paper, by applying the "thresholding" technique used in the ATNA, we propose and analyze a new algorithm named Adaptive Threshold Nonlinear Correlation Algorithm (ATNCA) that will effectively combat both types of impulse noise stated above, while realizing fast convergence.

2. IMPULSE NOISE MODELS

We identify two types of impulse noise; one is present in additive observation noise and the other is found at filter input. These are shown in Fig. 1 as $v(n)$ and $v_a(n)$.

First, the additive impulsive observation noise is often modeled as "Contaminated Gaussian Noise" (CGN) that is mathematically a combination of two independent Gaussian noise sources [4]. Namely,

$v_{<0>}(n)$: Gaussian noise source <0> with variance $\sigma_{v<0>}^2$ and probability of occurrence $p_{v<0>}$,

$v_{<1>}(n)$: Gaussian noise source <1> with variance $\sigma_{v<1>}^2$ and probability of occurrence $p_{v<1>}$.

Note that $p_{v<0>}+p_{v<1>}=1$ holds. Usually, $\sigma_{v<1>}^2 >> \sigma_{v<0>}^2$ and $p_{v<1>} < p_{v<0>}$. For "pure" Gaussian noise, $p_{v<1>}= 0$.

For the impulse noise at the filter input, we propose the following model. The filter input $b(n)$ is given by

$$b(n) = a(n) + \tau(n)\, v_a(n), \tag{1}$$

where n is time instant, and $\tau(n)$ is an independent Bernoulli trial taking 1 with probability of occurrence pv_a and 0 with $1- pv_a$. The impulse noise $v_a(n)$ itself is a *White & Gaussian* process with variance σv_a^2, independent of the reference input $a(n)$ whose variance is σ_a^2.

3. PERFORMANCE OF THE LMS ALGORITHM

In this section we revisit performance of LMS adaptive filters in the presence of both types of impulse noise. The tap weight update equation is given by

$$\mathbf{c}(n +1) = \mathbf{c}(n) + \alpha_c\, e(n)\, \mathbf{b}(n), \tag{2}$$

where $\mathbf{c}(n)$ is tap weight vector (N taps), $\mathbf{b}(n)=[b(n), \ldots, b(n-k), \ldots, b(n-N+1)]^T$ is filter input vector and α_c is step size. Error signal $e(n)$ is given by $e(n) = \boldsymbol{\theta}^T(n)\mathbf{b}(n) +v(n)$, where $\boldsymbol{\theta}(n) = \mathbf{h} - \mathbf{c}(n)$ is tap weight misalignment vector, $\mathbf{h}$

is response vector of stationary unknown system to be identified, and $v(n)$ is CGN in general.

For simulations (ensemble averaging over 5000 runs) with the LMS algorithm, we prepare:

Example #1 $N = 4$, $\mathbf{h} = [0.05, 0.994, 0.01, -0.1]^T$

W&G ref. input, $\sigma_a^2 = 1$ (0 dB), $\alpha_c = 2^{-8}$

Case 1: "pure" Gaussian noise $\sigma_{v<0>}^2 = 0.01$ (–20 dB)

no impulse noise at filter input $pv_a = 0$

Case 2: CGN $\sigma_{v<0>}^2 = 0.01$ (–20 dB); $p_{v<0>} = 0.9$

$\sigma_{v<1>}^2 = 10$ (+10 dB); $p_{v<1>} = 0.1$

no impulse noise at filter input $pv_a = 0$

Case 3: "pure" Gaussian noise $\sigma_{v<0>}^2 = 0.01$ (–20 dB)

impulse noise at filter input

$\sigma v_a^2 = 100$ (+20 dB); $pv_a = 0.1$

Case 4: CGN as in Case 2

impulse noise at filter input as in Case 3

The above parameter values for impulse noise are selected to reflect situations in practical systems [1].

Results of the simulations are shown in Fig. 2, where we observe considerable increase in steady-state ensemble average $\sigma_a^2 <\|\mathbf{\theta}(n)\|^2>$ for mean square tap weight misalignment (MSTWM) (see (10)) due to impulse noise. Note that the MSTWM increase is as large as 20 dB for the CGN. We can show that the LMS filter diverges if $\sigma v_a^2 > 161$ (with $pv_a = 0.1$) in theory.

4. ADAPTIVE THRESHOLD NONLINEAR CORRELATION ALGORITHM

To effectively combat both types of impulse noise while making filter convergence sufficiently fast, we propose Adaptive Threshold Nonlinear Correlation Algorithm.

