

Variable Forgetting Factor Algorithm Based on Sparsity for RLS-type Adaptive Algorithms

Kiyoshi Nishikawa and Junya Hasegawa

Dept. of information and Communications Systems, Tokyo Metropolitan University

Hino 191-0065 Tokyo Japan

E-mail: knishikawa@m.ieice.org Tel: +81-42-585-8423

Abstract—RLS-type adaptive algorithms employ the forgetting factor for improving the tracking performance. Under fixed forgetting factor configurations, it is known that a smaller value of the factor leads the better tracking but poor steady-state performance. Hence, a variable forgetting factor algorithm is desired to achieve better performance in both of stationary and non-stationary environments. In this paper, we propose a novel variable forgetting factor algorithm based on the sparsity of the adaptive filter instead of error and noise variances in the conventional methods. We propose a formula to adjust the forgetting factor according to the variation of the sparseness of the filter. Through the computer simulations, it is shown that both the tracking and the steady-state performance of the RLS-type adaptive filters could be improved by using the proposed method.

I. INTRODUCTION

In this paper, we propose a novel variable forgetting factor algorithm for RLS-type adaptive filters based on the sparsity. We show that the proposed method provides better tracking performance than conventional methods[8], [9] with maintaining the convergence property in the steady state.

The RLS (recursive least squares) is one of the typical adaptive algorithms and it is well known that its convergence speed is faster than other algorithms, such as LMS (least mean square)[2], by paying the cost of computational complexity. However, its performance would be degraded under non-stationary environments, namely it would be equivalent to that of NLMS (normalized LMS) algorithm under some conditions [3]. The ability to track the non-stationary environments is known as tracking performance. The purpose of the proposed method is to improve the tracking performance of the RLS-type algorithms.

The ERLS-DCD (exponentially weighted RLS - dichotomous coordinate descent) algorithm is proposed in order to decrease computational complexity of the RLS [4]. The complexity required for the ERLS-DCD algorithm is approximately same as that of the NLMS. However, its tracking performance becomes worse than or equal to that of the RLS.

For better tracking performance, several algorithms are proposed so far. For the LMS-type adaptive filters, they are known as the variable step size algorithms[2] and, for the RLS algorithm, variable forgetting factor algorithms[8], [9]. These algorithms adjust the value of the step-size or the forgetting factor according to the variance of the error signal. Namely, when the variance increases they consider environments are

varied and adjust the values of the parameters to track the environments.

In applications such as echo cancelers adaptive filters are frequently required to estimate sparse systems [5]. The algorithms specialized for sparse systems are proposed, i.e. PNLMS (proportionate normalized least mean square) algorithm and its modified versions[5], [6], [7]. The PNLMS algorithm improves the NLMS algorithm by computing step sizes for each tap in response to previously estimated tap coefficients. The PNLMS algorithm, however, is one of gradient-type algorithms, so its convergence property for correlated signals would be degraded.

In this paper, we propose a novel variable step-size method for RLS-type algorithms for sparse systems identification. The proposed method uses the sparseness[1] of the filter to estimate the variation of the environments instead of the error and noise variances in the conventional methods. We derive the formula for updating the value of the forgetting factor according to the sparseness. We show that the proposed method could improve the tracking performance in non-stationary environments and the steady-state performance in stationary environments by varying value of the forgetting factor.

II. PREPARATION

Here, we review adaptive algorithms which are used in this paper and the sparsity of a system shortly.

A. RLS algorithm[2]

The RLS algorithm minimizes the weighted sum of the squared errors by solving the normal equations with matrix inversion lemma [2]. Fig.1 shows a typical structure of the RLS adaptive filter, where $x(n)$, $d(n)$ and $v(n)$ are input signal, desired signal and the additive noise at time n , respectively.

