A NEW DESIGN METHODOLOGY FOR L-FILTERS

Remzi Öten and Rui J.P. de Figueiredo

Laboratory for Machine Intelligence, Neural and Soft Computing
Engineering Tower 616, University of California, Irvine, CA 92697-2625
Tel:(949)824-7043, Fax:(949)824-3408, e-mail:rui@uci.edu

ABSTRACT

L-filters are very successful at restoration of signals corrupted by noise. However, their on-line design in real-time applications is practically impossible, since the computation of optimal L-filter coefficients with the existing methods based on numerical approximations are extremely time consuming. In this paper, we present a new design methodology for L-filters. Our approach allows one to compute the desired filter coefficients by just evaluating the inverse cumulative distribution function and the probability density function of the underlying noise at n different points, where n is the sample size. The presented design method is derived by approximating the inverse covariance matrix of ordered random variables and simplifying the closed-form solution of optimal L-filter coefficients. The comparisons with the classical approach and the simulations show that the proposed method is very promising for a wide range of noise types.

1. INTRODUCTION

Nonlinear filters[1] have been shown to be more effective and robust than linear filters in many applications. However, the difficulty in design procedures and computational complexity of nonlinear filters have been major drawbacks of these filters.

Especially filters based on order statistics[2] (order filters [3][4]) are proven to be very effective and robust nonlinear filters. They have been successfully employed in restoration of signals and images corrupted by noise[5]. The most popular filter of this type is the median filter[6]. It is easy to implement and shows great performance on removing the impulsive noise types. Other simple and well-known filters of this type are the outer-mean (mid-point) filter, α-trimmed mean filter and L-filter[7]. As being the most general form among these filters, the L-filter has more flexible and complex structure. When designed properly it outperforms the linear filters and the above filters for most of the noise types.

However, for a given or estimated noise type, the classical design procedure of L-filters is computationally so complex. Especially, its design in real-time applications is prohibitive. Previous approaches that employ L-filters use iterative numerical approximation procedures[7][8][9][10] which are still computationally expensive and / or have convergence problems.

In this paper, we present a new fast design algorithm based on approximation theory and order statistics [11][2]. By proposing an analytical solution our method eliminates the extremely time consuming numerical approximation (integration) routines and makes the practical use of L-filters possible for many applications.

[1]This work was supported by the NSF Grant MIP 97-04262

The description of order filters and, particularly, L-filters is done in the next section. In section 3, optimal coefficient of L-filters are represented in a form such that the approximation of its components in sections 4 and 5 will be possible. The new closed form solution of these coefficients are presented in section 6. In section 7 we compared the method in section 6 with the classical method for various noise types. The outputs of these filters when applied to noisy signals are shown in section 8. The conclusion is drawn in section 9.

2. L-FILTERS

Consider the classical problem of estimating the constant amplitude signal θ from the samples, $x(i), x(i-1), ..., x(i-n+1)$, of a noisy observation data $\{x(i)\}$, where $n = 2N + 1$. Let

$$x(i) = \theta + \eta(i) \tag{1}$$

where the $\eta(i)$ are independent, identically distributed, zero mean noise samples with a symmetric probability density function(pdf), $f(.)$, i.e. $f(-\zeta) = f(\zeta)$, and a cumulative distribution function(cdf), $F(.)$, i.e. $F(\zeta) = 1 - F(-\zeta)$. In statistical estimation theory, estimation of θ is called "location estimation" and the sample averaging, median, and the outer mean are examples of location estimators, and they are the maximum likelihood estimates of Gaussian, Laplacian, and Uniform distributions, respectively. Thus, the filters designed to remove noise types with these specific distributions are the well-known Moving Average, Median, and Outer Mean Filters, respectively.

If the above data samples are arranged in ascending order of their magnitude, the order statistics result is

$$x_{(1)}(i) \leq x_{(2)}(i) \leq ... \leq x_{(n)}(i) \tag{2}$$

where $x_{(1)}(i)$ is the minimum, $x_{(n)}(i)$ is the maximum, and $x_{(N+1)}(i)$ is the median of the above set of observation data. The output of an order filter is:

$$y(i) = g(\mathbf{x_i}) \tag{3}$$

where g is a real-valued not-necessarily-linear function of n-vector $\mathbf{x_i} = [x_{(1)}(i)x_{(2)}(i)...x_{(n)}(i)]^T$.