Fig. 1 shows a schematic diagram of an adaptive filter using the proposed ATNCA. The tap weight $c_k(n)$ is adapted with the output of a *nonlinear threshold* device $f[\cdot]$ whose input $z_k(n)$ is *correlation* between the error $e(n)$ and the filter input at the kth tap $b(n-k)$ with $A_k^-(n)$ and $A_k^+(n)$ being *adaptive* thresholds.

The kth tap weight $c_k(n)$ is updated as

$$c_k(n+1) = c_k(n) + \alpha_c f[z_k(n) ; A_k^-(n), A_k^+(n)], \quad (3)$$

where $z_k(n) = e(n)b(n-k)$, $A_k^-(n)$ and $A_k^+(n)$ are lower and upper thresholds, respectively. We define

$$f(z ; A^-, A^+) = \begin{cases} z & \text{for } A^- \le z \le A^+ \\ 0 & \text{elsewhere.} \end{cases} \quad (4)$$

See Fig. 3 for a graph of the function $f(z ; A^-, A^+)$.

The upper and lower thresholds are adaptively controlled as follows.

$$A_k^\pm(n) = \mu z_k(n) \pm M_A \, \mathrm{v} z_k(n), \quad (5)$$
$$\mu z_k(n+1) = (1-\rho_\mu) \mu z_k(n) + \rho_\mu z_k(n), \text{ and} \quad (6)$$
$$\mathrm{v} z_k^2(n+1) = (1-\rho_\mathrm{v}) \mathrm{v} z_k^2(n) + \rho_\mathrm{v} [z_k(n) - \mu z_k(n)]^2, \quad (7)$$
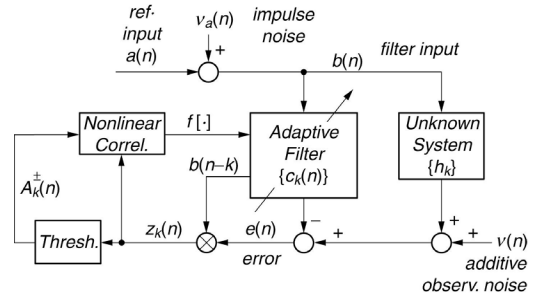

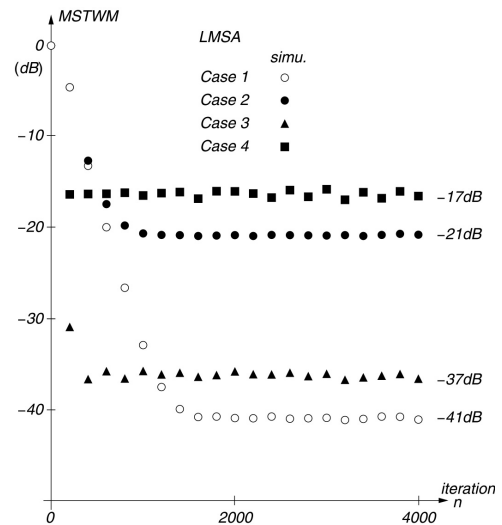Fig. 1   Schematic diagram of adaptive filtering system using ATNCA.

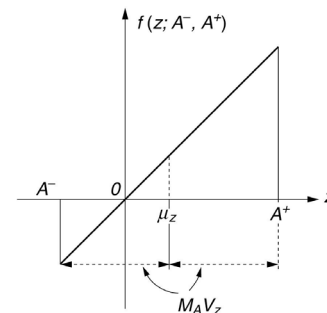Fig. 2   Adaptive filter convergence. LMS algorithm (*Example #1*)

Fig. 3   Graph of nonlinear correlation function.

where M_A is multiplier, $\mu z_k(n)$ and $\mathrm{v} z_k^2(n)$ are estimates of mean and variance of the correlation $z_k(n)$, respectively, and ρ_μ and ρ_v are leakage factors.

As in many cases, initial value of the tap weights is chosen $\mathbf{c}(0) = \mathbf{0}$. Initial thresholds $A_k^\pm(0)$ are determined by (5), (6) and (7) for zero tap weights prior to the tap weight adaptation, and usually have non-zero values.

5. ANALYSIS

5.1. Assumptions

For the transient analysis of the ATNCA to be developed in this section, we make the following assumptions.

A1) The reference input $a(n)$ is a stationary *White & Gaussian* process.