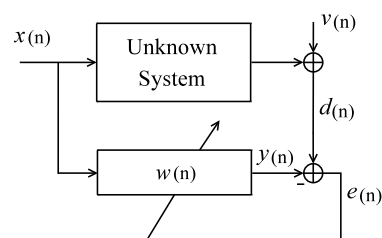


Fig. 1. Structure of RLS adaptive filter

We show the i -th coefficient of the adaptive filter as $w_i(n)$ and the output of the filter as $y(n)$. The error signal $e(n)$ is defined as $e(n) = d(n) - y(n)$. We define the filter input vector $\mathbf{x}(n)$ and filter coefficients vector $\mathbf{w}(n)$ as

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

$$\mathbf{w}(n) = [w_0(n), w_1(n), \dots, w_{M-1}(n)]^T, \quad (2)$$

respectively, where M shows the length of $w(n)$ and T denotes transposition of a matrix. The RLS algorithm updates $\mathbf{w}(n)$ as follows:

$$e(n) = d(n) - \mathbf{x}^T(n)\mathbf{w}(n-1) \quad (3)$$

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mathbf{K}(n)e(n) \quad (4)$$

$$\mathbf{K}(n) = \frac{\mathbf{P}(n-1)\mathbf{x}(n)}{\lambda_{RLS} + \mathbf{x}^T(n)\mathbf{P}(n-1)\mathbf{x}} \quad (5)$$

$$\mathbf{P}(n) = \lambda_{RLS}^{-1} (\mathbf{P}(n-1) - \mathbf{K}(n)\mathbf{x}^T(n)\mathbf{P}(n-1)) \quad (6)$$

where $\mathbf{K}(n)$, $\mathbf{P}(n)$ and λ_{RLS} are the Kalman gain vector, the inverse of auto-correlation matrix, and the forgetting factor, respectively. The forgetting factor is used to control the tracking performance and its value is in the range $0 < \lambda_{RLS} \leq 1$. It enables the filter to track time varying systems by weighting old data. Although the tracking performance can be improved by decreasing λ_{RLS} , it increases excess MSE at the same time. Besides, the RLS algorithm requires a complexity of $O(M^2)$ for updating the filter coefficients to achieve the faster convergence speed than LMS-type ones.

B. ERLS-DCD algorithm[4]

The ERLS-DCD algorithm is recently proposed for reducing the computational complexity of the RLS algorithm. Note that in the following, we use the letter \mathbf{h} to indicate the filter coefficient vector of the ERLS-DCD algorithm to avoid ambiguity with that of the RLS. Although the RLS algorithm requires order $O(M^2)$ calculations for updating the filter coefficients, the ERLS-DCD can be implemented in order $O(M)$.

The ERLS-DCD algorithm transforms the normal equations into the auxiliary normal equations below

$$\mathbf{R}(n)\Delta\mathbf{h}_o(n) = \beta_0(n), \quad \Delta\mathbf{h}_o(n) = \mathbf{h}_o(n) - \mathbf{h}_o(n-1) \quad (7)$$

where \mathbf{h}_o shows the optimum filter and updates the filter coefficients $\mathbf{h}(n)$, as follows:

$$\mathbf{R}(n) = \lambda_{ERLS}\mathbf{R}(n-1) + \mathbf{x}(n)\mathbf{x}^T(n) \quad (8)$$

$$\beta_0(n) = \lambda_{ERLS}\mathbf{r}(n-1) + e(n)\mathbf{x}(n) \quad (9)$$

$$\mathbf{h}(n) = \mathbf{h}(n-1) + \Delta\mathbf{h}(n), \quad (10)$$

where $\mathbf{R}(n)$ is a positive definite symmetric matrix; and $\mathbf{x}(n)$, $e(n)$ and $\mathbf{h}(n)$ are the filter input vector, the error signal and the filter coefficients vector, respectively. λ_{ERLS} is the forgetting factor ($0 < \lambda_{ERLS} \leq 1$). At each time, $\Delta\mathbf{h}(n)$ and $\mathbf{r}(n)$ are obtained by solving Eq.(9) using DCD algorithm which can be implemented without a multiplication[4]. Using a symmetric property of $\mathbf{R}(n)$, the algorithm can be implemented by the complexity equivalent to that of the NLMS.