If g is linear (3) becomes

$$y(i) = \mathbf{a^T x_i} = \sum_{j=1}^{n} a_j x_{(j)}(i) \tag{4}$$

where $\mathbf{a} = [a_1 a_2 ... a_n]^T$. This type of filters are called "OS-Filters" or "L-Filters".

Median filter is a special type of L-filter whose coefficients are $a_{N+1} = 1, a_i = 0, i = 1,2,..,n,\ i \neq N+1$. Moving Average Filters and Outer Mean Filters are also special type of L-filters whose coefficients are $a_i = 1/n, i = 1,2,...,n$, and $a_1 = a_n = 1/2, a_i = 0, i = 1,2,..,n, i \neq 1, n$, respectively.

In [7], the optimal coefficients $\mathbf{a}$ have been determined in Mean Square Error (MSE) sense, with the location invariance constraint, i.e. $\mathbf{a}^T e = 1$, where $\mathbf{e} = [11\ldots 1]^T$. The optimal coefficients are

$$\mathbf{a} = \frac{\mathbf{R}^{-1}\mathbf{e}}{\mathbf{e}^T\mathbf{R}^{-1}\mathbf{e}} \tag{5}$$

where $\mathbf{R}$ is the correlation matrix of the ordered noise variables $\eta_{(k)}, k = 1,\ldots,n$, and i^{th} row, j^{th} column element, R_{ij} can be computed by

$$R_{ij} = \iint_{-\infty}^{\infty} xy g_{\eta_{(i)}\eta_{(j)}}(x,y)dxdy \quad (i<j)$$
$$R_{ii} = \int_{-\infty}^{\infty} x^2 g_{\eta_{(i)}}(x)dx \tag{6}$$

where

$$g_{\eta_{(i)}}(x) = K_i f^{i-1}(x)[1-F(x)]^{n-i} f(x) \tag{7}$$
$$g_{\eta_{(i)}\eta_{(j)}}(x,y) = K_{ij} F^{i-1}(x)[F(y)-F(x)]^{j-i-1} \cdot [1-F(y)]^{N-i} f(x) f(y) \tag{8}$$

where $K_i = N!/[(i-1)!(N-i)!]$ and $K_{ij} = [N!/[(i-1)!(j-i-1)!(N-j)!]$

Since the above expressions in (8) are too complex, generally, computation of the optimal coefficients require the double and single numerical integration. This is computationally so expensive. Besides, in many cases, to obtain a reasonable precision, the intervals of the numerical integration routine should be kept very small [7], which makes it even more prohibitive.

3. OPTIMAL COEFFICIENTS OF L-FILTERS

In this section, for simplicity in notation, we will drop the time index, i, of the variables, i.e., $x = \theta + \eta$ and the ordered random variables in (2) are $x_{(1)} \leq x_{(2)} \leq \ldots \leq x(n)$.

Let the noise random variable,$\eta = \sigma v$, where $\sigma > 0$ is the standard deviation of η. With this normalization and previous assumptions, v is a zero mean, unit variance random variable with symmetric pdf f, and cumulative distribution function F.

The order in x holds for v, i.e., $x_{(j)} = \theta + \sigma v_{(j)}$ and $v_{(1)} \leq v_{(2)} \ldots \leq v_{(n)}$. Let $\mathbf{C}'$ is the covariance matrix of the random variable $v_{(i)}$ and c'_{ij} is the element of C' at i-th row and j-th column, i.e.,

$$c'_{ij} = \mathcal{E}\{v_{(i)}v_{(j)}\} - \mathcal{E}\{v_{(i)}\}\mathcal{E}\{v_{(j)}\} \tag{9}$$

Now, let us represent the L-filter in (4) as:

$$d = \frac{\sum_{i=1}^n \beta_i x_{(i)}}{\sum_{i=1}^n \beta_i} \tag{10}$$

where the coefficients, a_i in (4) relate to β_i by

$$a_i = \frac{\beta_i}{\sum_{i=1}^n \beta_i} \quad i = 1,\ldots,n \tag{11}$$

which makes the filter location invariant whatever the choice of the coefficients, β_i.

