Order Statistic-Based Nonlinear Filters: Stack Filters and Weighted Median Filters

Edward J. Coyle [†], Moncef Gabbouj [‡], and Jisang Yoo [†]

[†] School of Electrical Engineering
Purdue University
West Lafayette, Indiana 47907, USA

[‡] Signal Processing Laboratory
Tampere University of Technology
P.O. Box 553
SF-33101 Tampere, FINLAND

Abstract

Since the introduction of the median filter by John Tukey in 1971, many important classes of order statistic-based nonlinear filters have been developed. In this paper we review some recent results obtained for the two filter classes known as stack filters and weighted median filters. The highlights include new results on optimal filter design and fast training algorithms.

1. Introduction

Linear filters have long been the primary tool for signal and image processing. They are easy to implement and analyze and, perhaps most importantly, the linear filter which minimizes the mean squared error criterion can usually be found in closed form. Furthermore, they are optimal among the class of all filtering operations when the noise being considered is additive and Gaussian.

Unfortunately, a small deviation from this Gaussian assumption sometimes leads to a severe deterioration in the performance of linear filters. In the many applications in which non-Gaussian noise arises, linear methods have thus proven to be inadequate for signal smoothing and/or noise reduction. One such case occurs in the presence of speckle noise. Other types of non-Gaussian and/or signal dependent noise also cause problems. We believe that these cases occur more frequently than not; therefore, linear methods are not completely satisfactory when dealing with real signals and noise rather than simply computer simulations.

The obvious answer to this problem is to use a filter that is not linear. There are, however, many classes of non-linear filters, and the task of choosing the right class is itself a challenge. Each class of filters is good in certain applications. The user could consult a look-up table to determine which filter or class of filters best fits the problem at hand.

One filter that would certainly appear in any such catalogue would be the median filter. The median filter, or "running median" as it was called in the first publication in which it appeared [Tuk], consists of a window, usually of odd width, which is stepped one sample at a time along a signal. At each position of the window, the sample values inside are ranked according to their magnitude and the middle element in this ranking is defined to be the output.

Typically, the window is assumed to have width $2N+1$ where N is any positive integer. Suppose that the window is centered on the k'th sample in the input sequence and that the $2N+1$ points in the window, in time-order, are specified by the vector

$$\bar{x}_{N,k} = (x_{k-N}, x_{k-N+1}, \cdots, x_k, \cdots, x_{k+N}).$$

We want to find y_k, where

$$y_k = med(x_{k-N}, \cdots, x_{k+N}) = med(\bar{x}_{N,k})$$

which is the output of the median filter when the window is centered on the k'th sample.

First, the samples in the window are reordered according to their rank, with $x_{(i)}$ denoting the sample of i^{th} rank. The samples in the window, in rank-order, would then be

$$(x_{(1)}, x_{(2)}, \cdots, x_{(2N+1)}).$$

Suppose, for instance, that N=2 and that the samples, in time order, in the window at time k are

$$(x_{k-2}, x_{k-1}, x_k, x_{k+1}, x_{k+2}) = (8,1,6,4,1);$$

in rank-order they would be

$$(x_{(1)}, x_{(2)}, x_{(3)}, x_{(4)}, x_{(5)}) = (1,1,4,6,8).$$

The median value of $2N+1$ samples is given by $x_{(N+1)}$, which for the example just given would be: $x_{(3)} = 4$. We thus have $y_k = 4$. The window is then moved so it is centered over the k+1'st sample in the input sequence and the output value y_{k+1} is computed by following the above procedure.

Figure 1 shows the input and output for a window width three median filter, and also shows how the window moves along the signal. Note from the figure that the median filter can preserve edges while deleting impulses.

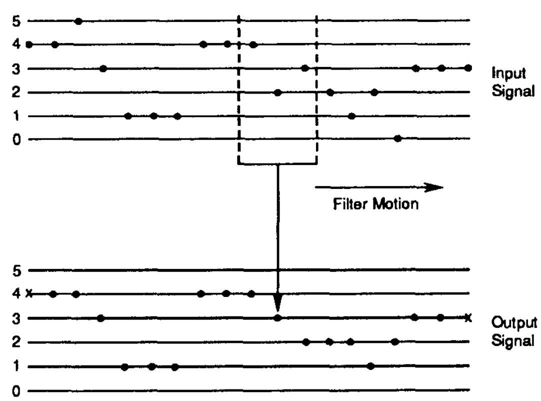

Figure 1: The window width 3 median filter.