C. Conventional Variable Forgetting Factor Algorithms

As we mentioned in Introduction, there are several variable forgetting factor algorithms proposed so far, e.g., [8], [9]. The update formula of the forgetting factor in [8] are given as Eq.(11), and [9] as Eq.(12):

$$\lambda(n) = \left[\lambda(n-1) - \frac{\mu}{1 - \lambda(n-1)} \cdot \frac{\partial \hat{\sigma}_e^2(n)}{\partial \lambda} \right]_{\lambda_{\min}}^{\lambda_{\max}} \quad (11)$$

$$\lambda(n) = \min \left\{ \frac{\hat{\sigma}_q(n)\hat{\sigma}_v(n)}{|\hat{\sigma}_e(n) - \hat{\sigma}_v(n)|}, \lambda_{\max} \right\}. \quad (12)$$

In these equations, $\hat{\sigma}_e(n)$, $\hat{\sigma}_v(n)$ are the estimated variance of the error and noise signals, and $\hat{\sigma}_q(n)$ is the variance of $q(n) = \mathbf{x}(n)\mathbf{P}(n-1)\mathbf{x}(n)$ of the RLS algorithm. Known from these equation, the conventional methods require the precise estimation of the variances of $e(n)$ and $v(n)$.

D. Sparsity of a system

The proposed method employs the sparsity, instead of variances of error and noise signals employed in [8], [9], that is used in several algorithms (e.g., PNLMS algorithm), as the index of sparseness of systems. The sparsity $\xi(\mathbf{w})$ of \mathbf{w} is defined as [1]

$$\xi(\mathbf{w}) = \frac{L}{L - \sqrt{L}} \left(1 - \frac{\|\mathbf{w}\|_1}{\sqrt{L}\|\mathbf{w}\|_2} \right), \quad (13)$$

where L is the length of impulse response of a system, $\|\mathbf{w}\|_1$ and $\|\mathbf{w}\|_2$ are the L_1 -norm and the L_2 -norm expressed as

$$\|\mathbf{h}\|_1 = \sum_{n=0}^{L-1} |w_n|, \quad \|\mathbf{h}\|_2 = \sqrt{\sum_{n=0}^{L-1} w_n^2} \quad (14)$$

respectively. We can confirm that when \mathbf{w} is a perfectly sparse system (e.g. Dirac delta function), the value of $\xi(\mathbf{w})$ becomes 1. On the other hand, as the sparsity of \mathbf{w} decreased, $\xi(\mathbf{w})$ approaches 0.

III. PROPOSED METHOD

Here, we describe the proposed method. We show that we can estimate the variation of an unknown system using Eq.(13) and, based on this expression, a novel variable forgetting factor algorithm is proposed.

A. Detection of variation based on sparsity

The proposed method focuses on the relation between the variation of an unknown system and that of the sparsity of filter \mathbf{w} . We propose to evaluate the variation of an unknown system in terms of $\xi(\mathbf{w})$.

If the unknown system is a sparse one, the adaptive filter also becomes sparse as it converges and so that the value of $\xi(\mathbf{w})$ increases. On the other hand, when the unknown system varies, the error signal increases and, in this case, values of $\xi(\mathbf{w})$ decreases due to those error components.

Fig.2 is an example of the variation of the sparsity. The figure shows the simulation results of the case where the unknown system is changed at time 6000, the length of the unknown and adaptive filters were 1000. We can confirm the value of sparsity has been decreased immediately at time 6000.