It is known that optimal coefficients β_i are symmetric about the median, i.e. $\beta_i = \beta_{n-i+1}$. Therefore (10) can be written as

$$d = \frac{\sum_{i=1}^r b_i[x_{(i)} + x_{(n-i+1)}]}{2\sum_{i=1}^r b_i} \tag{12}$$

where $r = \lceil n/2 \rceil$ and $b_i = \beta_i$, except for odd n $b_r = \beta_r/2$.

To obtain an ideal estimator one has to minimize the mean square error,with respect to the coefficients, b'_i, $i = 1,\ldots,r$. Mean square error is,

$$\mathcal{E}\{(d-\theta)^2\} = \frac{\sigma^2 \sum_{i=1}^r \sum_{j=1}^r b_i b_j c_{ij}}{\sum_{i=1}^r \sum_{j=1}^r b_i b_j} \tag{13}$$

where $c_{ij} = (c'_{ij} + c'_{i,n-j+1} + c'_{n-i+1,j} + c'_{n-i+1,n-j+1})/4$. Since u is symmetric, $c'_{ij} = c'_{n-i+1,n-j+1}$ and $c'_{ij} = c'_{ji}$. Therefore,

$$c_{ij} = \frac{c'_{ij} + c'_{i,n-j+1}}{2} \tag{14}$$

With some simple algebra, similar to [7], it can be seen that the result of the minimization of (13) is equal to the solution of

$$\sum_{i=1}^r b_i c_{ij} = 1 \quad j = 1,\ldots,r \tag{15}$$

or in vectoral form: $\mathbf{Cb} = \mathbf{e}$, where $\mathbf{C}$ is the matrix that contains c_{ij}s, $\mathbf{b} = [b_1 \ldots b_r]^T$ and $\mathbf{e} = [1\ldots 1]^T$. Then, the solution is

$$\mathbf{b} = \mathbf{C}^{-1}\mathbf{e} \tag{16}$$

Note that this solution is similar to the numerator of (5). And its denominator is nothing but the normalization over all the coefficients.

Let $\bar{c}_{ij}$ is the element of $\mathbf{C}^{-1}$ at the i-th row and j-th column. Then the solution can be written as

$$b_i = \sum_{j=1}^r \bar{c}_{ij} \quad i = 1,\ldots,r \tag{17}$$

In [11] the L-estimator with these coefficients is shown to be an asymptotically efficient estimator. In the following sections, we will follow the same steps as in [11], however, our focus will be on the coefficients, b_i of (12).

4. APPROXIMATION OF THE COVARIANCES

Let u be a random variable that is uniformly distributed over the interval $0 \leq u \leq 1$. For an ordered sample $u_{(i)}$,$i = 1,..,n$, taken from this distribution, the mean and the covariances are,

$$\mathcal{E}\{u_{(i)}\} = \frac{i}{n+1} = \lambda_i$$
$$\mathbf{cov}(u_{(i)}, u_{(j)}) = \frac{\lambda_i(1-\lambda_j)}{n+2} \tag{18}$$

To obtain the random variable v from u we used a transformation, S, such that

$$v = S(u) = F^{-1}(u) \tag{19}$$

By using a Taylor expansion on S, and neglecting the higher order terms we have,

$$S(u_{(i)}) \approx S(\lambda_i) + (u_{(i)} - \lambda_i) S'(\lambda_i) \tag{20}$$

Then, the mean and the covariances of $v_{(i)}$ can be approximated as,

$$\mathcal{E}\{v_{(i)}\} = \mathcal{E}\{S(u_{(i)})\} \approx S(\lambda_i) = w_i \tag{21}$$

and

$$\begin{aligned} c'_{ij} &= \mathbf{cov}(v_{(i)}, v_{(j)}) = S'(\lambda_i)S'(\lambda_j)\mathbf{cov}(u_{(i)}, u_{(j)}) \\ &= \frac{\lambda_i(1-\lambda_j)}{(n+2)f(w_i)f(w_j)} 1 \le i, j \le r \end{aligned} \tag{22}$$

since

$$S'(u) = \frac{dv}{du} = \frac{1}{f(v)}$$

Note that, here, we assume that the derivative of $S(u)$ exists at $u = \lambda_i = i/(n+1)$. This implies the cdf, F, to be differentiable at λ_i and pdf, f to be continuous at w_i. This will be our only constraint in the following sections.

