Proceedings of the 5th WSEAS International Conference on Applications of Electrical Engineering, Prague, Czech Republic, March 12-14, 2006 (pp59-64)

One Dimensional Nonlinear Adaptive Filters for Impulse Noise Suppression

MILAN STORK[1], DANIEL MAYER[2] AND JOSEF HRUSAK[1]
[1]Department of Applied Electronics and Telecommunications and
[2]Department of Theory of Electrical Engineering
University of West Bohemia
UWB, P.O.Box 314, 30614 Plzen
CZECH REPUBLIC
, http://home.zcu.cz/~stork

Abstract: - An adaptive filter is essentially a digital filter with self-adjusting characteristic. It adapts, automatically, to changes in its input signals. The contamination of a signal of interest by other unwanted, often lager, signals or noise is a problem often encountered in many applications. Typical applications where adaptive filters are appropriate are the following: Digital communication using a spread spectrum, where a large jamming signal, possibly intended to disrupt communication, could interfere with the desired signal. The interference often occupies a narrow but unknown band within the wideband spectrum, and can only be effectively dealt with adaptively. Digital data communication over the telephone channel at the high data rate.

Adaptive algorithms are used to adjust the coefficients of the digital filter such that error signal is minimized according to some criterion, for example in the least squares sense. The Nonlinear Normalized Mean Square algorithm is applicable to a wide variety of nonlinear filters. In this paper, algorithms are developed for an optimal time-varying step-size for FIR, Volterra, weighted median and weighted myriad filters.

Key-Words: - Adaptive filters, nonlinear filters, weighted median filter, weighted myriad filter, impulse noise.

1 Introduction

The Least Mean Square (LMS) algorithm [1] is widely used for adapting the weights of a linear Finite Impulse Response (FIR) filter that minimizes the mean square error (MSE) between the filter output and a desired signal. Consider an input (observation) vector of N samples, $\boldsymbol{x} \equiv [x_1, x_2, ..x_N]^T$, and a weight vector of N weights, $\boldsymbol{w} \equiv [w_1, w_2, ..w_N]^T$. Denoting the linear filter output by:

$$y = \boldsymbol{w}^T \boldsymbol{x} \tag{1}$$

The filtering error, in estimating a desired signal d, is:

$$e = y - d \tag{2}$$

The optimal filter weights minimize the MSE cost function:

$$J(\boldsymbol{w}) \equiv E\{e^2\}, \tag{3}$$

where $E\{.\}$ denotes statistical expectation. In an environment of unknown or changing signal statistics, the LMS algorithm [1] attempts to minimize the MSE by continually updating the weights as

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) - \mu e(n)\boldsymbol{x}(n), \tag{4}$$

where $\mu>0$ is the so-called step-size of the update.

The computational simplicity of the LSM algorithm has made it an attractive choice for several applications in linear signal processing. However, it suffers from a slow rate of convergence. Further, its implementation requires the choice of an appropriate step-size μ which affects the stability, steady-state MSE and convergence speed of the algorithm. The stability region for mean-square convergence [1, 2] of the LMS algorithm is given by:

$$0<\mu<(2/trace(\boldsymbol{R})), \tag{5}$$

where

$$\boldsymbol{R} \equiv E\{\boldsymbol{x}(n)\boldsymbol{x}^T(n)\}, \tag{6}$$

is the autocorrelation matrix of the input vector $\boldsymbol{x}(n)$. When the signal statistics are unknown or time-varying, it is difficult to choose a step-size that is guaranteed to lie within the stability region.

