SOME GENERALIZATIONS OF MEDIAN FILTERS*

Yong Hoon Lee and Saleem A. Kassam

Department of Systems Engineering
University of Pennsylvania
Philadelphia, Pennsylvania 19104

ABSTRACT

L-smoothers and M-smoothers are introduced as generalizations of the median filter for nonlinear smoothing of noisy data, and their properties are derived. In addition, a double-window smoothing algorithm which is shown to be a data-dependent modification of L- and M-smoothers is proposed for filtering noisy signals with sharp edges. Simulation results are given to demonstrate the performance characteristics of these smoothing algorithms.

I. INTRODUCTION

In most signal processing applications removal of unwanted components from an input discrete-time sequence has been done successfully by linear filtering. If a signal with sharp edges is corrupted by high frequency noise, however, as in some noisy image data, then linear filters designed to remove the noise also smooth out signal edges. In addition, impulse noise cannot be reduced sufficiently by linear filters. A nonlinear scheme called median filtering [1,2] has been used with success in these situations. Some interesting results and analyses for median filters have been obtained recently, and may be found in [3-5].

The degree of smoothing behavior of a median filter can be influenced only by the window size, which may not allow sufficient degree of control [3]. In addition, from experience in linear filtering, an averaging operation is essential in removing high frequency noise components. In fact median filters often fail to provide sufficient smoothing of high frequency noise components for which the smoothing was originally designed [2]. Thus in general, it is desirable to have a nonlinear filtering algorithm which includes some averaging operation. A way to include some averaging in the median filtering algorithm is to use a combination of median and linear filters [1,2]. However, it might be expected that some other nonlinear techniques which inherently combine linear and nonlinear processing, and which allow a degree of control over these characteristics, could give better results.

In this paper we will study L-smoothers and M-smoothers, examine their properties, compare their performances, and obtain some simulation results. It should be noted that such smoothers have been previously considered in the statistical literature [6], but they have not yet been exploited for applications such as speech and image processing. In addition, we present a new nonlinear smoothing algorithm, which may be viewed as a data-dependent modification of L- and M-smoothing.

*This research is supported by the Air Force Office of Scientific Research under Grant AFOSR 82-0022.

II. DEFINITION AND PROPERTIES OF L-SMOOTHERS

Figure 1 illustrates the operation of an L-smoother. Each output point is obtained as a weighted sum of ordered data values in a sliding window, the set of weights determining the characteristics of the smoother. Clearly, L-smoothers include odd- and even span median filters [4], non-median n-th ranked-order operations which were proposed in [5], as extensions of median filters, and the running mean, as their extremes. Therefore, by proper choices of weights, L-smoothers are tunable from linear smoothers, i.e. the running mean, to extreme nonlinear smoothers, i.e. median or n-th ranked order operations.

L-smoothers have properties similar to those of median filters [2,6]. These include the following properties, where the output Y_k of a smoother S at time k is $Y_k = S(\underline{X})_k$ for input sequence $\underline{X}$:

(a) Time invariance: $S(B^M\underline{X})_k = S(\underline{X})_{k-M}$ where B is the back shift operator, $(B\underline{X})_k \triangleq \underline{X}_{k-1}$, and M is an integer.
(b) Scale and translation invariance: $S(a\underline{X}+b)_k = aS(\underline{X})_k + b$ for all a, b.
(c) L-smoothers pass a linear trend if the weights are symmetric, i.e. $S(\alpha+\beta t) = \alpha + \beta t$ for all α,β.
(d) Narrow pulses in the data will not be severely smeared out as long as the duration of the pulses exceed N when window length is 2N+1.
(e) L-smoothers with window length 2N+1 introduce a delay of N.

Figure 2 shows a comparison among several smoothing algorithms for a test signal with sharp edges. The signal is corrupted by additive independent and identically distributed Gaussian noise with zero mean and unit variance. Figures 2(a)-(b) show respectively, the input sequence and the result of smoothing it by a combination of median and linear filters (a 5-point running median followed by a 3-point Hanning window). Figures 2 (c)-(e) show the outputs of L-smoothing with window size 5, specifically for the running median, the running mean, and an L-smoother with weights (0,1/4,1/2, 1/4,0) respectively. We see that the L-smoother

9.9

 CH1841-6/83/0000-0411 $1.00

with weights (0,1/4,1/2,1/4,0) and window length 5 has performance very similar to that of the combination filter whose effective window size is 7. Both are better in supression of high frequency noise but worse with regard to edge-preservation, compared to the median filter. It is seen that the running mean smears the edges severely; it does perform better than the median filter for noise suppression. The random noise in the input does not seem to have been removed sufficiently by any of the smoothers in this example. This arises from a lack of sufficient averaging operations in the smoothers, which can only be improved by increasing window sizes. We will see in Section IV that a data-dependent smoother we will define there can be interpreted as an L-smoother with time-varying weights, possibly non-symmetric. This allows it to combine both edge-preserving and noise-suppression characteristics, the resulting smoother having a variable-length window.

