<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>A Huber recursive least squares adaptive lattice filter for impulse noise suppression</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Zou, Y; Chan, SC</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>Icassp, Ieee International Conference On Acoustics, Speech And Signal Processing - Proceedings, 2001, v. 6, p. 3769-3772</td>
</tr>
<tr>
<td><strong>Issued Date</strong></td>
<td>2001</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10722/46257">http://hdl.handle.net/10722/46257</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>Creative Commons: Attribution 3.0 Hong Kong License</td>
</tr>
</tbody>
</table>
This paper proposes a new adaptive filtering algorithm called the Huber Prior Error-Feedback Least Squares Lattice (H-PEF-LSL) algorithm for robust adaptive filtering in impulse noise environment. It minimizes a modified Huber M-estimator based cost function, instead of the least squares cost function. In addition, the simple modified Huber M-estimate cost function also allows us to perform the time and order recursive updates of the conventional PEF-LSL algorithm so that the complexity can be significantly reduced to \( O(M) \), where \( M \) is the length of the adaptive filter. The new algorithm can also be viewed as an efficient implementation of the recursive least M-estimate (RLM) algorithm recently proposed by the authors [1], which has a complexity of \( O(M^2) \). Simulation results show that the proposed H-PEF-LSL algorithm is more robust than the conventional PEF-LSL algorithm in suppressing the adverse influence of the impulses at the input and desired signals with small additional computational cost.

1. INTRODUCTION

Impulsive interference, which results from nature or man-made electromagnetic waves, can significantly degrade the performance of linear adaptive filters. Nonlinear techniques are usually employed to suppress such adverse effects. For example, the median filtering has been applied to the LMS and the RLS algorithms to protect the filter weights from the effects of impulsive interference, giving rise to the order statistic least mean square (OSLMS) [2] and the order statistic fast Kalman filtering (OSFKF) algorithms [3]. While the adaptive threshold nonlinear (ATNA) [4] algorithm uses the clipping function to limit the transient fluctuation of the estimation error on filter coefficients in the LMS algorithm. Recently, the authors have proposed a new family of adaptive filters for robust adaptive filtering in impulse noise environment based on the concept of robust statistics [1, 5, 6]. Instead of minimizing the conventional mean squares error (MSE) or the least squares (LS) cost functions, the robust M-estimators based cost functions, which are more robust to impulse noise, were minimized. In particular, a recursive least M-estimate (RLM) [1], a least mean M-estimate (LMM) [5] and a robust gradient adaptive lattice - normalized LMS (RGAL-NLMS) [6] algorithms were developed to suppress the impulsive interference in the input or the desired signals. Simulation results showed that the RLM, LMM and RGAL-NLMS algorithms offer improved robustness to impulses in the desired and input signals over the conventional recursive least square (RLS), LMS and GAL-NLMS algorithms, respectively.

Although the recursive least squares (RLS) algorithm provides fast initial convergence rate and low steady state error as compared to the least mean squares (LMS) family [7, 8], research results indicated that it is more sensitive to quantization errors. In fact, it is numerical unstable if implemented in fixed point with less than 24-bit accuracy. Extensive research was devoted to this stability problem and the development of fast RLS algorithm with \( O(M) \) complexity, where \( M \) is the length of the transversal filter. One attractive class of the fast algorithms is the RLS lattice (RLSL) algorithm [7, 8]. They not only provide the exact LS solution, but also possess many distinctive properties such as low computational complexity, modular implementation, and better numerical stability than the conventional RLS algorithm [9]. Therefore, the RLSL algorithms have found many applications in speech signal processing and acoustic echo cancellation (AEC) where good convergence performance, numerical stability, and high computational complexity are the main concerns [9, 10].

In this paper, we generalize the robust statistic approach to the prior error feedback LSL (PEF-LSL) algorithm. In particular, a Huber PEF-LSL (H-PEF-LSL) algorithm is derived by minimizing the modified Huber M-estimate function. Simulation results show that the proposed H-PEF-LSL algorithm offers improved robustness over the conventional PEF-LSL algorithm in suppressing the adverse influence of the impulses both in the input and desired signals with small additional computational cost.

The paper is organized as follows: the formulation of the PEF-LSL algorithm is given in Section 2. The Huber PEF-LSL algorithm is introduced in Section 3. Simulation and comparison with other algorithms are described in Section 4. Finally, conclusions are drawn in Section 5.