The so-called Normalized LMS (NLMS) algorithm [1] addresses the problem of the step-size design in [1] by choosing a time-varying step-size $\mu(n)$ that minimizes the next-step MSE, $J_{n+1} \equiv E\{e^2(n+1)\}$. After incorporating an auxiliary fixed step-size $\mu_a>0$, the NLMS algorithm is written as

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) - \mu_a e(n)\frac{\boldsymbol{x}(n)}{||\,\boldsymbol{x}(n)\,||^2}, \tag{7}$$

Proceedings of the 5th WSEAS International Conference on Applications of Electrical Engineering, Prague, Czech Republic, March 12-14, 2006 (pp59-64)

where

$$\| \boldsymbol{x(n)} \|^2 = \sum_{i=1}^{N} x_i^2(n), \tag{8}$$

is the squared Euclidean norm of the input vector $\boldsymbol{x(n)}$. The theoretical bounds on the stability of NLSM algorithm are given by $0<\mu_a < 2$ [1]. Unlike the LMS step-size μ of (4), the auxiliary step-size μ_a is dimensionless and the stability region for μ_a is independent of the signal statistics. This allows for an easier step-size design with guaranteed stability of the algorithm. Further, the NLMS algorithm is known to converge much faster than the LMS algorithm [3, 4]. The equation (7) can be also interpret as a modified LMS algorithm, where the update term in (4) is divided (normalized) by the squared-norm $\|\boldsymbol{x(n)}\|^2$, to ensure stability under large change of the input vector $\boldsymbol{x(n)}$.

In this paper, the generalized NLMS algorithm of (7) is given by deriving a class of nonlinear normalized LMS-type algorithms that are applicable to a wide variety of nonlinear filter structures. Although linear filters are useful in a number of applications, several practical situations require nonlinear processing of the signals. Consider an arbitrary nonlinear filter whose output is denoted by $y= y(\boldsymbol{w},\boldsymbol{x})$. The LMS algorithm of (4) can be generalized to yield the following class of nonlinear LMS-type adaptive algorithm:

$$w_i(n+1) = w_i(n) - \mu e(n) \frac{\partial y}{\partial w_i}(n), \quad i = 1, 2..., N \tag{9}$$

Note that (9) can be applied to any nonlinear filter for which derivatives (10) exist.

$$\frac{\partial y}{\partial w_i}(n) \tag{10}$$

The above algorithm inherits the main problem of LMS algorithm, namely, the difficulty in choosing the step-size $\mu > 0$. Unlike the linear case where step-size bounds are available, the complexity inherent in most nonlinear filters has precluded a theoretical analysis of (9) to derive the stability range for μ. Just as the linear NLMS algorithm of (7) is developed from the classical LMS algorithm, we obtain a general nonlinear NLMS-type algorithm for the LMS-type algorithm of (9) by choosing a time-varying step-size $\mu(n)$, which minimizes the next-step MSE at each iteration. As in the linear case, we introduce a dimensionless auxiliary step-size whose stability range has the advantage of being independent of the signal statistics. The stability region could therefore be determined empirically for any given nonlinear filter.

2 Nonlinear LMS-type Filters Adaptive Algorithms

In this section, we briefly review the derivation of nonlinear LMS-type adaptive algorithms that have been used in the literature for the optimization of several types of nonlinear filters. Consider a general nonlinear filter with the filter output given by $y = y(\boldsymbol{w}, \boldsymbol{x})$, where $\boldsymbol{x}$ and $\boldsymbol{w}$ are the N-long input and weight vectors, respectively. The optimal filter weights minimize the mean square error (MSE) cost function:

$$J(\boldsymbol{w}) = E\{e^2\} = E\{(y(\boldsymbol{w}, \boldsymbol{x}) - d)^2\}, \tag{11}$$

where d is the desired signal and $e = y - d$ is the filtering error. The necessary conditions for filter optimality are obtained by setting the gradient of the cost function equal to zero:

$$\frac{\partial J(\boldsymbol{w})}{\partial w_i} = 2E\left\{ e \frac{\partial y}{\partial w_i} \right\} = 0, \ i = 1, 2, .., N \tag{12}$$