III. DEFINITION AND PROPERTIES OF M-SMOOTHERS

M-smoothers are defined exactly as in robust estimation of a location parameter in an input sequence. Specifically, let ϕ be some odd function. The output Y_k of an M-smoother at time index k is defined through values of inputs $X_{k-N}, \ldots, X_{k+N}$ in a window length 2N+1, as the solution of the equation

$$\sum_{i=k-N}^{k+N} \phi\,(X_i - Y_k) = 0. \quad (1)$$

When ϕ is linear M-smoothers reduce to computing running means, while M-smoothing approaches median filtering as ϕ approaches the hard-limiter (signum function); this will be shown through Theorem 2. We will say that an M-smoother is of the "limiter-type" if

$$\phi(x) = \begin{cases} f(p) & \text{if } x > p \\ f(x) & \text{if } |f| \le p \\ -f(p) & \text{if } x < -p \end{cases} \quad (2)$$

where f(x) is a strictly increasing odd function and p is some positive number. When $f(x) = ax (a>0)$ we get in particular an M-smoother which will be described as being of the "standard type". The properties of L-smoothers considered in Section II also hold for M-smoothers, except for the fact that M-smoothers are not scale-invariant, i.e. $S(a\underline{X})_k \neq aS(\underline{X})_k$ in general.

The following results will give criteria which have to be considered in choosing ϕ, where we assume that ϕ is a bounded, odd function.

<u>Theorem 1</u>: If ϕ is continuous and nondecreasing then there always exists a solution Y_k in Equation (1). The proof is simple, and is omitted.

<u>Corollary 1</u>: If ϕ is continuous and strictly increasing then the solution Y_k in Equation (1) exists and is unique. The proof for Corollary 1 is obvious. We now show the uniqueness of the solution for a specific class of continuous, non-decreasing ϕ functions.

<u>Theorem 2</u>: For limiter-type M-smoothers with filter parameter p there exists a unique output Y_k at each time k, satisfying

$$m_k - f^{-1}\{\frac{N}{N+1} f(p)\} \le Y_k \le m_k + f^{-1}\{\frac{N}{N+1} f(p)\} \quad (3)$$

where m_k is the median in the window.

The proof of Theorem 2 is somewhat long, and is omitted. Now it is clear that as p goes to 0, a limiter-type M-smoother approaches a median filter. Note that when ϕ is strictly increasing as in Corollary 1, we cannot find any general bound for Y_k.

For a limiter-type M-smoother we get the median as the output when all other data values are far from the median. For a standard-type M-smoother the mean of the values in a window is obtained as the output if all the data values are close to the median. Thus we can say that, especially for a standard-type M-smoother, the performance is in between that of a median filter and that of a running mean, depending on the data. The averaging operation of a standard-type M-smoother may not in general be enough to suppress high frequency noise sufficiently because a reasonably small window size is still essential. It can be seen that the parameter p is a very important one in designing M-smoothers. The value p can be chosen from knowledge of the noise variance at the input or it may be obtained from the input in an adaptive scheme. As noted above, the parameter p controls the degree to which the smoother acts as a linear smoother. M-smoothers generally do not completely ignore outliers such as impulse noise, or sharp discontinuities. Again we can say that M-smoothers have linear properties at the expense of nonlinear properties, just like L-smoothers or combination smoothers. From Inequality (3), however, it is observed that the output of a limiter-type M-smoother can be closer to that of a median filter around sharp edges of an input, compared to the output of an L-smoother or a combination filter, when all three have similar random noise suppression properties. Therefore among M-, L- and combination smoothers which suppress high frequency noise in roughly equal degree limiter-type M-smoothers can be expected to have the best edge characteristic.