2. FORMULATION OF THE PRIOR ERROR FEEDBACK LSL ALGORITHM

The transversal RLS and the PEF-LSL algorithms are mathematically identical. Both of them are developed to obtain the optimal solution of the following exponentially weighted least squares error cost function [8]

\[
J_p(n) = \sum_{i=0}^{\infty} h^2(i),
\]

(1)

where \( d(n) = d(n) - \hat{d}(n-1)X(n) \) is called the prior estimation error, \( X(n) = [x(n), \ldots, x(n-M+1)]^T \), \( d(n) \) and \( \hat{d}(n) \) are the input vector, desired signal and coefficient vector of the transversal filter, respectively. The superscript \( T \), \( M \) and \( 0 < \lambda < 1 \) denote respectively the transpose operator, order of the adaptive filter and the forgetting factor. The optimal LS solution \( \hat{d}(n) \) of (1) is governed by the following normal equation:

\[
R(n)\hat{d}(n) = P(n) = \hat{d}(n) = R^{-1}(n)P(n),
\]

(2)

where \( R(n) = \sum_{i=0}^{\infty} X(i)X^T(i) = \lambda R(n-1) + X(n)X^T(n) \),

\[
P(n) = \sum_{i=0}^{\infty} h^2(i)X(i) = \lambda P(n-1) + d(n)X(n),
\]

(3)

are the autocorrelation matrix of \( X(n) \), and the cross-correlation vector between \( X(n) \) and \( d(n) \). Direct inversion of \( R(n) \) in (2) will require \( O(M^3) \) arithmetic operations. For the transversal RLS algorithm, the arithmetic complexity is reduced to \( O(M^2) \) since the time recursive property of \( R(n) \) in (3) is used. The lattice structure based LSL-type algorithms explore both the time
recursive and order recursive properties of \( R(n) \), and its complexity is further reduced to \( O(M) \).

Fig. 1 shows the structure of an adaptive lattice ladder filter (ALLF). Signals \( x(n) \), \( d(n) \) and \( \bar{d}(n) \) are the input, output and desired signals of the ALLF, respectively, where \( n \) denotes the time index. It is assumed that the noise-free desired signal \( d(n) \) and the input are corrupted by statistically independent additive interference \( \eta_i(n) \) and \( \eta_o(n) \), respectively. The PEF-LSL algorithm is summarized as follows [8]:

**Lattice prediction part \((m = 1, \ldots, M)\):**

\[
\begin{align*}
\bar{x}_m(n) &= \bar{x}_m(n-1) + k_m(n-1)\bar{e}_m(n-1), \\
b_m(n) &= b_m(n-1) + k_m(n-1)\bar{x}_m(n),
\end{align*}
\]

where \( \bar{x}_m(n) = x(n) - \bar{\hat{x}}_m(n) \) and \( b_m(n) = x(n-M) - \bar{\hat{x}}_m(n-M) \) are respectively the \( m \) stage forward and backward prediction errors; \( k_m(n) \) and \( \delta_m(n) \) are respectively the \( m \) stage forward and backward reflection coefficients, which are updated at each iteration as

\[
\begin{align*}
k_m(n) &= k_m(n-1) - \gamma_m(n-1)\bar{e}_m(n-1)/E_m(n-1), \\
k_m(n) &= k_m(n-1) - \gamma_m(n-1)\bar{e}_m(n-1)/E_m(n-1),
\end{align*}
\]

where \( \gamma_m(n-1) = \bar{\gamma}_m(n-1) - \bar{\epsilon}_m(n-1)b_m(n-1)/E_m(n-1) \), \( \gamma_m(n-1) = \bar{\gamma}_m(n-1) - \bar{\epsilon}_m(n-1)b_m(n-1)/E_m(n-1) \), and \( \gamma_m(n-1) = \lambda \gamma_m(n-2) + \gamma_m(n-1) \delta_m(n-1)/E_m(n-1) \) are respectively the convention factor (or the maximum likelihood variable), the forward and backward time averaged prediction error energies.

