An Alternating Selection for Parallel Affine Projection Filters

Kwang-Hoon KIM\textsuperscript{a)}, Seong-Eun KIM\textsuperscript{b)}, Nonmembers, and Woo-Jin SONG\textsuperscript{a),c), Member

SUMMARY We present a new structure for parallel affine projection (AP) filters with different step-sizes. By observing their error signals, the proposed alternating AP (A-AP) filter selects one of the two AP filters and updates the weights of the selected filter for each iteration. As a result, the total computations required for the proposed structure are almost the same as that for a single AP filter. Experimental results show that the proposed alternating selection scheme extracts the best properties of each component filter, namely fast convergence and small steady-state error.

key words: adaptive filters, affine projection (AP), alternating selection

1. Introduction

Like LMS-type algorithms \cite{1}, \cite{2} the affine projection (AP) algorithm \cite{3} involves a tradeoff between convergence speed and accuracy depending on the step-size parameter. Hence, a number of variable step-size techniques have been proposed to address this tradeoff \cite{4}–\cite{6}. Recently, new approaches having different philosophy for solving the tradeoff have been devised \cite{7}–\cite{9}. As is well known, this scheme is easy to implement and can be simply extended to other kinds of adaptive filters without major modifications. Also, because of the robustness to the tuning parameters, this scheme is widely used for many applications. In spite of these advantages, it obviously requires twice as much computational burden as its individual filter. Furthermore, when extending this scheme to combining AP filters, the overall computation level becomes very high. In the sequel, combining the AP filters has not received as much attention to date.

In this paper, we propose a new scheme that is efficient to use in real-world applications. The proposed scheme, called the alternating AP (A-AP) filter, combines two AP filters with different step-sizes to exploit the advantages of both component filters. Consequently, the proposed A-AP chooses one of the two AP filters by considering the status of the combined filter and updates the weight vector of the corresponding filter.

In proposing this new scheme, we classify the adaptation into two modes: the tracking mode is the period when the adaptive filter converges quickly, and the steady-state mode is the period when it settles into the low steady-state error. The mode is determined by comparing the output error with a threshold derived from the steady-state mean-squared error (MSE) \cite{10}. By using this information, the proposed A-AP selects the more appropriate one between the two AP filters as shown in Fig. 1. In the tracking mode, the fast filter is to be selected to guarantee quick convergence. Subsequently, in the steady-state mode, the slow filter is to be selected, resulting in low steady-state error. Therefore, the proposed A-AP requires almost the same amount of computation as a single AP filter without degrading convergence performance.

2. Alternatively Selecting the AP Filter

Consider data \{d(i)\} that originate from the system identification model

\begin{equation}
    d(i) = u_i \mathbf{w}^o + v(i),
\end{equation}

where \( \mathbf{w}^o \) is the optimum column vector of length \( M \) of an unknown system that we wish to estimate, \( v(i) \) accounts for the measurement noise which is assumed to be white with a variance \( \sigma_v^2 \), and \( u_i \) denotes row input vectors of length \( M \) as follows:

\begin{equation}
    u_i = [u(i) \ u(i-1) \ \cdots \ u(i-M+1)].
\end{equation}

where \( u_i \) and \( d(i) \) can be expanded in the form of a matrix:
where $v$ is an advantage of the faster convergence of the fast filter to speed of each component filter. To assign the one. Thus, the modified adaptation rule for only applied when the fast filter is much better than the slow filter is selected to ensure low steady-state error. Accordingly, we propose an efficient method which chooses one of the two AP filters by considering the status of the combined filter and updates its weight vector dur- ing each iteration. In the tracking mode, the fast filter is selected to ensure quick convergence. In the steady-state mode, the slow filter is selected to ensure low steady-state error. With this motivation in mind, we develop a new combination scheme which updates only one AP filter as re- quired for each adaptation mode. The proposed scheme se- lects one of the two AP filters at iteration $i$, i.e., $y(i) = w_i(i)$, $k = 1, 2$, with $w_{k,i}$ being the adaptive filter coefficient vectors of the $k$-th filter, and $\lambda(\hat{i}) \in [0, 1]$ is the balancing parameter. The $\lambda(\hat{i})$ is updated in order to minimize the squared error of the overall filter. The update of $\lambda(i)$ at each iteration tries to extract best properties of each component filter. To assign the $\lambda(i)$ properly, for example, we borrow the following adaptation rule [7], [8]:

$$\lambda(i) = \text{sgm}[a(i)] = (1 + e^{-a(i)})^{-1},$$

and

$$a(i + 1) = a(i) + \mu e(i) [y(i) - y(i)] \cdot \lambda(i)[1 - \lambda(i)] + \rho[a(i) - a(i - 1)],$$