Figure 2(f) shows the result of M-smoothing with p=1, which is the same as the noise standard deviation, and window length 5. Sharp edges of the input are preserved almost as in median filtering, and high frequency noise is suppressed as in L-smoothing with weights (0,1/4,1/2,1/4,0) or combination filtering. Thus its performance looks superior to the other smoothers that we have considered so far.

IV. A DATA DEPENDENT DOUBLE-WINDOW SMOOTHING ALGORITHM

In the nonlinear smoothing algorithms discussed so far, large window sizes will generally imply a loss in the ability to preserve narrow pulses [3] and sharp discontinuities in the signal, but are required to suppress high frequency noise sufficiently. In addition we have observed that smoothers such as combination filters, L-smoothers and M-smoothers have linear characteristics at the expense of nonlinear characteristics. It does not appear that a single data-independent scheme can be designed to satisfy these requirements of good linear as well as nonlinear behavior at different sections of the input data. This leads to a consideration of data-dependent schemes; one reasonably simple and effective technique will now be introduced.

9.9

The algorithm we propose uses two windows centered at k simultaneously in producing the output Y_k. Let the windows be of length 2N+1 and 2L+1 where N>L. First the median m_k is computed from the small window of length 2L+1. For some positive number q an interval $[m_k-q, m_k+q]$ is defined. Then the mean of points lying within $[m_k-q, m_k+q]$ in the large window of length 2N+1 is computed as the output. Here q is a filter parameter which may be set a priori to some reasonable value depending on noise variance. When the statistics of the noise are not known, q may be estimated from the input in an adaptive scheme. This algorithm has good nonlinear characteristics because of the computation of the median from the small window, and also has sufficient linear characteristics because of the averaging of the noisy values in a large window. It can be seen that for some extreme cases the outputs of the algorithm can be the medians of the values in windows of length 2L+1, or can be means of the values in windows of length 2N+1. Therefore depending on the input data the performance varies between that of a median filter with window size 2L+1 and that of a running mean filter with window size 2N+1. Note that we observed similar properties for standard-type M-smoothers in Section III. However, the performance of a double-window smoothing algorithm is much better than that of M-smoothers because two windows are used and also because outliers are completely discarded.

It can be shown that a double-window smoothing scheme can be thought of as a modified data-dependent L-smoother. Suppose at some time k, a number γ of data values among the 2N+1 values in the large window are within the region $[m_k-q, m_k+q]$. Clearly averaging only over the γ data points is equivalent to weighting these γ values by $1/\gamma$ and assigning weights 0 to the remaining $2N+1-\gamma$ values. Note that the weights of L-smoothers are defined for the ordered sequence of data values inside a window (see Figure 1). Since the γ values in the region are consecutive in the ordered sequence of 2N+1 data values, the weighting consists of γ consecutive weights of $1/\gamma$ and 0 weights for the other portion of the ordered data. It should be noted that neither γ nor the segment of successive $1/\gamma$'s in the weights is fixed in location, but depends on the data. We have also observed the similarities between double-window smoothing and M-smoothers. Actually, if we use an M-smoother whose characteristic ϕ is a noise-blanker characteristic after finding the median from the small window, the resulting output will be very similar to that of a double-window smoothing.

Figure 2(g) shows the result of the double-window smoothing with 2L+1=3, 2N+1=7, and q=3 which is 3 times the standard deviation of the noise. Clearly the output is closest to the original signal, compared to the outputs of the other smoothing results. The edge characteristic is seen to be as good as that of the median filter (Figure 2(c)), and also high frequency noise is suppressed almost completely. When q is reduced to 2.75 the ripple before the last pulse disappears and the other parts of the output remain the same. If q is less than 2.5 some high frequency noise is not suppressed sufficiently. Therefore in this simulation the optimal q is around 2.75.

In summary, we observe that data-dependent double-window smoothing performs like a median filter with small window length as well as like a running mean with large window length, depending on the input, which can give the best result among the smoothers discussed so far. The performance may, however, be sensitive to the value chosen for q. It might be best to use a scheme for setting q which derives the value for q from measurements made on segments of the input data.