Due to the nonlinear nature of $y(\boldsymbol{w}, \boldsymbol{x})$, and consequently of the equation in (12), it is extremely difficult to solve for the optimal weights in closed-form. The method of steepest descent is a popular technique which attempts to minimize the MSE by continually updating the filter weights using the following equation:

$$w_i(n+1) = w_i(n) - \frac{1}{2}\mu \frac{\partial J}{\partial w_i}(n), \ i = 1, 2.., N, \tag{13}$$

where $w_i(n)$ is the ith weight at iteration n, $\mu > 0$ is the step-size of the update, and the ith component of the gradient at the nth iteration is given from (12) by

$$\frac{\partial J}{\partial w_i}(n) = 2E\left\{ e(n) \frac{\partial y}{\partial w_i}(n) \right\} \tag{14}$$

In a situation where the signal statistics are either unknown or rapidly changing (as in nonstationary environment), we use instantaneous estimates for the gradient.

To this end, removing the expectation operator in (14) and substituting into (13), we obtain the following class of nonlinear LMS-type adaptive algorithms:

$$w_i(n+1) = w_i(n) - \mu e(n) \frac{\partial y}{\partial w_i}(n) \tag{15}$$

Note that for a linear filter ($y = \boldsymbol{w}^T \boldsymbol{x}$), we have:

$$\frac{\partial y}{\partial w_i} = x_i, \tag{16}$$

and (15) reduces to the LMS algorithm of [1]. The development of automatic step-size that guarantees the stability of (15) is derived in following section.

3 Normalized LMS-Type (NLMS) Adaptive Algorithms

We derive the class of nonlinear NLMS-type algorithms by choosing a time-varying step-size $\mu(n)>0$ in the LMS-type algorithm of (15). Rewriting of (13), using (12) is:

$$w_i(n+1) = w_i(n) - \mu E\left\{ e(n)\frac{\partial J}{\partial w_i}(n) \right\} \qquad (17)$$

Now, the next-step MSE at the *n*th iteration is defined by:

$$J_{n+1} \equiv J(w(n+1)) = E\{e^2(n+1)\}, \qquad (18)$$

where the next-step filtering error $e(n+1)$ is:

$$e(n+1) = y(n+1) - d(n+1) =$$
$$= y(\mathbf{w}(n+1), \mathbf{x}(n+1)) - d(n+1) \qquad (19)$$

Note, that J_{n+1} depends on the updated weight vector $\mathbf{w}(n+1)$, which in turn is a function of $\mu > 0$. We obtain the NLMS-type algorithm from (17) by determining the optimal step-size, denoted by $\mu_o(n)$, that minimizes $J_{n+1} \equiv J_{n+1}(\mu)$:

$$\mu_o(n) \equiv \arg\min_{\mu>0} J_{n+1}(\mu) \qquad (20)$$

To determine $\mu_o(n)$, we need an expression for the derivative function $(\partial/\partial\mu)\ J_{n+1}(\mu)$. Using (18) and (19):

$$\frac{\partial}{\partial\mu} J_{n+1}(\mu) = \sum_{j=1}^{N} \frac{\partial J_{n+1}(.)}{\partial w_j(n+1)} \frac{\partial w_j(n+1)}{\partial\mu} \qquad (21)$$

Detailed evaluation for optimal step size $\mu_o(n)$ is derived in [5]. Simplified expression for the optimal step-size is:

$$\mu_o(n) \approx \frac{1}{\sum_{j=1}^{N}\left(\frac{\partial y}{\partial w_j}(n)\right)^2} \qquad (22)$$

After incorporating an auxiliary step-size $\mu_a > 0$, just in the conventional (linear) NLMS algorithm of (7), we can then write the time-varying step-size, to be used in the steepest-descent algorithm of (17), as:

$$\mu(n) = \mu_a . \mu_o(n) \approx \frac{\mu_a}{\sum_{j=1}^{N}\left(\frac{\partial y}{\partial w_j}(n)\right)^2} \qquad (23)$$

