NORMALIZED CORRELATION-NEWTON ALGORITHM WITH VARIABLE CONTROL OF q-NORM

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

ABSTRACT

This paper proposes a new adaptation algorithm named *normalized correlation-Newton* (NC-Newton) algorithm and a novel *variable q-norm* control method (NC-Newton-Var$q$-norm) for complex-domain adaptive filters. First, stochastic models are presented for two types of impulse noise intruding adaptive filters: one is present in observation noise and another at filter input. After reviewing q-norm and NC-Newton algorithm, we propose a variable q-norm control method. Analysis of the NC-Newton-Var$q$-norm algorithm is developed for theoretically calculating transient and steady-state convergence behavior. Through experiment with some examples, we demonstrate effectiveness of the proposed variable q-norm control method in improving filter convergence speed while preserving robustness of the NC-Newton algorithm in impulsive noise environments. Good agreement between simulated and theoretical convergence behavior validates the analysis.

Index Terms— Adaptive filter, normalized correlation, impulse noise, Newton's method, variable q-norm control.

1. INTRODUCTION

Adaptive filtering is one of the core technologies that play a crucial role in implementing essential functions and realizing required performance in many latest info/commun systems.

The LMS algorithm for adaptive filters is most intensively studied and most extensively applied to practical systems [1], [2]. The LMS algorithm has born many "children" such as the NLMS algorithm, the sign algorithm (SA), the sign-sign algorithm (SSA), and others [3]-[7].

Although the LMS algorithm is attractive, it is known to be vulnerable to impulse noise that intrudes adaptive filtering systems [8], [9]. We identify two types of impulse noise: one is present in observation noise and another at filter input. The latter type of impulse noise is sometimes found in such applications as "active noise cancellation."

In the real-number domain, the SSA is highly robust against both types of impulse noise stated above. In the complex-number domain, least mean modulus (LMM) algorithm and correlation phase algorithm (CΦA) are robust algorithms [10], [11]. The author proposed a "normalized" type adaptation algorithm named *normalized correlation* (NC) algorithm that is also highly robust against *both* types of impulse noise [12].

When the reference input is highly correlated, the filter convergence becomes considerably slower. To solve this problem, we introduce an estimate of the *inverse* covariance matrix in the tap weight adaptation. Simple recurrence calculation of the inverse covariance matrix is possible using the well-known Newton's method. The least mean modulus-Newton (LMM-Newton) algorithm is the LMM algorithm combined with the Newton's method [13]. Likewise, *normalized correlation-Newton* (NC-Newton) algorithm effectively improves the filter convergence for a correlated input. However, still faster convergence for the NC-Newton algorithm is desired.

In general, for "normalized" type algorithms, the normalizing factor can be "q-norm (or l_q norm)" of the filter input. q-norm of a vector $\mathbf{x}$, denoted by $\| \mathbf{x} \|_q$, is calculated as $[\sum_{l=0}^{N-1} | x(l) |^q]^{1/q}$ for $q \geq 1$. For the NLMS algorithm, the normalizing factor is "2-norm square."

In this paper, we propose NC-Newton algorithm with normalization by q-norm of the filter input, and a novel *variable q-norm* control method to further improve the filter convergence, yielding "NC-Newton-Var$q$-norm" algorithm.

2. IMPULSE NOISE MODELS

2.1. Impulsive Observation Noise

Impulse noise found in the additive observation noise is often modeled as *contaminated Gaussian noise* (CGN) that is mathematically a combination of two independent Gaussian noise sources, i.e., noise source #0 with variance $\sigma^2_v{}^{(0)}$ and probability of occurrence $p_v^{(0)}$, and noise source #1 with variance $\sigma^2_v{}^{(1)}$ and probability of occurrence $p_v^{(1)}$. Note that $p_v^{(0)} + p_v^{(1)} = 1$ holds. Usually, $\sigma^2_v{}^{(1)} >> \sigma^2_v{}^{(0)}$ and $p_v^{(1)} < p_v^{(0)}$. For "pure" Gaussian noise, $p_v^{(1)} = 0$ and $\sigma^2_v = \sigma^2_v{}^{(0)}$ [14].