Three questions arise from this example. Why does the median filter preserve edges and eliminate impulses? Are there other filters which have similar behavior? Can a design theory be developed for a class of filters related to the median filter?

For the last two decades, these questions have attracted the attention of many researchers. The result of this attention is a large number of papers, dissertations and books written on the median and other order statistic-based filters.

Thorough reviews of the development of these filters can be found in [CLG][GCG]. An overview of the field of nonlinear filters can be found in [PiV1]. In this brief conference paper we only have space to review two classes of order statistic-based nonlinear filters: stack filters and weighted median filters.

2. Stack Filters

In 1984, an important theoretical tool for analyzing median and median-type filters was developed by Fitch et al, [FCG1][FCG2]. It was shown that all rank-order-based filters possess a weak superposition property called the *threshold decomposition* property. It says that median filtering any sequence whose elements take on values in the set $Q = \{0, 1, \cdots, M-1\}$ is equivalent to decomposing the signal into binary sequences by thresholding at each level from 1 through $M-1$, filtering each resulting binary sequence by a (binary) median filter, and then adding up the results.

To express this property more precisely, define $T_i[\cdot]$ to be the operator which thresholds its argument at level i:

$$T_i(x) = \begin{cases} 0 & \text{if } x < i \\ 1 & \text{if } x \geq i \end{cases}$$

For instance, $T_3(5) = 1$ and $T_3(2) = 0$. We will also apply thresholding to vectors, and will define it as follows, even though it means a slight abuse of notation

$$T_i\Big[(x_1, \cdots, x_n)\Big] = \Big[T_i(x_1), \cdots, T_i(x_n)\Big]$$

For example, we would have

$$T_3\Big[(6,2,1,3,7)\Big] = (1,0,0,1,1).$$

Suppose we are to apply a median filter of window width 3 to a sequence $\{x_k\}$ that takes on values in Q. The threshold decomposition states that for every k

$$Med(x_{k-1}, x_k, x_{k+1}) = Med\left[\sum_{i=1}^{M-1} T_i\Big[(x_{k-1}, x_k, x_{k+1})\Big]\right]$$
$$= \sum_{i=1}^{M-1} Med\Big[T_i[(x_{k-1}, x_k, x_{k+1})]\Big].$$

This weak superposition property, which is illustrated in Figure 2 with a median filter of window width 3, allows the analysis of the effects of the median on multi-valued sequences to be reduced to the study of its effects on binary sequences.

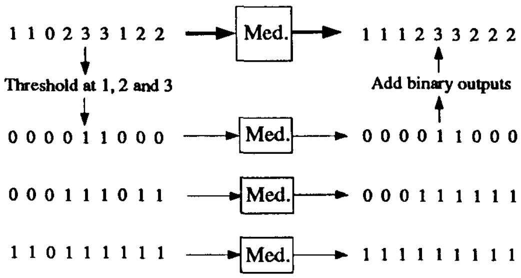

Figure 2: Threshold Decomposition

Note that the median filter on each threshold level in Figure 2 has binary data at its input and produces binary data at its output. A median filter with binary input is thus a binary operator which we denote as $f(\cdot)$. In the case of the window width three median, $f(\cdot)$ is simply a majority logic operation. If the three bits in the window at time k are b_{k-1}, b_k, and b_{k+1}, then the output is

$$f(b_{k-1}, b_k, b_{k+1}) = b_{k-1}*b_k + b_k*b_{k+1} + b_{k+1}*b_{k-1}$$

where "*" denotes the logical *and* operation, and "+" denotes the logical *or* operation.