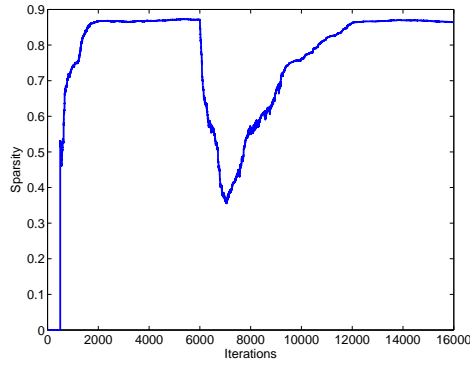


Fig. 2. Values of sparsity of filter coefficients when the unknown system changed at time 6000

B. Sparsity-based variable forgetting factor RLS algorithm (SVFF-RLS)

Here, we describe the outline of the proposed variable forgetting factor method based on the sparsity. Note that the detail of the derivation is omitted because of limited space.

In the propose method, the value of λ_{RLS} is adjusted according to the value of $\xi(\mathbf{w})$ as follows:

$$\lambda_{RLS}(n) = \frac{1 - \lambda_{min}}{1 - \exp[-a\{\hat{\xi}(n) + b(n)\}]} + \lambda_{min}, \quad (15)$$

$$\xi(n) = 20\{\xi(\mathbf{w}(n)) - 0.5\}. \quad (16)$$

where λ_{min} and a show the lower bound of the forgetting factor and a constant gain respectively; $b(n)$ is a parameter whose value are updated according to the system variation and we describe the update recursion in the following; and $\hat{\xi}(n)$ is an expanded function of the sparsity $\xi(\mathbf{w}(n))$. In this paper, we propose to update the parameter $b(n)$ according to the recursion in Table I. Note that the derivation of these recursion is omitted because of the limitation of the space.

Finally, we use the variable forgetting factor defined in Eq.(15) into Eqs (5) and (6). We refer the RLS algorithm based on the proposed method as the sparsity-based variable forgetting factor RLS (SVFF-RLS).

C. Application to the ERLS-DCD algorithm (SVFF-ERLS-DCD)

The proposed method can be applied to other RLS-type algorithms, such as ERLS-DCD. Namely, we propose to update the value of the forgetting factor of ERLS-DCD as

$$\lambda_{ERLS}(n) = \frac{1 - \lambda_{min}}{1 - \exp[-a\{\hat{\xi}(n) + b(n)\}]} + \lambda_{min}. \quad (17)$$

We call the ERLS-DCD based on the proposed method the SVFF-ERLS-DCD (sparsity-based variable forgetting factor - ERLS-DCD).

IV. SIMULATION

Here, we provide results of the system identification using the proposed method to confirm its effectiveness.

TABLE I
UPDATE RECURSION OF THE PARAMETER $b(n)$. λ_{MAX} AND λ_{MIN} ARE THE MAXIMUM AND THE MINIMUM VALUES OF λ

Initial values:
$b(0) = 0, b^-(0) = -\Delta b, b^+(0) = \Delta b$
$\Delta b = a^{-1} \log_e (c/(\lambda_{max} - \lambda_{min}) - 1)$
At each time n :
if $(\xi_w(n) < b^-(n-1))$:
$b^-(n) = \xi_w(n)$
$b(n) = b^-(n) + \Delta b$
$b^+(n) = b(n) + \Delta b$
elseif $(b^-(n-1) \leq \xi_w(n) < b^+(n-1))$:
$b(n) = b(n-1)$
$b^-(n) = b(n) - \Delta b$
$b^+(n) = b(n) + \Delta b$
elseif $(b^+(n-1) \leq \xi_w(n))$:
$b^+(n) = \xi_w(n)$
$b(n) = b^+(n) - \Delta b$
$b^-(n) = b(n) - \Delta b$

A. Performance of SVFF-RLS algorithm

First, let us show the results of the simulations using the SVFF-RLS algorithm to confirm the effect of the proposed method. We used an FIR filter whose sparsity was 0.8956 as the unknown system. The length of the unknown system and the adaptive filter were 500 taps. The parameters for the proposed method are $a = 10$, $\lambda_{max} = 0.9999$, and $\lambda_{min} = 0.97$. We simulated that the unknown system changed drastically at time $n = 12000$. The input signal was a speech signal sampled at 8 kHz and a white Gaussian noise of zero-mean was added to the desired signal (SNR = 40[dB]). In this simulation, we compared the standard RLS algorithm with the forgetting factor of $\lambda_R = 0.995$ and $\lambda_R = 0.9999$ and the results were ensemble averages of 10 independent trials.