**LS filtering part \((m = 1, \ldots, M+1)\):**

\[
\begin{align*}
\bar{e}_m(n) &= \bar{\epsilon}_m(n-1) - \bar{\gamma}_m(n-1)\bar{e}_m(n-1), \\
\bar{w}_m(n) &= \bar{\epsilon}_m(n-1) - \bar{\gamma}_m(n-1)\bar{e}_m(n-1)/E_m(n-1),
\end{align*}
\]

where \( \bar{e}_m(n) \) and \( \bar{w}_m(n) \) are respectively the \( m \) stage prior estimation error and the ladder coefficient at time instant \( n \). The PEF-LSL algorithm has the following important properties: (1) the lattice predictors perform the Gram-Schmidt orthogonalization of the input data with very good numerical property [11]; (2) the direct update of the forward and backward coefficients in (6) and (7) also leads to better numerical behavior of this algorithm [8]; (3) the backward prediction errors \( b_m(n) \) in \((m = 1, \ldots, M+1)\) at different stages are uncorrelated and orthogonal to the space spanned by the input vector \( X_m(n) = [x(n), \ldots, x(n-m+1)]^T \).

From the above formulations, it can be seen that if \( x(n) \) and/or \( d(n) \) are corrupted by additive impulsive noise, then \( R(n) \) and/or \( P(n) \) in (3) and hence \( \hat{w}(n) \) in (2) will exhibit momentary fluctuation which might take many iterations to recover, affecting the convergence speed of the RLS adaptive filter [1]. For the PEF-LSL algorithm, since \( f_0(n) = b_0(n) = x(n) \) and \( e_0(n) = d(n) \), the effect of the impulses will propagate through the order and time recursion of the algorithm by disturbing the variables such as reflection coefficients, etc. Thus, it can be expected that the PEF-LSL algorithm will be significantly degraded by the impulses in \( x(n) \) and \( d(n) \).

3. HUBER PEF-LSL (H-PEF-LSL) ALGORITHM

In this section, the proposed robust prior error feedback recursive least squares lattice algorithm will be developed using the robust statistics approaches and the conventional PEF-LSL algorithm [8]. In fact, this work is motivated by our previous work on the robust RLM algorithm [1]. First at all, we shall give a brief introduction to the recursive least M-estimate (RLM) algorithm [1]. Then, the proposed Huber PEF-LSL algorithm will be discussed in detail.

3.1 Recursive Least M-estimate (RLM) Algorithm [1]

In [1], the authors have proposed a new class of adaptive filter based on the concept of robust statistics. Instead of the LS cost function \( J_{LS}(n) \) in (1), the following M-estimator based cost function is minimized.

\[
J_{M}(n) = \sum_{i=1}^{N} \lambda_i \rho(\epsilon(i)),
\]

where \( \rho(\cdot) \) is an M-estimate function. In [1], \( \rho(\cdot) \) is chosen as the Hampel's three parts redescending M-estimate function for its computational simplicity and more flexibility in choosing the interval parameters for impulse noise suppression. The optimal weight vector for this objective function was found to be governed by the following M-estimate normal equation:

\[
R_{M}(n)\hat{w}(n) = P_{M}(n),
\]

where \( R_{M}(n) = \sum_{i=1}^{N} \lambda_i \rho(\epsilon(i))X(i)X(i)^T \), \( P_{M}(n) = \sum_{i=1}^{N} \lambda_i \rho(\epsilon(i))X(i)X(i)^T \), and \( \rho(\cdot) \) is chosen as the modified Huber function, then (12) will be a simplification of the Hampel's three parts redescending M-estimate function used in [1].
signals. Because we shall see later in this section, allows us to utilize again the order and time update of \( R_x(n) \) in (13). First of all, let’s consider that case where \(|e(n)|\) is larger than \( \xi \). This indicates that there might be an impulse in the input or desired signals. Because \( q(e(n)) \) is equal to zero, (12) and (13) will be simplified to

\[
R_x(n) = \lambda R_x(n-1) + P_{y_x}(n-1)
\]

That is, \( R_x(n) \) and \( P_{y_x}(n) \) are not updated but just multiplied by \( \lambda \).