A2) The filter input and the tap weights are mutually independent (*Independence Assumption*).

A3) The error $e(n)$ and the filter input $b(n)$ are jointly Gaussian distributed [3].

5.2. "Gaussian Product Distribution"

Under Assumption A3), the correlation $z_k(n)$ is a product of two correlated Gaussian random variables [7]. The author names the probability distribution of this product "Gaussian Product Distribution" whose details are summarized in **APPENDIX**.

5.3. Difference Equations for Tap Weights

For the kth tap weight misalignment $\theta_k(n)$, we can derive the following difference equations of mean $m_k(n)=E[\theta_k(n)]$ and variance $R_k(n)=E\{[\theta_k(n) - m_k(n)]^2\}$.

$$m_k(n+1) = m_k(n) - \alpha_c p_k(n), \text{ and} \qquad (8)$$

$$R_k(n+1) = [1-2\alpha_c W_k(n)]R_k(n) + \alpha_c^2[T_k(n) - p_k^2(n)], \quad (9)$$

where

$$p_k(n) = \Sigma_{\text{all } \{\tau\}} \text{Pr}_{\{\tau\}} \Sigma_{i=0}^1 p_{v<i>} p_{\{\tau\}<i>k}(n),$$

$$W_k(n) = \partial p_k(n)/\partial m_k(n) \text{ [6]},$$

$$T_k(n) = \Sigma_{\text{all } \{\tau\}} \text{Pr}_{\{\tau\}} \Sigma_{i=0}^1 p_{v<i>} T_{\{\tau\}<i>k}(n).$$

Here, we define a set of the Bernoulli trials $\{\tau\}=\{\tau(n), \dots, \tau(n-N+1)\}$ and its probability of occurrence $\text{Pr}_{\{\tau\}}=(1-pv_a)^{n\tau 0} pv_a^{n\tau 1}$ in which $n\tau 0$ and $n\tau 1$ are numbers of 0's and 1's in the set $\{\tau\}$, respectively. Referring to **5.2.**, we calculate:

$$p_{\{\tau\}<i>k}(n) = \sigma_{e\{\tau\}<i>}(n) \sigma_{b\{\tau\}k}$$
$$\times\{F_Z[r^+_{\{\tau\}<i>k}(n);\alpha_{\{\tau\}<i>k}(n)] - F_Z[r^-_{\{\tau\}<i>k}(n);\alpha_{\{\tau\}<i>k}(n)]\},$$

$$T_{\{\tau\}<i>k}(n) = \sigma_{e\{\tau\}<i>}^2(n) \sigma_{b\{\tau\}k}^2$$
$$\times\{G_Z[r^+_{\{\tau\}<i>k}(n);\alpha_{\{\tau\}<i>k}(n)] - G_Z[r^-_{\{\tau\}<i>k}(n);\alpha_{\{\tau\}<i>k}(n)]\},$$

where

$$\sigma_{e\{\tau\}<i>}^2(n) = \Sigma_{l=0}^{N-1}\sigma_{b\{\tau\}l}^2 [m_l^2(n)+R_l(n)] + \sigma_{v<i>}^2,$$

$$\sigma_{b\{\tau\}k}^2 = \sigma_a^2 + \tau(n-k) \sigma_{v_a}^2,$$

$$r^\pm_{\{\tau\}<i>k}(n) = E[A_k^\pm(n)] / [\sigma_{e\{\tau\}<i>}(n) \sigma_{b\{\tau\}k}], \text{ and}$$

$$\sin[\alpha_{\{\tau\}<i>k}(n)] = [\sigma_{b\{\tau\}k} /\sigma_{e\{\tau\}<i>}(n)]m_k(n).$$

See **APPENDIX** for the functions $F_Z(r\,;\,\alpha)$ and $G_Z(r\,;\,\alpha)$.