Another crucial property of the median filter is the ordering property known as the stacking property. For any two vectors $\bar{x} = (x_1, x_2, \cdots, x_n)$ $\bar{y} = (y_1, y_2, \cdots, y_n)$ of length n, we say that

$$\bar{x} \geq \bar{y}$$

if and only if $x_i \geq y_i$ for all values of i. A logical operator $f(\cdot)$ is said to obey the stacking property if and only if

$$f(\bar{x}) \geq f(\bar{y}) \quad whenever \quad \bar{x} \geq \bar{y}.$$

The threshold decomposition architecture shown in Figure 2 and the stacking property are the defining properties of stack filters. Any filter $S_f(\cdot)$ is a stack filter if it has the superposition architecture of Figure 3 and if the Boolean function $f(\cdot)$ used on every threshold level of the architecture has the stacking property.

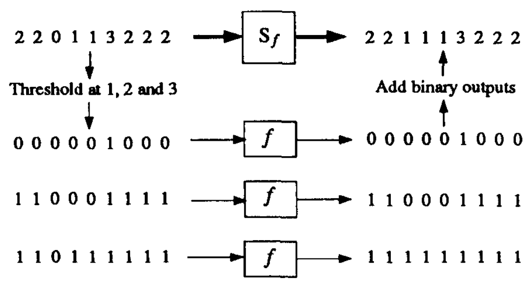

Figure 3: A stack filter: an asymmetric median.

$$f(b_1, b_2, b_3) = b_1*b_3 + b_2$$
$$S_f(x_1, x_2, x_3) = max(min(x_1, x_3), x_2)$$

Note that the AND and OR operations are simply the binary version of the *max* and *min* operations [NaR]. Because every rank order operator, including the *max* and *min*, possesses the threshold decomposition property [FCG2], each stack filter also obeys the threshold decomposition property and each can be expressed as a max of mins operation [MaS] as illustrated in the caption of Figure 3.

Mathematically, the following superposition property holds for the window width $2N+1$ stack filter $S_f(\cdot)$ based on the positive Boolean function $f(\cdot)$ of window width $2N+1$:

$$\begin{aligned} S_f(x_{k-N}, \cdots, x_{k+N}) &= S_f\left\{\sum_{i=1}^{M-1} T_i((x_{k-N}, \cdots, x_{k+N}))\right\} \\ &= \sum_{i=1}^{M-1} S_f\left[T_i((x_{k-N}, \cdots, x_{k+N}))\right] \\ &= \sum_{i=1}^{M-1} f\left[T_i((x_{k-N}, \cdots, x_{k+N}))\right]. \end{aligned}$$

The importance of the threshold decomposition and stacking properties of stack filters follows from an application to estimation.

Suppose that the input sequence $\{x_k\}$ for a stack filter is a signal sequence $\{s_k\}$ that has been distorted in some fashion by a noise sequence $\{n_k\}$. We will assume that the desired signal sequence and the noise sequence are jointly stationary and that all sequences take on values in $Q = \{0, 1, 2, \cdots, M-1\}$. The goal is to choose a stack filter of window width $2N+1$ such that, for input sequence $\{x_k\}$, the output sequence $\{y_k\}$ produced by the stack filter is "close" to $\{s_k\}$.

Our measure of closeness will be the mean absolute error

$$MAE_f = E_f\left[\ |s_k - y_k|\ \right]$$

where the choice of k is arbitrary because of the stationarity assumption made above and where the stack filter $S_f(\cdot)$ is being used. Our goal is to find the positive Boolean function $f(\cdot)$ which minimizes MAE_f.

The threshold decomposition and stacking properties allow us to significantly simplify MAE_f:

$$MAE_f = E_f\left[\ |s_k - y_k|\ \right]$$

$$= E_f\left[\ |s_k - S_f(y_{k-N}, \cdots, y_{k+N})|\ \right]$$

Applying the threshold operations and the threshold decomposition property yields:

$$E_f\left[\ |\sum_{i=1}^{M-1} T_i\left(s_k\right) - \sum_{i=1}^{M-1} S_f\left(T_i(y_{k-N}), \cdots, T_i(y_{k+N})\right)|\ \right]$$

Now apply the stacking property:

$$E_f\left[\sum_{i=1}^{M-1} |\ T_i\left(s_k\right) - S_f\left(T_i(y_{k-N}), \cdots, T_i(y_{k+N})\right)|\ \right]$$

$$E_f\left[\sum_{i=1}^{M-1} |\ T_i\left(s_k\right) - f\left(T_i(y_{k-N}), \cdots, T_i(y_{k+N})\right)|\ \right]$$

In this last equation we see that the mean absolute error is simply the sum of the mean absolute errors made on each level of the threshold decomposition architecture.