When \(|e(n)|\) is less than or equal to \( \xi \), no impulse is detected. In this case, \( q(e(n)) \) is equal to one and (13) becomes identical to (3). Therefore, the order and time updates in the PEF-LSL algorithm can be performed, significantly reducing the arithmetic complexity. Before proceeding to the detailed implementation, let’s consider the estimation of the threshold parameter \( \xi \) in (15). The solution to this problem has been addressed previously by the authors in [1]. The error signal is modeled as a Gaussian signal corrupted by additive impulse noise. Then, the threshold \( \xi \) in (15) can be chosen as [1]

\[
\xi = k_3 \hat{\sigma}_1(n),
\]

where \( k_3 \) is a constant and \( \hat{\sigma}_1(n) \) is the variance of the error signal without the impulses. Because of the Gaussian assumption, we have 99% confidence that there is an impulse in \( e(n) \) (and hence \( d(n) \) and \( x(n) \)) when \(|e(n)| > \xi \) with \( k_3 = 2.576 \) [1].

Moreover, \( \hat{\sigma}_1(n) \) can be in (16) be estimated using the following formula [1]:

\[
\hat{\sigma}_1(n)^2 = \lambda \hat{\sigma}_1(n-1)^2 + C (1 - \lambda_0) \text{med}(A(n)),
\]

where \( A(n) = (e^n(1), \ldots, e^n(N-n+1)) \), \( N \) is the length of the estimation window, \( \lambda_0 \) is the forgetting factor and \( C = 1.483 \sigma_1^2 / \text{med}(A(n)) \) is a finite sample correction factor ([14], p.44). Due to the recursive nature of the estimation in (17), the estimation window is of infinite length, giving rise to a more stable estimation against impulse noise [1]. The tracking ability of this estimation is also very good, as suggested by computer simulations [1, 5]. Interested readers are referred to [1] for details.

Due to the order recursion, the estimation error \( e(n) \) will be available at the last stage of the ALLF, see Fig. 1, \( e(n) = e_{w(n)}(n) \).

If \(|e(n)| > \xi \), \( d(n) \) or \( x(n) \) are suspected to be corrupted by impulses. In this case, as mentioned earlier, no updating will be performed and the previous filtering parameters should be used instead of those generated in the current iteration. However, when \( x(n) \) is corrupted by an impulse, the case will become more complicated. Although we can also detect this impulse from \( e(n) \), replacing the filtering parameters by their values in the previous iteration cannot effectively suppress the effect of this impulse on the lattice parameters. If we replace both the lattice and the filtering parameters by their values in the past iteration, the exact order and time recursions of the PEF-LSL algorithm will be damaged. Therefore, it is better to suppress the impulse in \( x(n) \) before it enters the ALLF, following the method proposed in [6]. Similar idea has also been proposed by Kim [15], in which the authors suggested a pre-filter to remove the impulses in \( x(n) \).

Following our work in [6], there is no need to add another pre-filter since the lattice prediction part actually implement the prediction of \( e(n) \). As a result, impulse can be detected directly from \( f_y(n) \), and a predictor for \( x(n) \) can be formed as follows

\[
\hat{x}(n) = \sum_{i=0}^{N} b_i(n-1) x_i(n-i).
\]

The corresponding threshold to detect the impulse in \( f_y(n) \) follows the same way as described in (16) and (17). Details of the H-PEF-LSL algorithm are summarized in Table 1. This algorithm can also be view as an effective implementation of the RLM algorithm in impulse noise environment, while preserving the advantages of the lattice-based algorithms.

4. SIMULATIONS

The performance of the proposed H-PEF-LSL algorithm is evaluated and compared with the RLS, the RLM, and the PEF-LSL algorithms in impulse noise environment. The adaptive filter shown in Fig. 1 is used to identify the unknown system \( w' \), which is a 9th order lowpass FIR filter with coefficients \( w' = [2, -4.6, -8.1, -8.6, -4.2] \). To evaluate its tracking ability to sudden system change, \( w' \) is changed to \(-w'\), at \( n = 3000 \). The input signal \( x(n) \) is generated by passing a zero-mean, unit variance Gaussian process through a linear time-invariant filter with coefficients \([3.887, 1, 3.887]\) [8]. The output of the unknown system \( d(n) \) is corrupted by the additive noise \( \eta(n) \), which is modeled as the frequently used contaminated Gaussian noise, \( \eta(n) = \eta_0(n) + b(n) \eta_1(n) \) [1, 2]. In fact, \( \eta_0(n) \) and \( \eta_1(n) \) are independent identically distributed (i.i.d.) zero mean Gaussian noises with variance \( \sigma_{\eta_0}^2 \) and \( \sigma_{\eta_1}^2 \), respectively. \( \eta_0(n) \) is a Bernoulli random variable with occurrence probability \( P_i(b(n)=1) = p \). The ratio \( \gamma_0 = \sigma_{\eta_0}^2 / \sigma_{\eta_1}^2 = p \sigma_{\eta_1}^2 / \sigma_{\eta_0}^2 \) determines the impulsive characteristic of \( \eta_1(n) \). For fixed value of \( \sigma_{\eta_0}^2 \), the larger the \( \gamma_0 \), the more impulsive \( \eta_1(n) \) becomes. The signal-to-noise ratio at the system output is defined as \( \text{SNR} = 10 \log_{10} (\sigma_{w'}^2 / \sigma_{\eta_0}^2) \), where \( \sigma_{\eta_0}^2 \) is the variance of \( d(n) \).