In order to measure adaptive filter performance in the transient phase and in the steady state as well, we define mean square tap weight misalignment (MSTWM) as

$$\xi(n) = \sigma_a^2 E[\|\boldsymbol{\theta}(n)\|^2]$$
$$= \sigma_a^2\Sigma_{l=0}^{N-1}[m_l^2(n)+R_l(n)]. \qquad (10)$$

5.4. Difference Equations for Thresholds

From (5), (6) and (7), we find

$$E[A_k^\pm(n)] = E[\mu z_k(n)] \pm M_A E[vz_k^2(n)]^{1/2}, \qquad (11)$$

$$E[\mu z_k(n+1)] = (1-\rho_\mu)E[\mu z_k(n)] +\rho_\mu E[z_k(n)], \text{ and} \quad (12)$$

$$E[vz_k^2(n+1)] = (1-\rho_v)E[vz_k^2(n)]$$
$$+\rho_v\{E[z_k^2(n)]-2E[z_k(n)]E[\mu z_k(n)]+E^2[\mu z_k(n)]\}, \quad (13)$$

where we calculate

$$E[z_k(n)] = \beta v_a \sigma_a^2 m_k(n), \text{ and}$$

$$E[z_k^2(n)] = \beta v_a \sigma_a^2 [\beta v_a \xi(n) + \sigma_v^2]$$
$$+(2\beta v_a^2 +3\eta v_a)[m_k^2(n)+R_k(n)]$$

with $\beta v_a = 1 + pv_a \sigma_{v_a}^2/\sigma_a^2$ and $\eta v_a = (1- pv_a) pv_a \sigma_{v_a}^4/\sigma_a^4$. Here, $\sigma_v^2 = \Sigma_{i=0}^1 p_{v<i>} \sigma_{v<i>}^2$ is "average" CGN variance.

6. EXPERIMENT

In this section, experiment is carried out with simulations and theoretical calculations of adaptive filter convergence using the ATNCA. The example used is basically the same as *Example #1* in Section **3**.

Example #2 filter parameters: same as *Example #1*
for the ATNCA: M_A=1.25, ρ_μ= ρ_v = 2^{-8}
Cases 1 to 4: same as *Example #1*

Fig. 4 shows the results of the experiment for Case 1 (no impulse noise), where convergence curves of the MSTWM and the thresholds $A_1^\pm$ are plotted. We calculate the theoretical convergence with the difference equations developed in Section **5**. As seen in the figure, the MSTWM converges as fast as that of the LMS algorithm, and the thresholds have fairly large initial magnitudes but they converge to smaller values equal with opposite polarities. Also we observe that the theoretical convergence curves are in good agreement with the simulated ones.

Fig. 5 depicts results of the experiment for Cases 1 to 4. For the CGN (Case 2), the increase in the MSTWM is about 5 dB that is much smaller than the 20 dB increase with the LMS algorithm, whereas in the presence of impulse noise at the filter input (Case 3) the MSTWM increase remains within a fraction of dB. Even if both types of impulse noise are present (Case 4), the MSTWM increase does not exceed 8 dB, an improvement of more than 16 dB in comparison with the LMS algorithm. The theoretical convergence mostly exhibits good agreement with the simulation data. This proves the validity of the analysis for practical use in filter design.

7. CONCLUSION

In this paper, we have presented mathematical models for two types of impulse noise; one in additive observation noise and another at filter input. To successfully combat both types of impulse noise and to ensure fast convergence of the filter, we have proposed a new algorithm named Adaptive Threshold Nonlinear Correlation Algorithm (ATNCA) which adds a moderate increase of computational complexity to the LMS algorithm (see (5), (6) and (7)).
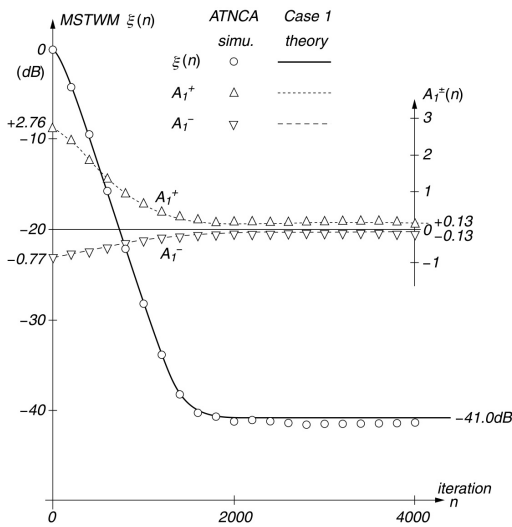
Fig. 4   Adaptive filter convergence.
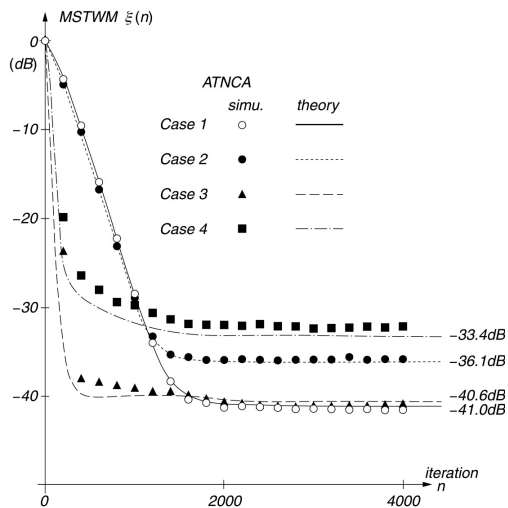ATNCA (*Example #2*; Case 1)