Remark: The empirical studies show that the approximation of the covariance c'_{ij} in (22) gives more precise results if $(n+2)$ term in the denominator of the last expression is replaced by n. However, using either representation does not make any difference in our computations since that term will be cancelled during the normalization of the optimal coefficients.

5. APPROXIMATE REPRESENTATION OF THE INVERSE MATRIX ELEMENTS

Since optimal coefficients can be written in terms of the elements, $\bar{c}_{ij}$, of the inverse of the covariance matrix, $\mathbf{C}^{-1}$ as in (17), deriving a simple closed form representation for $\bar{c}_{ij}$s will be adequate. From the results in section 4 and (14), the c_{ij}s are approximately equal to:

$$c_{ij} \sim \frac{\lambda_i}{nf(w_i)f(w_j)} \quad 1 \le i, j \le r \tag{23}$$

where $\lambda_i = \frac{i}{n+1}$ and $w_i = F^{-1}(\lambda_i)$, $i = 1, \ldots, r$. If we take the inverse of the matrix, C whose elements are represented as in (23), then $\bar{c}_{ij}$s can be written as [12]:

$$\begin{array}{lll} \bar{c}_{11} \sim & 2nf^2(w_1)\left[\frac{1}{\lambda_1} + \frac{1}{\Delta\lambda_1}\right] & \\ \bar{c}_{ii} \sim & 2nf^2(w_i)\left[\frac{1}{\Delta\lambda_{i-1}} + \frac{1}{\Delta\lambda_i}\right] & 2 \le i < r-1 \\ \bar{c}_{rr} \sim & 2nf^2(w_r)\frac{1}{\Delta\lambda_{r-1}} & \\ \bar{c}_{ij} \sim & \bar{c}_{ji} = -2n\frac{f(w_i)f(w_j)}{\Delta\lambda_i} & i = j-1 = 1, \ldots, r-1 \\ \bar{c}_{ij} \sim & 0 & |i-j| > 1 \end{array} \tag{24}$$

In our case, $\Delta\lambda_i = \frac{1}{n+1}$. Therefore, $\Delta\lambda = \Delta\lambda_i = \lambda_1 = \frac{1}{n+1}$, and (25) becomes

$$\begin{array}{lll} \bar{c}_{ii} \sim & 4\frac{nf^2(w_i)}{\Delta\lambda} & 1 \le i \le r-1 \\ \bar{c}_{rr} \sim & 2nf^2(w_r)\frac{1}{\Delta\lambda} & \\ \bar{c}_{ij} \sim & \bar{c}_{ji} = -2n\frac{f(w_i)f(w_j)}{\Delta\lambda} & i = j-1 = 1, \ldots, r-1 \\ \bar{c}_{ij} \sim & 0 & |i-j| > 1 \end{array} \tag{25}$$

6. COEFFICIENTS

If we substitute, these approximate representations of $\bar{c}_{ij}$s into (17), and cancel the coefficients which appear both at the numerator and denominator of (12):

$$\begin{array}{ll} b_1 & = f(w_1)\left[-2f(w_1) + f(w_2)\right] \\ b_i & = f(w_i)\left[f(w_{i-1}) - 2f(w_i) + f(w_{i+1})\right] \quad i = 2 \ldots r-1 \\ b_r & = f(w_r)\left[f(w_{r-1}) - f(w_r)\right] \end{array} \tag{26}$$

where $w_i = F^{-1}(i/n+1)$.

Finally, the coefficients, β_i, $i = 1 \ldots, n$, of (10) are computed from b_i's, $i = 1 \ldots, r$ by

$$\beta_i = \beta_{n-i+1} = b_i \quad i = 1, \ldots, r \tag{27}$$

However, because of our new representation in (12), for odd n, $\beta_r = 2b_r$.

The coefficients a_i, $i = 1, \ldots, n$, are derived from β_is by the relation in (11).