Finally, on using instantaneous estimates by removing the expectation operator in the steepest-descent algorithm of (17), we obtain the following Nonlinear Normalized LMS-type Adaptive Filtering Algorithm:

$$w_i(n+1) = w_i(n) - \frac{\mu_a}{\sum_{j=1}^{N}\left(\frac{\partial y}{\partial w_j}(n)\right)^2} e(n)\frac{\partial y}{\partial w_i}(n) \qquad (24)$$
$$i = 1,2,\ldots,N$$

This algorithm has following important advantages:

- It is applicable to wide variety of nonlinear filters; in fact, to any nonlinear filter for which the filter output y is an analytic function of each of the filter weights w_i (so that derivatives of all orders exist).
- The auxiliary step-size μ_a is dimensionless and the stability region for μ_a is independent of the signal statistics. As a result, the stability region could be determined empirically for any particular nonlinear filter.
- This algorithm has a potentially faster convergence than its LMS-type counterpart of equation (15).
- It can also be interpreted as a modification of the LMS-type algorithm of (15) in which the update term is divided (normalized) by the Euclidean squared-norm of the set of values:

$$\frac{\partial y}{\partial w_i}(n), \quad i = 1,2,..,N, \qquad (25)$$

in order to ensure algorithm stability when these values become large in magnitude.

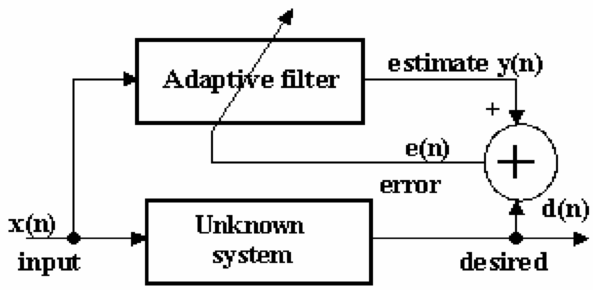

Figure. 1. Unknown system identification by using adaptive filter

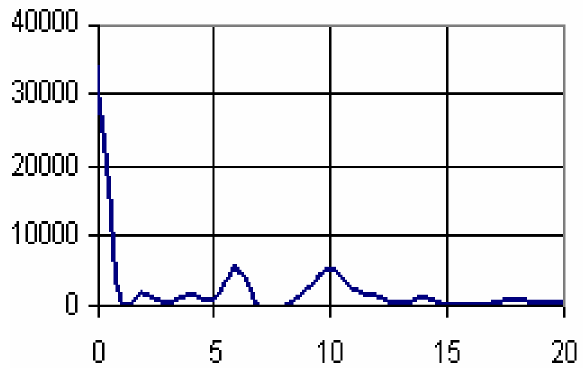

Figure 2. Adaptive FIR filter. MSE as a function of algorithm iterations ($\mu_a = 1$).

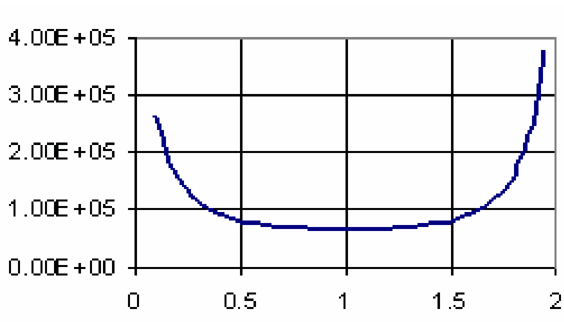

Figure 3. Adaptive FIR filter. MSE as a function of auxiliary step-size ($\mu_{a.}=1$ is optimal).

4 Simulation Results

In this part, the normalized adaptive algorithms used in different examples are described. The linear and nonlinear system identification, Figure 1, was tested The linear FIR filter, $N=11$, was tested. Zero-mean white Gaussian signal was used. MSE learning curve is shown in Figure 2. Figure 3 shows a MSE as a function of auxiliary step-size $\mu_{a.}$.