Fig.3 shows the results and we can confirm that the proposed method achieves initial convergence speed which is equivalent to the RLS with $\lambda_R = 0.9999$ and good tracking performance which is equivalent to the RLS of $\lambda_R = 0.995$.

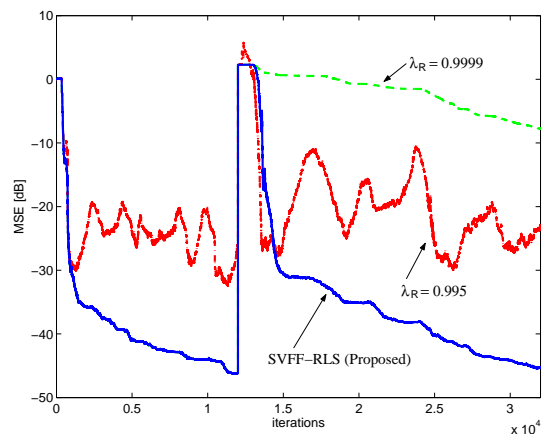


Fig. 3. Comparison of the convergence properties of the proposed SVFF-RLS and the standard RLS algorithms.

B. Comparison with conventional algorithms

Next, we compared the performance of the proposed method with the conventional variable forgetting factor algorithms, namely GVFF-RLS[8], and VFF-RLS[9]. In this simulation, the length of unknown and adaptive filters were 64 because the selection of parameters of VFF-RLS was sensitive to the order of the filter. The parameters for the proposed method was $\alpha = 10$, $\lambda_{\max} = 0.9999$, and $\lambda_{\min} = 0.85$. We applied the algorithms to 20 independent speech signals and the results were averaged. The performance comparison is shown in Fig.4. From the figure, we could see that the proposed method provides better performance than the conventional ones.

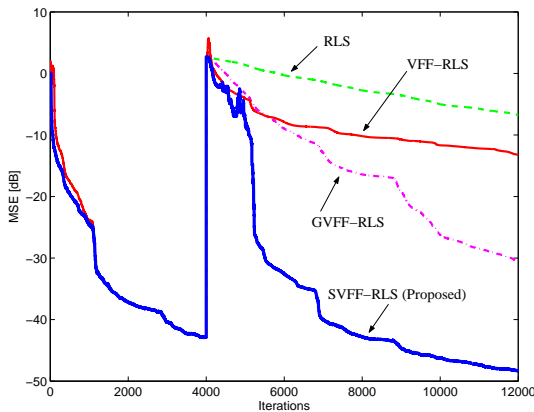


Fig. 4. Comparison of the convergence properties of the proposed, GVFF-RLS, VFF-RLS, ERLS-DCD algorithms. (Average of the results for 20 independent speech signals)

C. SVFF-ERLS-DCD algorithm

Here, we confirm that the proposed method can be applied to the ERLS-DCD algorithm. We compared the proposed method with GVFF-RLS, and VFF-RLS. The values for the parameters were identical to those in IV-B.