As it is seen from (26), the basic component of the computation of these coefficients is the sampling (evaluation) of the pdf, $f(.)$, and inverse cdf, $F^{-1}(.)$, i.e. $w_i = F^{-1}(\lambda_i)$ and $f(w_i)$. The nature of this sampling is shown in Figure 1.

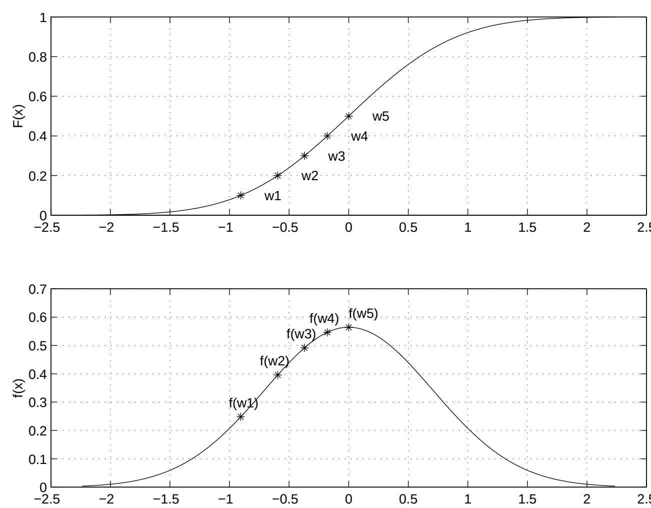


Figure 1: The sampling of the probability distribution function according to the sampling of the cumulative distribution function.

7. EXAMPLES

In this section, we will give examples to the L-filters designed with equations in (27). Due to the nature of our method, we are only interested in the $w_1 \le v \le w_r$ portion of the pdf and cdf domain, where $w_1 < 0$ and $w_r \le 0$. This implies that, in this region, $0 < F(w_1) < F(w_r) \le 0.5$. Therefore, the following expressions for different noise pdf's and cdf's may not be valid for the rest of the domain.

We will start with two simple and very well-known noise distributions: Uniform and Laplacian. Let the sample size n=5 and r=3.

The pdf and cdf of a symmetric uniform distribution is $f_u(x) = 1$ and $F_u(x) = x + 0.5$. Since $w_i = F_u^{-1}(i/6)$, for $i = 1, 2, 3$, $w_i = (i/6) - 0.5$ and $f_u(w_i) = 1$. If we substitute these into (26) then we have $b_1 = -1$ and $b_2 = b_3 = 0$. From (11) and (27) the coefficients, a_i, of the L-filter, (4), are equal to: $a_1 = a_5 = 0.5$ and $a_2 = a_3 = 0$, which is nothing but the 'outer mean filter'.

In the given domain region, the Laplacian distribution pdf and cdf are $f_l(x) = F_l(x) = (1/2)e^x$ and inverse cdf is $F_l^{-1}(x) = \ln(2x)$. Therefore, $w_1 = \ln(1/3)$, $w_2 = \ln(2/3)$, and $w_3 = 0$. Substitute these into $f_l(w_i)$: $f_l(w_1) = 1/6$, $f_l(w_2) = 1/3$, and $f_l(w_3) = 1/2$. Then, $b_1 = (1/6)[-2(1/6) + (1/3)] = 0$, $b_2 = (1/3)[(1/6) - 2(1/3) + (1/2)] = 0$ and $b_3 = (1/2)[(1/3) - (1/2)] = (-1/12)$. From (11) and (27) the coefficients, a_i, of the L-filter, (4), are equal to: $a_3 = 1$ and $a_1 = a_2 = a_4 = a_5 = 0$, which is the 'median filter'.

As it is seen from these examples, the method presented generated the optimal coefficients for uniform and Laplacian distributions with very simple calculations when the sample size, n=5. The computation complexity is $O(n)$. The speed of this method is enormously higher than the method in [7] since we do not employ any numerical integration procedure. In the rest of the section we compare the coefficients that are generated with the method presented in this paper to the coefficients in [7] for several distributions.These distributions are: U-shaped, Parabolic, Gaussian, Laplacian, and Uniform distributions.