As a second example, the normalized adaptive Volterra filter was tested [4, 5, 6, 7].

The Volterra filter belongs to a particularly simple class of nonlinear filters having the property, that the filter output is linear in the filter parameters (weights). Given N x 1 input (observation) vector $\boldsymbol{x}$, the filter output in this class is given by:

$$y = \boldsymbol{h}^T \boldsymbol{z} = \boldsymbol{h}^T f(\boldsymbol{x}), \tag{26}$$

where $\boldsymbol{h}$ is a M x 1 vector of filter parameters, and f is a (generally nonlinear) mapping that transforms the N x 1 input vector $\boldsymbol{x}$ into an M x 1 modified observation vector $\boldsymbol{w}$. Consider now the special case of the Volterra filter, which has found wide-spread use in nonlinear signal processing [6, 9]. The output of this filter is given by:

$$\begin{aligned} y &= \sum_{i=1}^{N} w_1(i)x_i + \sum_{i=1}^{N}\sum_{j=1}^{N} w_2(i,j)x_i x_j + .. \\ &= \boldsymbol{h}_1^T \boldsymbol{z}_1 + \boldsymbol{h}_2^T \boldsymbol{z}_2 + .. = \boldsymbol{h}^T \boldsymbol{z} \end{aligned} \tag{27}$$

In order to obtain the NLMS-type adaptive filtering algorithm for this filter, from (26) and (27) can be derived:

$$\begin{aligned} \| \boldsymbol{z}(n) \|^2 &= | \boldsymbol{z}_1(n) \|^2 + | \boldsymbol{z}_2(n) \|^2 + ... \\ &= \sum_{i=1}^{N} x_i^2(n) + \sum_{i=1}^{N}\sum_{j=1}^{N} x_i^2(n) x_j^2(n) + ... \end{aligned} \tag{28}$$

The NLMS-type algorithm can be written as:

$$\boldsymbol{h}(n+1) = \boldsymbol{h}(n) - \frac{\mu_a}{\| \boldsymbol{z}(n) \|^2} e(n)\boldsymbol{z}(n) \tag{29}$$

The algorithm is similar to the linear NLMS algorithm of (7), since the filter output is linear in filter parameters. Therefore, the stability region for μ is the same as in linear case. The nonlinear truncated Volterra filter, $N=5$, was tested. The Volterra filter has a 20 parameters (linear and quadratic):

$$\begin{aligned} & x_1, x_2, x_3, x_4, x_5, \\ & x_1^2, x_1x_2, x_1x_3, x_1x_4, x_1x_5, \\ & x_2^2, x_2x_3, x_2x_4, x_2x_5, \\ & x_3^2, x_3x_4, x_3x_5, \;\; x_4^2, x_4x_5, x_5^2 \end{aligned} \tag{30}$$

Zero-mean white Gaussian signal was used. MSE learning curve is shown in Figure 4.

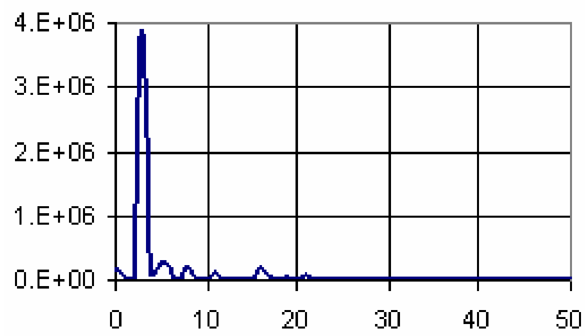

Figure 4. Adaptive Volterra filter. MSE as a function of algorithm iterations ($\mu_a=1$).