2.2. Impulse Noise at Filter Input

A "noisy" filter input $b(n)$ at time instant n with impulse noise added to the reference input $a(n)$ is given by $b(n) = a(n) + \tau(n)\ va(n)$, where $\tau(n)$ is an independent Bernoulli random variable taking 0 with probability of occurrence $1 - p_{va}$ and 1 with p_{va}. The impulse noise $va(n)$ itself is an independent White & Gaussian noise with variance σ^2_{va}.

3. q-NORM AND NC-NEWTON ALGORITHM

3.1. q-Norm

Let $\mathbf{x} = [x(0), \cdots, x(k), \cdots, x(N-1)]^T$ be a complex-valued vector. q-norm (or l_q norm) of the vector $\mathbf{x}$ is defined by

$$\| \mathbf{x} \|_q = [\sum_{l=0}^{N-1} | x(l) |^q]^{1/q}, \qquad (1)$$

where $| \cdot |$ is *modulus* of a complex number and $q \geq 1$. In (1), typical values of q for practical use are: $q = 1$, 2 and ∞. $\| \mathbf{x} \|_2$ is Euclidean norm and $\| \mathbf{x} \|_\infty$ is called "infinity-norm." It can be shown that the infinity-norm is calculated as $\| \mathbf{x} \|_\infty = \max\{| x(0) |, \cdots, | x(k) |, \cdots, | x(N-1) |\}$.

Next, let $x(\cdot)$ be a complex-valued zero-mean Gaussian process, colored in general, with covariance matrix $\mathbf{R}_\mathbf{x} = E(\mathbf{x}\mathbf{x}^H)/2$ and unit variance. Let us denote expectation of q-norm of $\mathbf{x}$ and its square by $\gamma(q) = E(\| \mathbf{x} \|_q)$ and $\gamma_2(q) = E(\| \mathbf{x} \|_q^2)$. Note that $\gamma(q)$ or $\gamma_2(q)$ depends on N and the covariance matrix $\mathbf{R}_\mathbf{x}$. Generally, it is difficult to calculate $\gamma(q)$ or $\gamma_2(q)$ analytically. However, we find $\gamma(1) = (\pi/2)^{1/2}N$ and $\gamma_2(2) = 2N$.

For a general value of q, we run simulation to calculate ensemble average $<\| \mathbf{x} \|_q>$ and $<\| \mathbf{x} \|_q^2>$ for $\gamma(q)$ and $\gamma_2(q)$, respectively. An example of simulation results is given below, where $N = 32$ and $x(\cdot)$ is an AR1 Gaussian process with regression coefficient $\eta = 0.9$.

Table 1. Simulation results for q-norm.

q	$<\| \mathbf{x} \|_q>$	$<\| \mathbf{x} \|_q^2>$	$<\| \mathbf{x} \|_q>/<\| \mathbf{x} \|_q^2>$
1	40.11	1716	0.0234
2	7.773	64.00	0.121
4	3.703	14.45	0.256
8	2.756	7.980	0.345
32	2.436	6.230	0.391
128	2.417	6.132	0.394
∞	2.416	6.128	0.394

In Table 1, the ratio $<\| \mathbf{x} \|_q>/<\| \mathbf{x} \|_q^2> \approx \gamma(q)/\gamma_2(q)$ increases as q increases to converge to an upper bound.

3.2. NC-Newton Algorithm

First, we define correlation between the error and the reference input at the kth tap as

$$z_k(n) = e^*(n)\, a(n-k).$$

Then, we form a correlation vector

$$\mathbf{z}(n) = e^*(n)\, \mathbf{a}(n),$$

where $\mathbf{a}(n) = [a(n), \cdots, a(n-k), \cdots, a(n-N+1)]^T$ is the reference input vector. Using this vector $\mathbf{z}(n)$, we derive a tap weight update equation for the NC algorithm as given by

$$\mathbf{c}(n+1) = \mathbf{c}(n) + \alpha_c\, \mathbf{z}(n) / \| \mathbf{z}(n) \|_q \ \ [12],$$

where $\mathbf{c}(n)$ is the tap weight vector and α_c is the step size.