In Tables 1-5, we present the L-filter coefficients generated with the classical method and our method for n=9. These methods are labeled as "I" and "II", respectively. Due to the symmetry, only the coefficients, a_i, $i = 1 \ldots r$, are presented. For method I, the coefficients, a_i, are the result of (5) and taken from [7]. For method II, the coefficients are the result of (26), (27), and (11).

In all these cases, the method presented in this paper produces coefficients that are very close to the output of (5) with very few computations.

8. SIMULATIONS

In this section, we will apply the L-filter with proposed coefficients and some other filters to a noisy signal. The signal is a portion of a received cable channel data stream. It is up-sampled and displayed in Figure 2-a. This data is obtained from Signal Processing Information Base (SPIB) data repository of Rice University.

Since the performances of the filters for the other noise types are discussed in section 7 and / or more predictable, we added U-shaped noise ($\sigma = 0.5$) to the signal. The noisy signal is shown in Figure 2-b.

The filter types, we use for the restoration of this signal are median filter, α-trimmed mean filter ($\alpha = 0.3$), moving average filter, outer-mean filter and, L-Filters with coefficients obtained by Method I and Method II. The signal-to-noise ratios for n=3, n=9, and n=25 are listed in Table 6.

The output of the Method II is presented in Figure 2-c.

9. CONCLUSION

In this paper, we have presented a very fast design method for L-filters by approximating the inverse of the covariance matrix of the ordered noise random variables.

From the results in section 7 and section 8, it is concluded that, our proposed approach to compute the coefficients of an L-filter can be employed in all applications where the speed of the design

Table 1: U-shaped distribution for n=9

Method	a_1	a_2	a_3
I	0.5548	-0.0263	-0.0146
II	0.4586	0.0068	0.0092

Method	a_4	a_5
I	-0.0099	-0.0081
II	0.0254	0.0000

Table 2: Parabolic distribution for n=9

Method	a_1	a_2	a_3
I	0.3403	0.0613	0.0443
II	0.3105	0.0755	0.0513

Method	a_4	a_5
I	0.0362	0.0357
II	0.0426	0.0402

Table 3: Gaussian distribution for n=9

Method	a_1	a_2	a_3
I	0.1111	0.1111	0.1111
II	0.1300	0.1072	0.1054

Method	a_4	a_5
I	0.1111	0.1111
II	0.1049	0.1048

Table 4: Laplacian distribution for n=9

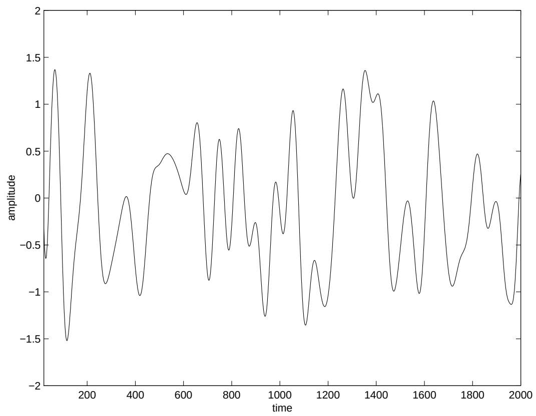
Method	a_1	a_2	a_3
I	-0.0190	0.02904	0.0697
II	0.0000	0.0000	0.0000