where $e(i) = d(i) - u_i w_{k,i}$, $\mu$ is the step-size of the adapta- tion of $a(i)$, and $\rho$ is a positive constant. The performance of the combined filter can be further improved if we take ad- vantage of the faster convergence of the fast filter to speed up the convergence of the slow one when an abrupt change appears. This can be done by step by step transferring a por- tion of weights $w_1$ to $w_2$. The weight transfer procedure is only applied when the fast filter is much better than the slow one. Thus, the modified adaptation rule for $w_2$ is obtained as follows [8]:

$$w_{2,i} = \alpha w_{2,i} + (1 - \alpha) w_{1,i}, \quad \text{for } \lambda(i) > 1$$

where $\alpha$ is a parameter close to 1. In general, $\alpha = 0.9$ and $t = 0.98$ are good choices, and these are the settings we have used in all the simulations.

2.2 Alternately Selecting the AP Filter

From (4)–(8), we extract the best properties of each com- ponent filter by means of the two AP filters with fixed step- sizes. However, the computational demand grows in direct proportion to the number of filters. Therefore, it is not suit- able to use in practical applications without modification.

We see from (5) that before the combined filter con- verges, $\lambda(i) \approx 1$; thus, $y(i) \approx y_1(i)$, the combined filter acts as a fast filter. On the other hand, when it comes to the steady-state, $\lambda(i) \approx 0$ and $y_2(i)$, the combined filter behaves as a slow filter. Based on this observation, we can clearly establish that updating the weights of both the fast and slow filters simultaneously at every iteration is inefficient. Furthermore, for combining the two AP filters such as (4)–(8), the overall computations are very high since the AP filter requires much more computations than LMS-type filters.

Accordingly, we propose an efficient method which choices one of the two AP filters by considering the status of the combined filter and updates its weight vector dur- ing each iteration. In the tracking mode, the fast filter is selected to ensure quick convergence. In the steady-state mode, the slow filter is selected to ensure low steady-state error. With this motivation in mind, we develop a new combination scheme which updates only one AP filter as re- quired for each adaptation mode. The proposed scheme se- lects one of the two AP filters at iteration $i$ by compar- ing the squared output error with the threshold. Therefore, the update equation of each AP filter can be formalized as follows:

$$w_{1,i} = \begin{cases} w_{1,i-1} + \mu_1 U_i (e I + U_i U_i)^{-1} e_{1,i}, & \text{if } e^2(i) \geq \delta \\ w_{1,i-1}, & \text{otherwise} \end{cases}$$

$$w_{2,i} = \begin{cases} w_{2,i-1} + \mu_2 U_i (e I + U_i U_i)^{-1} e_{2,i}, & \text{if } e^2(i) \geq \delta \\ w_{2,i-1}, & \text{otherwise} \end{cases}$$

where $e_{k,i} = d(i) - u_i w_{k,i-1}$, $k = 1, 2$ and $\delta$ denotes the threshold. If $e^2(i)$ is larger than $\delta$, then the fast filter is selected to ensure a quick convergence. As the slow filter is not selected, the weight coefficient of the slow filter is main- tained as the previous value. In the opposite case, the slow filter is selected to ensure low steady-state error. Accord- ingly, the fast filter is not selected and the weight coefficient of the slow filter is maintained as the previous value. If the fast filter is significantly outperforming the slow one, the weight transfer procedure is applied, as in (8). Based on this
line of thought, the proposed A-AP reduces the heavy computational burden while maintaining its convergence performance.

2.3 Convergence Behavior of the Proposed A-AP

The convergence behavior of the proposed A-AP is closely related to the evolution of the \( \lambda(i) \). As mentioned in previous subsection, the \( \lambda(i) \) is evolved to minimize the quadratic error of the combined filter. When the error signal \( e(i) \) is very large (for instance, when there is an abrupt change), the combined filter should behave as a fast AP filter for quick convergence. When the error signal \( e(i) \) gets much larger for the purpose of minimizing the steady-state performance. But once it has reached the “steady-state mode”, thus decreasing \( a(i) \) so that \( \lambda(i) \) approaches 0, the proposed scheme behaves as the slow AP filter. As a result, we can obtain the small residual error during the steady-state.