The error signal is given by $e(n) = \epsilon(n) + v(n)$. $\epsilon(n) = \boldsymbol{\theta}^H(n)\mathbf{a}(n)$ is the excess error, $\boldsymbol{\theta}(n) = \mathbf{h} - \mathbf{c}(n)$ is the tap weight misalignment vector, $\mathbf{h}$ is the response vector of an unknown system, and $v(n)$ is the additive observation noise.

When the reference input $a(n)$ is highly correlated, the filter convergence becomes considerably slow. For "de-correlation," we calculate an estimate of the *inverse* covariance matrix $\mathbf{R}_\mathbf{a}^{-1}$ of the input. In this paper, as in [13], we apply the well-known Newton's method to calculation of $\mathbf{R}_\mathbf{a}^{-1}$, yielding NC-Newton algorithm.

The tap weight update equation is given by

$$\mathbf{c}(n+1) = \mathbf{c}(n) + \alpha_c\, \mathbf{P}(n)\, \mathbf{z}(n) / \| \mathbf{z}(n) \|_q, \qquad (2)$$

where $\mathbf{P}(n)$ is the estimate of $\mathbf{R}_\mathbf{a}^{-1}$ recurrently calculated as

$$\mathbf{P}(n+1) = (1+\rho)\mathbf{P}(n) - \rho\, \mathbf{g}(n)\mathbf{g}^H(n)/[\| \mathbf{a}(n) \|_2^2/N]$$

with $\mathbf{g}(n) = \mathbf{P}(n)\mathbf{a}(n)$ and ρ being an adaptation coefficient.

4. NC-NEWTON ALGORITHM WITH VARIABLE CONTROL OF q-NORM

As is demonstrated in Section 6, when $q = 1$ for the q-norm, the steady-state error for the NC-Newton algorithm can be made small enough by selecting an appropriate step-size value, but the filter convergence becomes very slow as a trade-off. However, when we select $q = \infty$, the filter convergence with the same step size is much faster, but the steady-state error becomes very large (see Fig. 2).

The results above inspire us to vary the value of q so that initially $q = \infty$ (or 128) and $q = 1$ close to the steady state. Based on this idea, we propose a novel "variable q-norm" control method to combine with the NC-Newton algorithm.

In (2), q is now time-variant as given by

$$q(n) = q_0 + (q_\infty - q_0) \exp[-g_s P_s^m(n)],$$

where g_s is a coefficient, $P_s(n)$ is an "error variance estimator" and m (≥ 1) is a power of $P_s(n)$. Practically, we select $q_0 = 128$, $q_\infty = 1$ and $m = 2$. The estimator $P_s(n)$ is calculated by $P_s(n) = \| \mathbf{s}(n) \|_2^2$ where $\mathbf{s}(n+1) = (1-\rho_s)\mathbf{s}(n) + \rho_s \mathbf{P}(n)\, \mathbf{z}(n)/\| \mathbf{z}(n) \|_{q(n)}$ and ρ_s is a leakage factor. We name this algorithm NC-Newton-Var$q$-norm algorithm.

5. ANALYSIS

In this section, we develop analysis of the NC-Newton-Var$q$-norm algorithm. Due to space limitations, detailed derivation process cannot be fully described, but only main results are summarized. However, the validity of the analysis will be verified through experiment in Section 6.

5.1. Assumptions

In the analysis, *Long Filter Assumption* (N>>1) and *Independence Assumption* are adopted. Further, we assume that the matrix $\mathbf{P}(n)$ is uncorrelated with the correlation $\mathbf{z}(n)$.