REFERENCES

1. J. W. Tukey, Exploratory Data Analysis, Addison-Wesley, Reading, 1977.
2. L. R. Rabiner, M. R. Sambur, and C. E. Schmidt, "Applications of a Non-linear Smoothing Algorithm to Speech Processing", IEEE Trans. Acoust., Speech, Sig., Proc., ASSP-23, pp. 552-557, 1975.
3. T. A. Nodes, N. C. Gallagher, "Some Results on the Median Filtering of Signals and Additive White Noise", Proc. 19th Allerton Con. on Comm., Control, Comp., pp. 99-108, 1981.
4. R. Kuhlmann, G. L. Wise, "Performance of Median Filters With Random Inputs", Proc. Internat. Conf. on Comm., Jun. 1982, Sect. 1H, 2.
5. T. A. Nodes, N. C. Gallagher, Jr., "Median Filters: Some Modifications and Their Properties", IEEE Trans. Acoust., Speech, Sig. Proc., ASSP-30, pp. 739-746, 1982.
6. C. L. Mallows, "Some Theoretical Results on Tukey's 3R Smoother", Proc. Heidelberg Workshop on Smoothing Tech. for Curve Est., pp. 77-90, 1979.

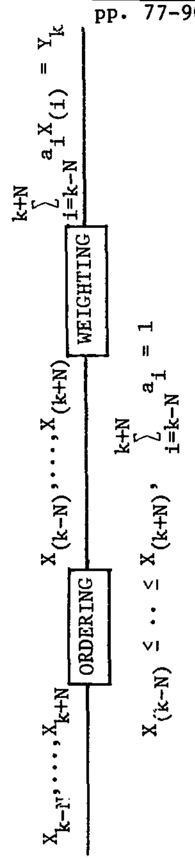


Figure 1. Definition of L-Smoothers.

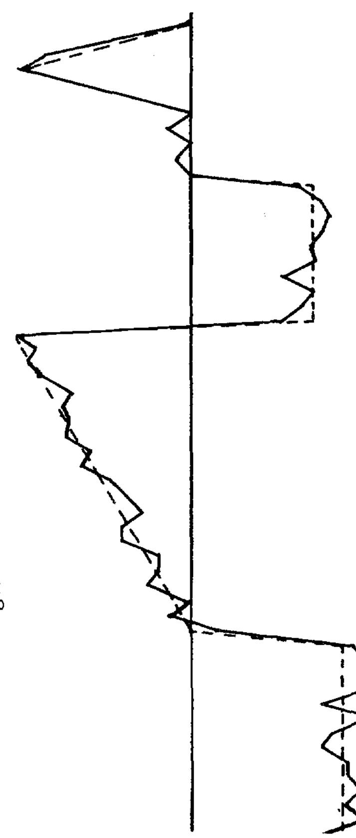
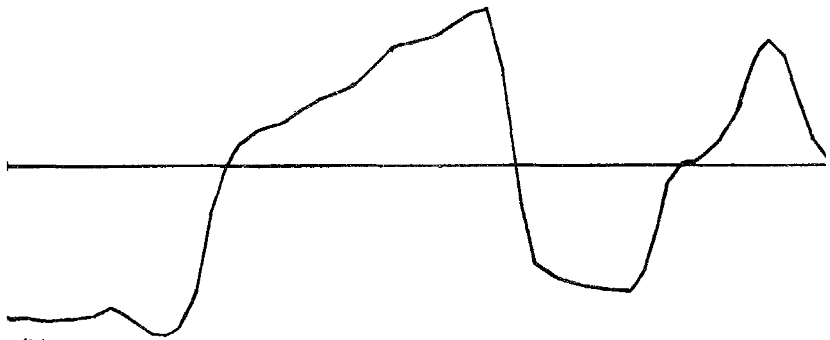
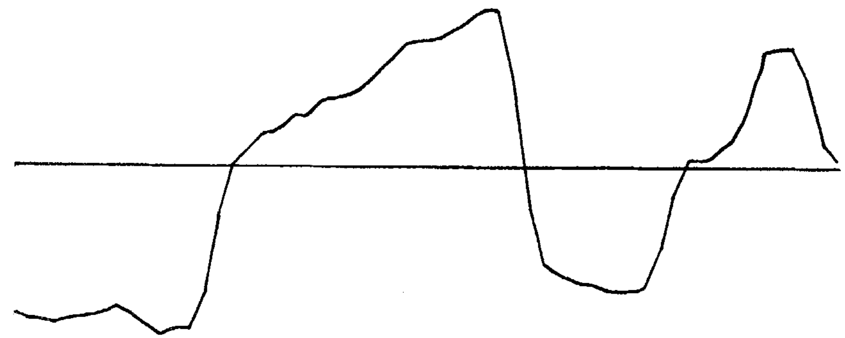


Figure 2. (a) Input; A Signal Plus Gaussian Additive Noise with Mean 0, Unit Variance.