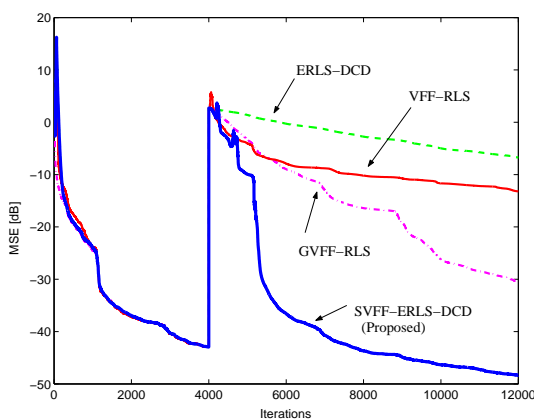


Fig. 5. Comparison of the convergence properties of the proposed (SVFF-ERLS-DCD), GVFF-RLS, VFF-RLS, ERLS-DCD algorithms. (Average of the results for 20 independent speech signals)

D. Comparison with PNLMS algorithm

Finally, we compared the tracking performances of VFF-RLS and VFF-ERLS-DCD with that of the PNLMS algorithm with its step size parameter set as 1. Fig.6 shows simulation results. Both proposed algorithms outperform the PNLMS algorithm in terms of the steady-state and the tracking performances simultaneously.

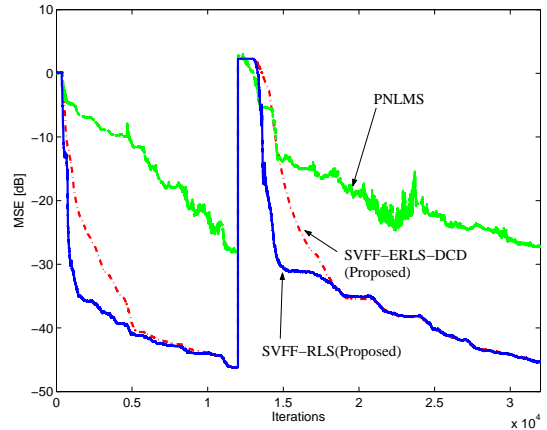


Fig. 6. Comparison with PNLMS algorithm

V. CONCLUSION

In this paper, we proposed the sparsity-based variable forgetting factor algorithm for RLS-type algorithms. The forgetting factor in the proposed method is adjusted based on the sparsity of the filter instead of the error variance in the conventional methods. We confirmed that the proposed method could provide a better performance than the conventional methods by computational simulations.

REFERENCES

- [1] Y.Huang, J.Benesty, and J.Chen "Acoustic MIMO Signal Processing (Signal and Communication Technology)." Secaucus, NJ, USA:Springer-Verlag New York, Inc., 2006.
- [2] A. H. Sayed, Fundamentals of Adaptive Filtering. Hoboken, NJ: Wiley, 2003.
- [3] Eweda Eweda, "Comparison of RLS, LMS, and Sign Algorithms for Tracking Randomly Time-Varying Channels" IEEE Trans. on Signal Process., vol. 42, no.11, pp2937-2944, November 1994.
- [4] Zakharov et al., "Low-Complexity RLS Algorithms Using Dichotomous Coordinate Descent Iterations" IEEE Trans. Signal Process.,VOL. 56, NO.7, July 2008.
- [5] Ligang Liu et al., "Proportionate Adaptive Algorithm Guided by Estimation of Channel Sparsity" IEICE SIP 2007., pp587-591, November 2007.
- [6] Sachio Saiki, et al., "Convergence Characteristics of Variable Step Size IPNLMS Algorithm" IEICE SIP 2007., pp592-597, November 2007.
- [7] Pradeep Loganathan, et al., "A Class of SparsenessControlled Algorithms for Echo Cancellation" IEEE Trans. Audio Speech Language Process., vol. 17, pp1591-1601, Nov 2009.
- [8] Shu-Hung Leung, et al., "Gradient-Based Variable Forgetting Factor RLS Algorithm in Time-Varying Environments" IEEE Trans. Signal Process., vol.53, pp.3141-3150, Aug 2005.
- [9] Constantin Paleologu, et al., "A Robust Variable Forgetting Factor Recursive Least-Squares Algorithm for System Identification" IEEE Signal Process. Letters, VOL.15, pp.597-600, 2008.