2.4 Determination of the Threshold

As suggested by (9a) and (9b), the alternating selection method is accomplished by comparing the squared output error with the threshold, \( \delta \). Therefore, the threshold must reflect the stage of the AP filter. As is well known, the AP algorithm converges with the following steady-state MSE in a stationary system [10], assuming that the noise is statistically independent of the regression matrix \( \{ U_i \} \) and \( \{ U_i \} \) is statistically independent of \( e_i \) at a steady-state:

\[
MSE = \frac{\mu \sigma^2}{2 - \mu} \text{Tr}(R_u) E \left[ \frac{K}{||u||^2} \right] + \sigma_v^2, \tag{11}
\]

where \( R_u = E[u u^T] \) and Tr(\( \cdot \)) denotes the trace of the matrix. Since it is not feasible to obtain the exact value in (11), we replace it with an instantaneous value. For a large \( M \), the fluctuations in the input signal energy \( ||u||^2 \) from one iteration to the next are small enough to justify the approximation below [1],[15].

\[
E \left[ \frac{1}{||u||^2} \right] \approx \frac{1}{E[||u||^2]} \tag{12}
\]

By applying (12) to (11), and using the relation \( \text{Tr}(R_u) = E[||u||^2] \), Eq. (11) can be rewritten as

\[
MSE \approx \frac{\mu \sigma^2 K}{2 - \mu} + \sigma_v^2. \tag{13}
\]

When the squared output error reaches the MSE of the fast filter, the status of the combined filter can be regarded as steady-state. For this reason, we set the threshold \( \delta \) by substituting \( \mu_1 \) into (13). We can now obtain the threshold \( \delta \) as follows:

\[
\delta = \frac{\mu_1 \sigma^2 K}{2 - \mu_1} + \sigma_v^2. \tag{14}
\]

In practice, the noise variance, \( \sigma_v^2 \), can be easily estimated during silences [11],[12] and online [13],[14]. Table 1 summarizes the proposed A-AP.

3. Experimental Results

The performance of the proposed scheme was demonstrated...
by carrying out computer experiments on the system identification configuration. The length of the unknown system is $M = 16$. The unknown system was randomly generated, and the adaptive filter and the unknown system are assumed to have the same number of taps. The input signal $u(i)$ is obtained by filtering a white, zero-mean, Gaussian random sequence through a first-order system $G(z) = 1/(1 - 0.9z^{-1})$. The signal-to-noise ratio (SNR) is calculated by $\text{SNR} = 10\log_{10}(E[y^2(i)]/E[v^2(i)])$, where $y(i) = u_i w_0$. The measurement noise $v(i)$ was added to $y(i)$ with SNR=30 dB.

Figure 3 shows the plot of the MSD curves for the AP filters ($\mu = 1.0, 0.1, 0.02$, and 0.002), VS-AP [6], and the proposed A-AP. For comparison purposes, we have chosen the VS-AP ($K = 8, M = 16, \mu_{\text{max}} = 1.0, \alpha = 0.99$, and $C = 0.15$) as a good representative of variable step-size AP algorithms because it has generally obtained the best performance. The step-size corresponding to the fast AP filter is $\mu = 1.0$, which is selected to provide the fastest possible convergence rate, whereas the step-size of the slow AP filter is selected as $\mu = 0.002$ to obtain the similar steady-state MSE attained by the VS-AP. As we can see, the proposed scheme retains the rapid convergence from the fast filter and the small steady-state error from the slow filter. In Fig. 3, we can see that the proposed A-AP has lower MSD than AP filter with $\mu = 0.002$, since the proposed scheme updates the weights of the slower AP filter only when the error is small. Note also that the proposed A-AP results in a similar convergence performance to the VS-AP.

Figure 4 shows the result of suddenly multiplying the unknown system by $-1$. It can be seen that the proposed A-AP keeps track of the sudden weight change without degrading the convergence speed or the steady-state accuracy.

Averaged selection indicator for over 100 independent trials. The unknown system is suddenly changed from $w_0$ to $-w_0$. All of the parameter values are the same as in Fig. 3.
Fig. 6 MSD curves of the proposed A-AP with various estimated noise variances when the exact noise variance is $\sigma^2$. All of the parameter values are the same as in Fig. 3 except for the noise variance.

4. Conclusions

We have presented an alternating selection scheme for parallel AP filters. For each adaptation stage, the selection of the filter is accomplished by comparing the squared output error with the threshold which reflects the status of the adaptive filter. As a consequence, the proposed framework adapts either one or the other AP filter depending on the level of the steady-state MSE. Through experimental results, we have shown that the proposed A-AP requires almost the same amount of computation as a single AP filter without degrading the convergence performance in terms of speed and precision. Furthermore, since the proposed A-AP is based on the combination scheme, the parameters of the same value can be used for most applications. Therefore, the proposed A-AP can be more popular in the area of adaptive signal processing than the VS-AP.

Acknowledgments

This work was supported in part by the Brain Korea (BK) 21 Program funded by the MEST, in part by HY-SDR Research Center at Hanyang University under the ITRC Program of MKE and in part by the IT R&D program of MCST/ITIA(2008-F-031-01), Korea.

References