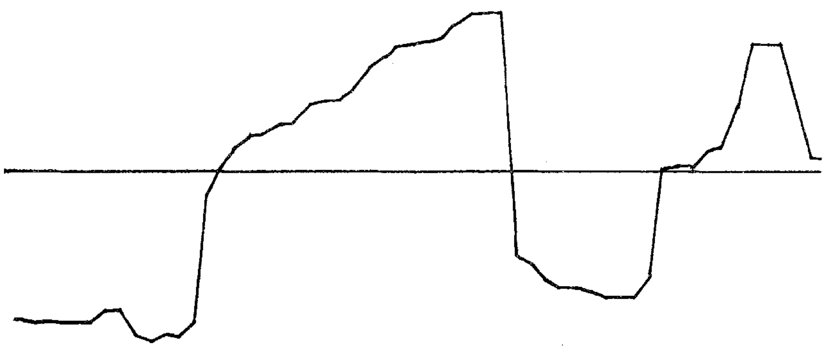
9.9



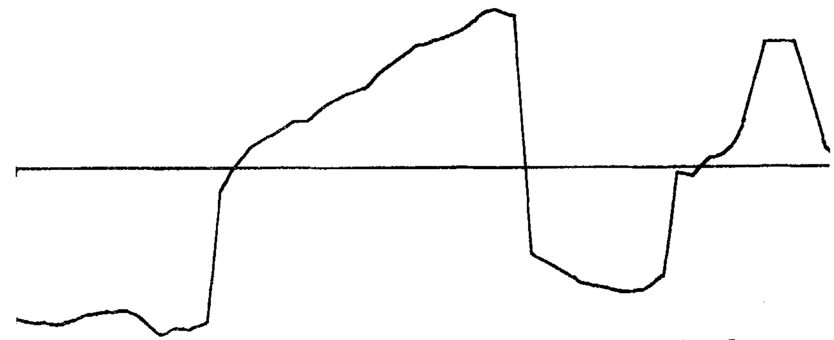
(b). Combination of 5-point Median and 3-point Hanning Window.



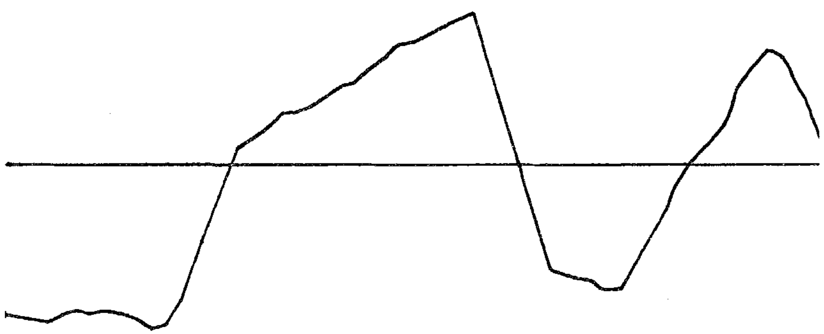
(e) L-Smoother with Weights (0, 1/4, 1/2, 1/4, 0).



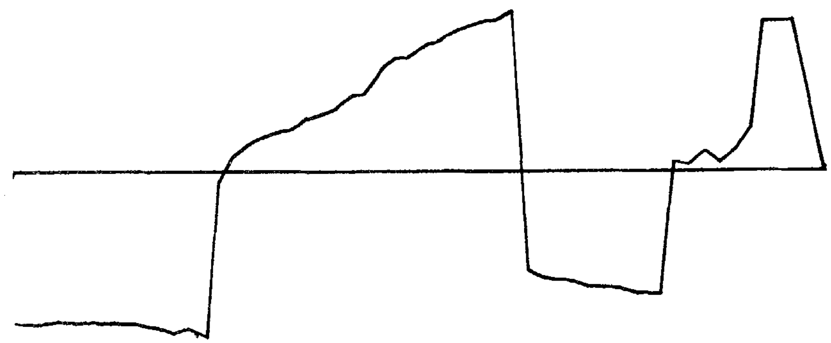
(c). Median Filter with Window Size 5.



(f). Standard-Type M-Smoother with p=1 and Window Size 5.



(d). 5-point Running Mean.



(g). Double Window Smoothing with 2L+1=3, 2N+1=7 and q=3.

Figure 2. (Continued) Smoothed Outputs