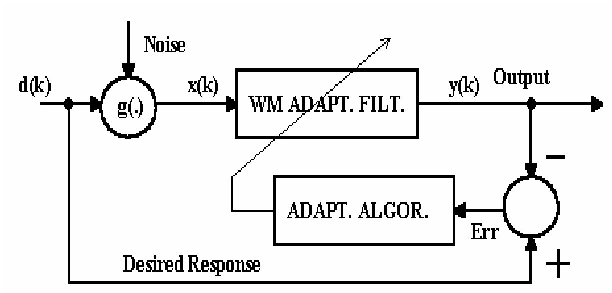

Figure 5. Adaptive weighted median filter block diagram

As third example, the adaptive weighted median filter (WM), $N=5$, was tested, Figure 5. For the discrete-time continuous-valued input vector $\boldsymbol{x} = [x_1, x_2,..., x_R]$, the output y of the WM filter of span N associated with the integer weights:

$$\boldsymbol{w} = [w_1, w_2,..., w_R] \tag{31}$$

is given by

$$y = MED[w_1 \Diamond x_1, w_2 \Diamond x_2, ..., w_R \Diamond x_R] \tag{32}$$

where $MED[.]$ denotes the median operation and $\Diamond$ denotes duplication, i.e.:

$$p \Diamond x = \overbrace{x, x ... x}^{p - times} \tag{32}$$

(x is used p - times)

Example: Consider a length 5 WM filter with integer weights *[1, 2, 3, 2, 1]*. Now apply the filter to the following sequence so that the window is centered at the sample value 9:

$$x = [-2, 4, 9, 12, -3]$$

After sorting and duplication, the samples inside filter window are *[12, 12, 9, 9, 9, 4, 4, -2, -3]*. The WM filter output is $y=9$ (middle).

The adaptive WM filter was used for filtering of signal corrupted by impulsive noise (Figure 6). Final filtering algorithm for adaptive WM filter is given by:

$$w_i(n+1) = w_i(n) - \frac{\mu_a[d(n)-y(n)]sign\{sign[w_i(n)x(n)]-y(n)\}}{1+\sum_{j=1}^{N} abs\{sign[w_i(n)]x(n)\}}, \quad (34)$$

where $i=1,2,....N$, and weights of adaptive WM filter during learning are shown in Figure 7.

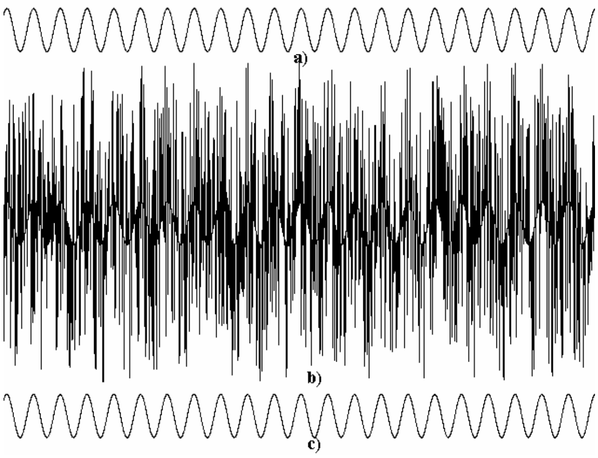

Figure 6. Example of sinusoidal signal a), corrupted by impulsive noise b) and signal filtered by NLMS-type adaptive weighted median filter c), N=5.

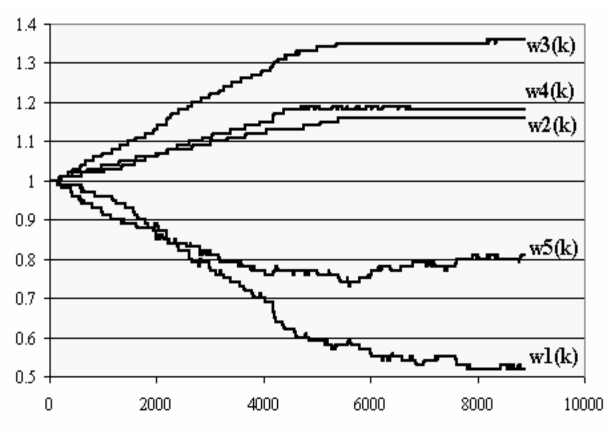

Figure 7. Weights of adaptive median filter evaluation during the learning. Weights after learning, (9000 iterations) are:
$\boldsymbol{w}$= *[0.51, 1.16, 1.35, 1.18, 0.81]*.