5.2. Difference Equations for Tap Weight Misalignment

First, from (2), we find an update equation for the tap weight misalignment vector as

$$\boldsymbol{\theta}(n+1) = \boldsymbol{\theta}(n) - \alpha_c\, \mathbf{P}(n)\, e^*(n) \,/\, |\, e(n)\,| \cdot \mathbf{a}(n) \,/\, \|\, \mathbf{a}(n)\, \|_q, \quad (3)$$

From (3), defining $\mathbf{m}(n) = E[\boldsymbol{\theta}(n)]$ and $\mathbf{K}(n) = E[\boldsymbol{\theta}(n)\boldsymbol{\theta}^H(n)]$, we derive the following difference equations:

$$\mathbf{m}(n+1) = \mathbf{m}(n) - \alpha_c\, \mathbf{p}(n)$$

$$\mathbf{K}(n+1) = \mathbf{K}(n) - \alpha_c[\mathbf{V}(n)+\mathbf{V}^H(n)] + \alpha^2_c\mathbf{T}(n),$$

where $\mathbf{p}(n) = E[\mathbf{P}(n)]\mathbf{W}(n)\mathbf{m}(n)$, $\mathbf{V}(n) = E[\mathbf{P}(n)]\mathbf{W}(n)\mathbf{K}(n)$, $\mathbf{T}(n)=E[\mathbf{P}(n)]\mathbf{Ta}E[\mathbf{P}(n)]$, $\mathbf{W}(n)\approx(\pi/2)^{1/2}\mathbf{Ra}/\sigma_{eCGN}(n)/[\gamma(q)\sigma_a]$, $\mathbf{Ta} \approx 2\mathbf{Ra}/[\gamma_2(q)\sigma^2_a]$, $\sigma_{eCGN}(n) = 1 / \sum_{i=0}^{1} p_v^{(i)}/\sigma_e^{(i)}$ and $\sigma^2_e{}^{(i)} = \varepsilon(n)+\sigma^2_v{}^{(i)}$ for a Gaussian process $a(n)$. Here, $\varepsilon(n) = E[|\epsilon(n)|^2]/2 = \mathrm{tr}[\mathbf{Ra}\mathbf{K}(n)]$ is *excess mean square error* (EMSE).

5.3. Newton's Method

We can derive a difference equation for $E[\mathbf{P}(n)]$ as

$$E[\mathbf{P}(n+1)] = (1+\rho)E[\mathbf{P}(n)] - \rho E[\mathbf{P}(n)]\mathbf{Qa}E[\mathbf{P}(n)],$$

where, referring to [15], we calculate

$$\mathbf{Qa} = 2N \int_0^\infty u\, |\mathbf{A}(u)|^{-1}\, \mathbf{D}(u)\, du$$

with $\mathbf{D}(u) = \mathbf{A}^{-1}(u)\, \mathbf{Ra}$ and $\mathbf{A}(u) = \mathbf{I} + u^2\, \mathbf{Ra}$.

5.4. Variable *q*-Norm Control Method

We calculate expectation of "error variance estimator" as

$$E[P_s(n)] = E[\|\, \mathbf{s}(n)\, \|_2^2],$$

where for $E[\mathbf{s}(n)]$ and $E[\|\mathbf{s}(n)\|_2^2]$ we have

$$E[\mathbf{s}(n+1)] = (1-\rho_s)E[\mathbf{s}(n)] + \rho_s\, \mathbf{p}(n) \text{ and}$$

$$E[\|\mathbf{s}(n+1)\|_2^2] = (1-\rho_s)^2\, E[\|\mathbf{s}(n)\|_2^2] + 2(1-\rho_s)\rho_s\, \mathbf{p}^H(n)E[\mathbf{s}(n)] + \rho^2_s\, \mathrm{tr}[\mathbf{T}(n)].$$

5.5. Steady-State Solution

Assuming filter convergence as $n\rightarrow\infty$, with $E[\mathbf{P}(\infty)] = \mathbf{Qa}^{-1}$ and $q(\infty) = 1$, we derive the following equation to solve the steady-state EMSE iteratively.

$$\varepsilon(\infty) \approx \alpha_c(2/\pi)^{1/2}[\gamma(1)/\gamma_2(1)]\, \mathrm{tr}(\mathbf{Ra}\mathbf{Qa}^{-1})/\sigma_a \cdot \sigma_{eCGN}(\infty).$$

6. EXPERIMENT

In this section, we present results of experiment, where we calculate simulated and theoretical filter convergence behavior for NC, NC-Newton, NC-Newton-Var*q*-norm and NLMS algorithms. The simulation result is an ensemble average of squared excess error $<|\epsilon(n)|^2>/2$ over 1000 independent runs of filter convergence.