Simulation parameters and the initial values for various algorithms are shown in Fig. 2. For illustration purpose, from \( n = 1 \) to 1700, \( \eta(n) = \eta_0(n) \) is used. Whereas from \( n = 1701 \) to \( 2650 \), \( \eta(n) = \eta_0(n) + b(n) \eta_1(n) \) with \( p = 0.005 \) and \( \gamma_0 = 300 \) is used. To visualize clearly the effect of impulses in \( d(n) \), their locations generated by \( b(n) \) are fixed and marked in Fig. 2 but their amplitudes are varied according to \( \eta_0(n) \), which is generated statistically independent in each run. Also in order to visualize the effect of the impulse in the input signal, one impulse located at \( n = 500 \) is added to the filter input signal. The MSE results averaged over 200 independent runs are plotted in Fig. 2. From the Fig. 2, we have the following observations: (1) the performances of the H-PEF-LSL and the RLM algorithms are very close to each other; (2) The RLS and the PEF-LSL algorithms are significantly degraded by the impulses. The effect of a single impulse in \( d(n) \) and \( x(n) \) will last for more than 250 and 800 iterations for the RLS and the PEF-LSL algorithms, respectively; (3) the performance of the RLM and the H-PEF-LSL algorithms is very robust to the impulses in \( d(n) \) and \( x(n) \); (4) the initial convergence, steady state-error, and the tracking ability to the sudden change of the H-PEF-LSL algorithm are comparable to other algorithms considered.

5. CONCLUSION

A Huber Prior Error-Feedback Least Squares Lattice (H-PEF-LSL) algorithm for robust adaptive filtering in impulse noise environment is presented. It minimizes the modified Huber M-estimator based cost function, instead of the conventional least squares based cost function. In additional to improved robustness
to impulse, the simple modified Huber M-estimate cost function also allows us to perform the time and order recursive updates in the conventional PEF-LSL algorithm so that the arithmetic complexity can be significantly reduced to $O(M')$. The new algorithm can also be viewed as an efficient implementation of the recursive least M-estimate (RLM) algorithm recently proposed by the authors [3], which has a complexity of $O(M^2)$.

Simulation results show that the proposed H-PEF-LSL algorithm is more robust than the conventional PEF-LSL algorithm in suppressing the adverse influence of the impulses at the input and desired signals with small additional computational cost. Its initial convergence, steady state error, and tracking ability to sudden system change are comparable to the RLM algorithm.

\[ e_{\text{w},n}(n-1) = e_{\text{w},n}(n-1) - w_{\text{w},n}(n-2) b_{\text{w},n}(n-1), \]
\[ w_{\text{w},n}(n-1) = w_{\text{w},n}(n-2) + \gamma_{\text{w},n}(n-1)b_{\text{w},n}(n-1)e_{\text{w},n}(n-1), \]
\[ e_{\text{w},n}(n) = e_{\text{w},n}(n) - w_{\text{w},n}(n-1)b_{\text{w},n}(n), \]
\[ A_n = [a_{\text{w},n}(n), \ldots, a_{\text{w},n}(n-N_n + 1)], \]
\[ \delta^{\text{w}}(n) = \delta^{\text{w}}(n-1) + C_n(1-\lambda) \text{med}(A_n), \]
\[ \xi_n = k_2 \delta^{\text{w}}(n), \]
\[ \text{if } |\xi_n| > \xi_n, \text{ then } w_{\text{w},n}(n-1) = w_{\text{w},n}(n-2), e_{\text{w},n}(n) = \delta^{\text{w}}(n) \text{ end}. \]

6. REFERENCES