As the last example, the weighted myriad filter, $N=11$, was tested. Weighted myriad smoothers [8, 9, 10, 11] are derived from the sample myriad, which is defined as the Maximum-Likelihood estimate of location of data following the Cauchy distribution. Consider a set of N independent and identically distributed random variables $\{X_i\}^N_{i=1}$, each following a Cauchy distribution with location parameter ψ and scaling factor $K>0$. Thus, $X_i \sim Cauchy(\psi, K)$, with the density function:

$$f_{X_i}(x_i;\psi,K) = \frac{K}{\pi}\frac{1}{K^2+(x_i-\psi)^2} = \frac{1}{K} f(\frac{x_i-\psi}{K}) \quad (35)$$

Given a set of observations $\{x_i\}^N_{i=1}$, the sample Myriad maximizes the likelihood function:

$$\prod_{i=1}^{N} f_{X_i}(x_i;\psi,K) \quad (36)$$

Using some manipulations, we can write:

$$\begin{aligned}\hat{\psi} &\triangleq myriad(x_1, x_2, .., x_N ; K) = \\ &= \arg\min_{\psi} \prod_{i=1}^{N}\left[1+(\frac{x_i - \psi}{K})^2\right] = \\ &= \arg\min_{\psi} \sum_{i=1}^{N} log[K^2 + (x_i - \psi)^2] \end{aligned} \quad (37)$$

It is important to note, that $log(.)$ is strictly increasing function. For NLMS-type adaptive algorithm (33) is rewritten as:

$$\begin{aligned} &y \equiv y(\boldsymbol{h},\boldsymbol{x}) \triangleq \\ &\arg\min_{\psi} \sum_{i=1}^{N} log[1+|h_i|(\psi - sig(h_i)x_i)^2], \end{aligned} \quad (38)$$

where $\boldsymbol{h}$ = $\boldsymbol{w}/K^2$. The filter can therefore be adaptively optimized by updating the parameter vector $\boldsymbol{h}$. The corresponding normalized adaptive (NLMS-type) algorithm is given by:

$$h_i(n+1) = h_i(n) + \mu_a e(n)\frac{\Delta(n)}{\sum_{j=1}^{N}\delta_j^2(n)}, \quad (39)$$

$$i = 1, 2, ..., N,$$

where

$$\delta_i = \frac{u_i}{(1+|h_i|u_i^2)^2}, \quad i = 1,2,..,N, \quad (40)$$

and

$$\Delta = \sum_{j=1}^{N} |h_j| \frac{1-|h_j|u_j^2}{(1+|h_j|u_j^2)^2}, \quad (41)$$

with $u_i = sgn(h_i)y - x_i$, $i=1, 2,...,N$.

The example of NLMS-type of adaptive weighted myriad filter was used to extract corrupted impulsive noise from sinusoidal signal. The observed signal was given by:

$$x(n) = s(n) + v(n) \quad (42)$$

The additive noise process *v(n)* was chosen to have a zero-mean symmetric α-stable distribution. The example in is shown in Figure 8.

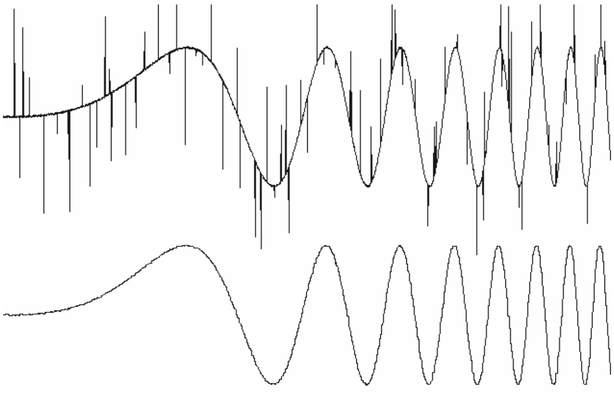

Figure 8. Example of sinusoidal chirp signal corrupted by impulsive noise (top) and signal filtered by NLMS-type adaptive weighted myriad filter (bottom) N=11.