Four examples are carefully prepared. In the examples, $N = 32$, the reference input is an AR1 Gaussian process with variance $\sigma^2_a = 1$ (0 dB) and regression coefficient $\eta = 0.9$ or 0. The step size is $\alpha_c = 2^{-9}$. For Newton's method, $\rho = 2^{-8}$. In *Examples* #1 to #3, noise is "pure" Gaussian noise with $p_v^{(1)} = 0$; $\sigma^2_v = 0.01$ (–20 dB) and no impulse noise is present at filter input.

Example #1 q-norm: $q = 1$
NC, NC-Newton and NLMS algorithms
NLMS algorithm: $\eta = 0$ (W&G) and
step size $\alpha_c = 2^{-6}$

Example #2 q-norm: $q = 1$ and 128
NC-Newton algorithm

Example#3 NC-Newton-Var*q*-norm algorithm
variable q-norm control: $q_0 = 128$, $q_\infty = 1$,
$g_s = 2$, $\rho_s = 2^{-11}$

Example #4 NC-Newton-Var*q*-norm and NLMS algorithms
variable q-norm control: same as above
NLMS algorithm: same as in *Example #1*
Case 1: "pure" Gaussian noise
no impulse noise at filter input
Case 2: CGN $\sigma^2_v{}^{(0)} = 0.01$; $p_v^{(0)} = 0.9$
$\sigma^2_v{}^{(1)} = 10$; $p_v^{(1)} = 0.1$
Case 3: "pure" Gaussian noise
impulse noise at filter input
$\sigma^2_{va} = 1000$ (+30 dB); $p_{va} = 0.1$
Case 4: CGN as in Case 2
impulse noise at filter input
as in Case 3

For *Example #1*, filter convergence curves for NC algorithm are depicted in Fig. 1 for $q = 1$. Clearly, we observe that the convergence for η=0 (W&G) is much faster than that for a highly correlated input (η=0.9). For NC-Newton algorithm with η=0.9, the convergence becomes as fast as that for NC algorithm with η=0, showing the effectiveness of the Newton's method. However, we see still slower convergence than for the NLMS algorithm with η=0 and the step size $\alpha_c = 2^{-6}$ which gives $\varepsilon(\infty)$ of about –40 dB.

For *Example #2*, convergence curves for NC-Newton algorithm with $q = 1$ and 128 are compared in Fig. 2. Here, clearly the convergence for $q = 128$ is much faster, but the steady-state EMSE becomes much larger than for $q = 1$.

Fig. 3 shows filter convergence for *Example #3*, where the value of q for the q-norm is varied according to the proposed control method (NC-Newton-Var*q*-norm algorithm). The figure also shows how $q(n)$ is varied from the initial value 128 to the final value 1. We observe that the variable q-norm control method significantly improves the filter convergence speed for the NC-Newton algorithm with a fixed q-norm (q=1).

In *Example #4*, either or both types of impulse noise are present for NC-Newton-Var*q*-norm algorithm (Cases 2 to 4). In Case 2, for CGN, the increase in the EMSE from that for Case 1 is only a fraction of dB. For Cases 3 and 4, only simulation results are given. Even though the convergence becomes slower (reason unknown), the steady-state EMSE is much smaller than that for Case 1 or 2. These results demonstrate the high robustness of the algorithm against both types of impulse noise. It is observed that the filter convergence for NC-Newton-Var*q*-norm algorithm is even faster than that for the NLMS algorithm with a W&G input, showing the effectiveness of the proposed method.

In the examples above, we observe good agreement between simulation and theory that validates the analysis.