Fig. 5   Adaptive filter convergence.
ATNCA (*Example #2*; Cases 1 to 4)

Analysis of the ATNCA has been developed to derive a set of difference equations to calculate transient and steady-state behavior in terms of mean square tap weight misalignment (MSTWM). Through experiment we have demonstrated how effectively the proposed ATNCA works in combating the impulse noise. Also we have observed fairly good agreement between simulations and theoretical calculations that has proven the validity of the analysis.

The paper has compared the ATNCA only with the LMS algorithm. Comparison with other robust algorithms, *e.g.* normalized sign algorithm, is left as a future work.

APPENDIX

"Gaussian Product Distribution"
Probability Distribution of Product of Two Correlated Gaussian Random Variables [7]

Let x and y be zero mean Gaussian random variables with variance $\sigma_x^2 = E(x^2)$, $\sigma_y^2 = E(y^2)$ and covariance $R_{xy} = E(xy)$. Define a product $z = xy$. Normalizing z as $\zeta = z / (\sigma_x \sigma_y)$ and defining a correlation parameter $\sin \alpha = R_{xy} / (\sigma_x \sigma_y)$, we find probability density function (pdf) of ζ

$$p_Z(\zeta \,;\, \alpha) = \sec \alpha \, \exp(\zeta \sec^2 \alpha) \, p_H(\zeta \sec^2 \alpha),$$

where

$$p_H(\eta) = \pi^{-1} K_0(|\eta|)$$
$$= \pi^{-1} \int_0^1 u^{-1} \exp[-(|\eta|/2)(u+u^{-1})]du,$$

$K_0(x)$ zero-th order *Modified Bessel Function* of the second kind [8].

Next, we define the following functions related to the "Gaussian Product Distribution."

$$Q_Z(\zeta \,;\, \alpha) = \int_{-\infty}^{\zeta} p_Z(\lambda \,;\, \alpha)d\lambda, \quad prob.\ distribution\ func.$$
$$F_Z(\zeta \,;\, \alpha) = \int_{-\infty}^{\zeta} \lambda\, p_Z(\lambda \,;\, \alpha)d\lambda, \text{ and}$$
$$G_Z(\zeta \,;\, \alpha) = \int_{-\infty}^{\zeta} \lambda^2\, p_Z(\lambda \,;\, \alpha)d\lambda.$$

It can easily be shown that the above functions are calculated by definite integrals over $[0, 1]$ with respect to u.

REFERENCES

[1] D. Middleton, "Man-made noise in urban environments and transportation systems: models and measurements," *IEEE Trans. Commun.*, vol.21, pp.1232-1241, Nov. 1973.

[2] K. L. Blackard, T. S. Rappaport, and C. W. Bostian, "Measurements and models of radio frequency impulsive noise for indoor wireless communications," *IEEE J. Select. Areas Commun.*, vol.11, pp.991-1001, Sept. 1993.

[3] V. J. Mathews and S. H. Cho, "Improved convergence analysis of stochastic gradient adaptive filters using the sign algorithm," *IEEE Trans. Acoust., Speech and Signal Processing*, vol.35, pp.450-454, Apr. 1987.

[4] S. Koike, "Adaptive threshold nonlinear algorithm for adaptive filters with robustness against impulse noise," *IEEE Trans. Signal Processing*, vol.45, pp.2391-2395, Sept. 1997.

[5] Y. Zou, S. C. Chan, and T. S. Ng, "Least mean M-estimate algorithms for robust adaptive filtering in impulse noise," *IEEE Trans. Circuits Syst. II*, vol.47, pp.1546-1569, Dec. 2000.

[6] S. Koike, "Analysis of the sign-sign algorithm based on Gaussian distributed tap weights," *IEICE Trans. Fundamentals*, vol.E83-A, pp.1551-1558, Aug. 2000.

[7] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, Singapore: McGraw-Hill, 1984.

[8] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, New York: Dover, 1964.