Method	a_4	a_5
I	0.2380	0.3647
II	0.0000	1.0000

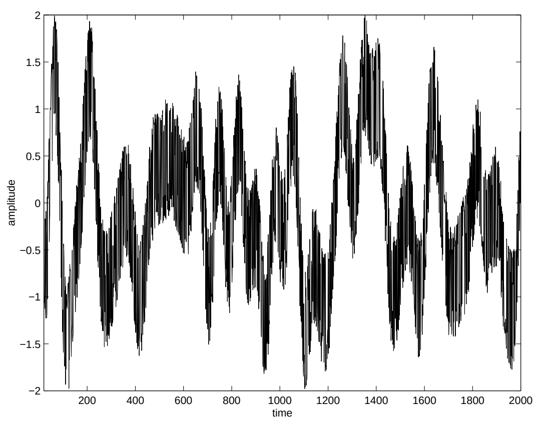
Table 5: Uniform distribution for n=9

Method	a_1	a_2	a_3
I	0.5000	0.0000	0.0000
II	0.5000	0.0000	0.0000

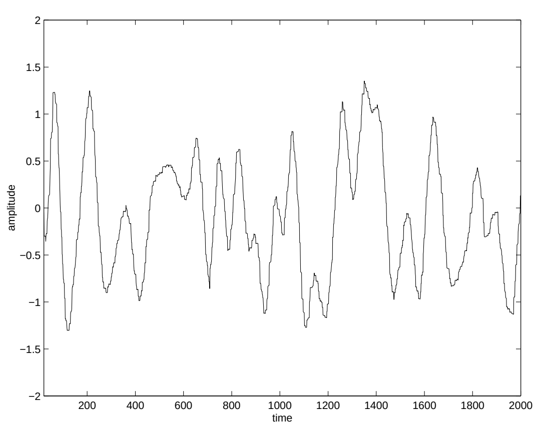
Method	a_4	a_5
I	0.0000	0.0000
II	0.0000	0.0000



(a) Original Signal



(b) Noisy Signal



(c) Restored by the Proposed Method

Figure 2: (a) The original signal is obtained by up-sampling a portion of a received cable channel data stream. (b) The signal in (a) is corrupted with U-shaped additive noise $\sigma = 0.5$ (c) The corrupted signal in (b) is restored by the L-filter whose coefficients are computed by the proposed method. The filter window size is 25

Table 6: Signal-to-Noise Ratio's (dB)

	$n = 3$	$n = 9$	$n = 25$
No Filter	0.2798	0.2798	0.2798
Median	0.3702	0.5229	0.8082
TriMean	0.3702	0.6886	1.0702
Mean	0.7449	1.2056	1.4923
MidPoint	0.8622	1.8106	1.5794
Method I	0.8669	1.8106	1.5829
Method II	0.8622	1.6153	1.5899

is important. An example to these applications is the real-time open-loop adaptive filtering of noisy signals. In this application, the noise type has to be estimated and a corresponding L-filter has to be designed in real-time for every data set. Especially when the sample size (or equivalently the window size) is large, the use of the proposed approach is necessary.

10. REFERENCES

[1] Jaakko Astola and Pauli Kuosmanen, *Fundamentals of Nonlinear Digital Filtering*, CRC Press, Florida, 1997.

[2] Herbert A. David, *Order Statistics*, Wiley Publications, New York, second edition, 1980.

[3] Rui J.P. de Figueiredo and Hemant D. Tagare, "Order filters," in *Conference Record of the 1983 Intnl. Symposium on Information Theory*, St. Jovite, Quebec, Canada, September 1983.

[4] Hemant D. Tagare and Rui J.P. de Figueiredo, "Order filters," *Proceedings of the IEEE*, vol. 73, pp. 163–165, January 1985.

[5] Ioannis Pitas and Anastasios N. Venetsanopoulos, "Order statistics in digital image processing," *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1892–1921, December 1993.

[6] J. W. Tukey, "Nonlinear (nonsuperposable) methods for smoothing data," in *Conf. Rec. EASCON*, 1974, p. 673.

[7] Alan C. Bovik, Thomas S. Huang, and David C. Munson, "A generalization of median filtering using linear combinations of order statistics," *IEEE Trans. on Acoust., Speech, Signal Processing*, vol. 31, no. 6, pp. 1342–1349, December 1983.

[8] Laith Naaman and Alan C. Bovik, "Least squares order statistic filters for signal restoration," *IEEE Transaction on Circuits and Systems*, vol. 38, no. 3, pp. 244–257, March 1991.

[9] I. Pitas and A.N. Venetsanopoulos, "Adaptive filters based on order statistics," *IEEE Transactions on Signal Processing*, vol. 39, no. 2, pp. 518–522, February 1991.

[10] C. Kotropoulos and I.Pitas, "Constrained adaptive lms l-filters," *Signal Processing*, vol. 26, pp. 335–358, 1992.

[11] Carl A. Bennett, *Asymptotic Properties of Ideal Linear Estimators*, Ph.D. thesis, University of Michigan, 1952.

[12] Frederick Mosteller, "On some useful inefficient statistics," *Annals of Mathematical Statistics*, vol. 17, pp. 377–408, 1946.