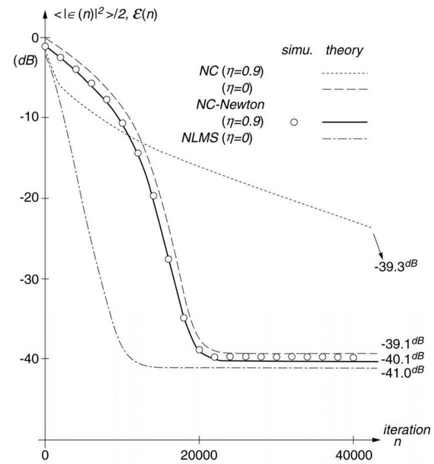

Fig. 1. Adaptive filter convergence (*Example #1*, p=1).

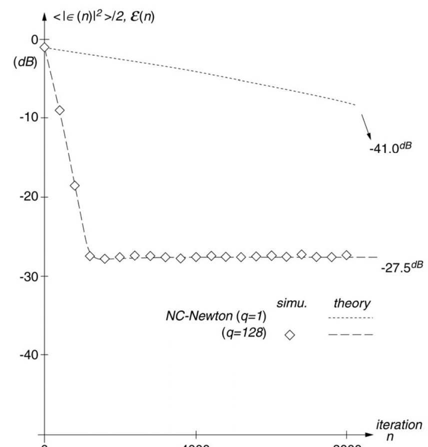

Fig. 2. Adaptive filter convergence (*Example #2*, p=1 & 128).

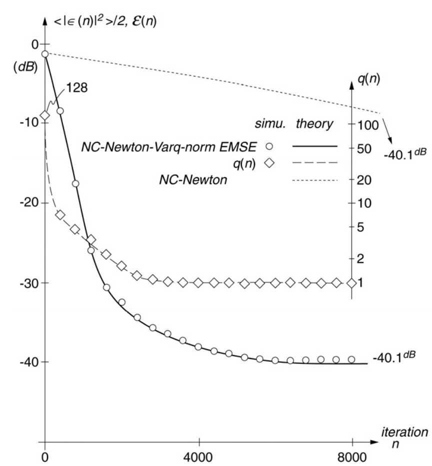

Fig. 3. Adaptive filter convergence (*Example #3*, Var$q$-norm).

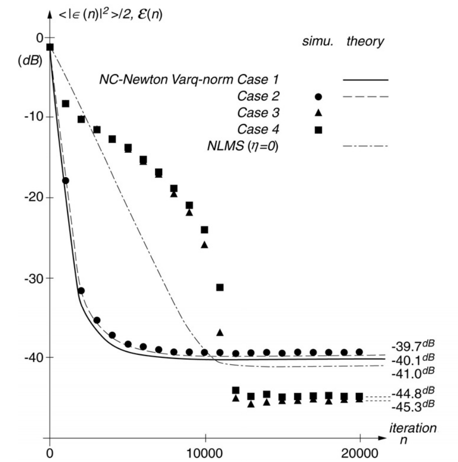

Fig. 4. Adaptive filter convergence (*Example #4*, Cases 1 to 4).

7. CONCLUSION

In this paper, we have proposed normalized correlation-Newton algorithm and a novel variable q-norm control method. Through analysis and experiment, effectiveness of the proposed NC-Newton-Var$q$-norm algorithm has been demonstrated in successfully improving the convergence speed while preserving the robustness against both types of impulse noise. Theoretical analysis in the presence of impulse noise at filter input is left as a future work.

8. RELATION TO PRIOR WORK

In order to improve filter convergence speed for adaptation algorithms, many variable (or adaptive) step-size control methods have been proposed. The paper by Kwong et al. may be the first that proposed a variable step size based on the error power [16]. Then, various methods have been proposed to improve performance in practical situations [17] – [20], [10] [11].

Recursive least type algorithms and affine projection algorithms are also effective in improving the filter convergence [2], [21], [22]. However, for these algorithms computational complexity is normally very high.

So far, to the best of the author's knowledge, there have been no other methods proposed for accelerating filter convergence than those stated above. Therefore, this paper proposes a completely novel approach in which we aim at achieving faster convergence by varying the value of q for q-norm of the normalizing factor.