5 Conclusion

It is well known that, for any algorithm with fixed value of tap weight adaptation step size, a trade-off between the filter convergence rate and the steady-state error does exist. If a larger value of the step size is used, then a faster convergence is attained as long as the filter remains stable. On the other hand, the smaller the step size, the more accurate the estimation in presence of observation noise.

In this paper, examples of the NLMS-type adaptive filters were described. By providing for an automatic choice for the step-size, the NLMS-type algorithms eliminate the difficult problem of step-size design. Experimental results indicate (computer simulations examples: linear adaptive FIR filter, adaptive Volterra filter and adaptive weighted median filter), that the NLMS-type algorithms, in general, converge faster than their LMS-type counterparts. Examples presented in [12, 13] show that nonlinear filters can be used in industry applications and also in biomedical signal processing.

ACKNOWLEDGMENT

This research work was partly supported from research project *Diagnostics of interactive phenomena in electrical engineering*, MSM 49777513110 and Dept. of Applied Electronics and Telecommunication, University of West Bohemia, Plzen, Czech Republic

References:

[1] S. Haykin, Adaptive Filter Theory, Englewood Cliffs, Prentice Hall, 1996.

[2] M. Rupp, "The behavior of LMS and NLMS algorithms in the presence of spherically invariant processes", IEEE Transaction on Signal Processing, vol. 41, Mar., 1993, pp. 1149-1160.

[3] S. Kalluri, G. R. Arce, "A general class of nonlinear normalized LMS-type adaptive algorithms", Proceedings of the 1998 CISS, Princeton, Marc 1998.

[4] M. Schetzen, The Volterra and Wiener Theories of Nonlinear Systems. New York, Willey, 1980.

[5] S. Benedetto, E. Biglieri, V. Castellani, V.: Digital Transmission Theory. Englewood Cliffs, NJ: Prentice Hall, 1987.

[6] T. Koh, E. J. Powers, "Second-order Volterra filtering and its application to nonlinear system identification" IEEE Transaction on Acoustics, Speech and Signal Processing. V ASSP-33, December 1985, pp. 1445-1455.

[7] I. Pitas, A. N. Vanetsanopoulos, Non-linear Digital Filters - Principles and Applications. Boston, MA: Kluwer, 1990.

[8] I. Pitas, A. N. Vanetsanopoulos, "Order Statistics in Digital Image Processing", Proceedings of the IEEE, vol. 80, December 1992.

[9] S. Kalluri, G. R. Arce, "Adaptive weighted myriad filter algorithms for robust signal processing in α-stable noise environments", IEEE Transaction on Signal Processing, vol. 46, Feb. 1998, pp. 322-334.

[10] G. R. Arce, "A general weighted median filter structure admitting negative weights", IEEE Transactions on Signal Processing. Vol. 46, December 1998, pp. 3195-3205.

[11] C. L. Nikias, M. Shao, Signal Processing with Alpha-Stable Distributions and Applications. New York, Willey, 1995.

[12] M. Stork, "Standard Median and Recursive Median Filter Applications", IEEE Design and Diagnostic of Electronics Circuits and Systems, Conference. ISBN 80-214-2094-4, Brno, April 2002, pp. 142-149.

[13] M. Stork, "Median Filters For Some Artifacts And Impulse Noise Filtering", 16-th Biennal International Eurasip Conference Biosignal 2002, Brno, ISSN 1211-412X, ISBN 80-214-2120-7, pp. 126-128