9. REFERENCES

[1] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. New Jersey: Prentice-Hall, 1984.

[2] S. Haykin, *Adaptive Filter Theory*. New Jersey: Prentice-Hall, 1996.

[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, no. 4, pp. 450-454, Apr. 1987.

[4] S. Koike, "Analysis of adaptive filters using normalized signed regressor LMS algorithm," *IEEE Trans*. *Signal Processing*, vol. 47, no. 10, pp. 2710-2723, Oct. 1999.

[5] S. Koike, "Adaptive threshold nonlinear algorithm for adaptive filters with robustness against impulse noise," *IEEE Trans*. *Signal Processing*, vol. 45, no. 9, pp. 2391-2395, Sept. 1997.

[6] 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*, *Analog Digit*. *Signal Process*., vol. 47, no. 12, pp. 1564-1569, Dec. 2000.

[7] S. Koike, "Analysis of the sign-sign algorithm based on Gaussian distributed tap weights," in *Proc*. *IEEE ICASSP 98*, May 1998, pp. 1673-1676.

[8] D. Middleton, "Man-made noise in urban environments and transportation systems: models and measurements," *IEEE Trans*. *Commun*., vol. 21, no. 11, pp. 1232-1241, Nov. 1973.

[9] 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.

[10] S. Koike, "Variable step-size least mean modulus algorithm for robust filtering in digital QAM systems," in *Proc*. *EUSIPCO 2007*, Sept. 2007, pp. 193-197.

[11] S. Koike, "Variable step-size correlation phase algorithm: analysis and simulations," in *Proc*. *IEEE ISPACS 2007*, Nov.-Dec. 2007, pp. 17-20.

[12] S. Koike, "Analysis of normalized correlation algorithm for adaptive filters in impulsive noise environments," in *Proc*. *EUSIPCO 2011*, Aug.-Sept. 2011, pp. 141-145.

[13] S. Koike, "Performance analysis of least mean modulus-Newton algorithm," in *Proc*. *IEEE ISPACS 2009*, Dec. 2009, pp. 413-416.

[14] N. J. Bershad, "On error saturation nonlinearities for LMS adaptation in impulse noise," *IEEE Trans*. *Signal Processing*, vol.56, no. 9, pp.4526-4529, Sept. 2008.

[15] N. J. Bershad, "Analysis of the normalized LMS algorithm with Gaussian inputs," *IEEE Trans*. *Acoust*., *Speech and Signal Processing*, vol. 34, no. 8, pp. 793-806, Aug. 1986.

[16] R. H. Kwong and E. W. Johnston, "A variable step size LMS algorithm," *IEEE Trans*. *Signal Processing*, vol. 40, no. 7, pp. 1633-1642, Jul. 1992.

[17] T. Aboulnasr and K. Mayyas, "A robust variable step-size LMS-type algorithm," Analysis and simulations," *IEEE Trans*. *Signal Processing*, vol. 45, no. 3, pp. 631-639, Mar. 1997.

[18] V. J. Mathews and Z. Xie, "A stochastic gradient adaptive filter with gradient adaptive step size," *IEEE Trans*. *Signal Processing*, vol. 41, no. 6, pp. 2075-2087, June 1993.

[19] S. Koike, "A class of adaptive step-size control algorithms for adaptive filters," *IEEE Trans*. *Signal Processing*, vol. 50, no. 6, pp. 1315-1326, June 2002.

[20] Hsu-Chang Huang and Junghsi Lee, "A new variable step-size NLMS algorithm and its performance analysis," *IEEE Trans*. *Signal Processing*, vol. 60, no. 4, pp. 2055-2060, Apr. 2012.

[21] S. Koike, "Performance analysis of recursive least moduli algorithm for fast convergent and robust adaptive filters," in *Proc*. *IEEE ICASSP 2009*, Apr. 2009, pp. 3133-3136.

[22] K. Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *Electronics and Communications in Japan*, vol. 67-A, no. 5, 1984.