With the simplification shown above, the problem of minimizing MAE_f can be reduced to a linear program [CoL]. This linear program is difficult to solve, though, for window sizes greater than 10. A stack filter training algorithm [LSC] was thus developed and has been used to design stack filters with windows containing up to 25 points.

In Figure 4, we show an image of Einstein. Figure 5 is the same image after it has been corrupted with a combination of Gaussian noise and salt and pepper noise.

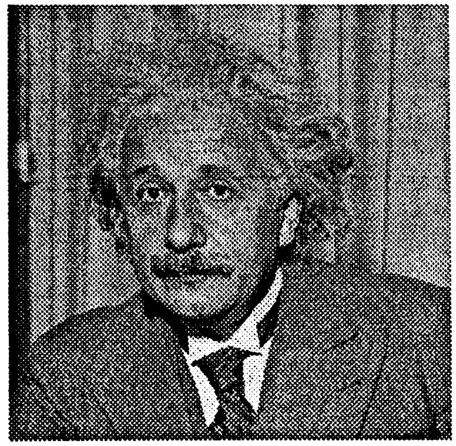

Figure 4: *Original noise-free image.*

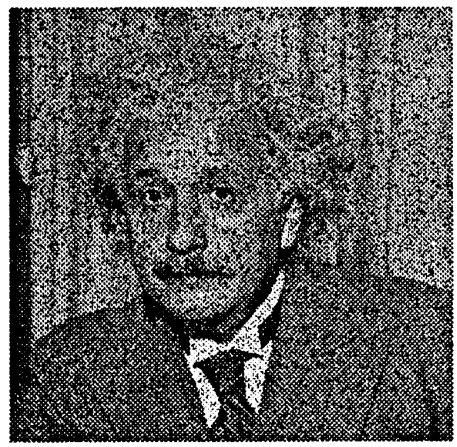

Figure 5: *Noisy image.*

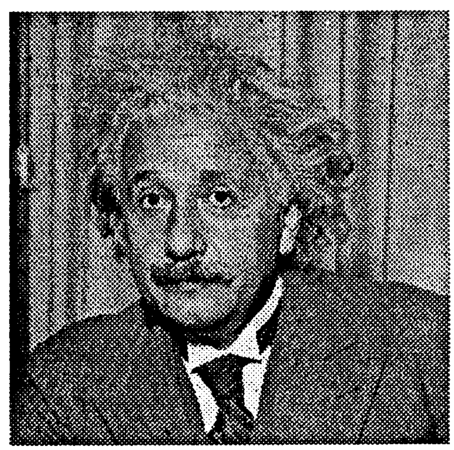

Figure 6: *Ouput of* 4×4 *stack filter.*

A 4×4 stack filter was trained to restore this image. The results obtained are shown in Figure 6. The detail in the image has been retained while the noise has been well suppressed.

One of the most recent applications of stack filters has been to the problem of edge detection in images corrupted with noise. The use of stack filters has led to a new approach to edge detection [YBD].

Consider a noise-free, one-dimensional signal $\{x(n)\}$. We would like to estimate the magnitude of the first difference, $|s(n+1)-s(n)|$, for this signal since the result can then be thresholded to determine if an edge is located at position n. We will estimate the first difference by applying a spatially invariant nonlinear filter $F(\cdot)$ to the noise corrupted signal, $\tilde{x}(n)$.

First define the windowing operation

$$W_{\tilde{x}}(n) = \left[\tilde{x}(n-k), \tilde{x}(n-k+1), \cdots, \tilde{x}(n+k)\right]$$

The estimated first difference is then $F(W_{\tilde{x}}(n))$ where $F(\cdot)$ is the nonlinear function to be determined. Our goal is then to determine the function $F(\cdot)$ which minimizes

$$e(F) = E\left[\ | F(W_{\tilde{x}}(n)) - |x(n+1)-x(n)|\ |\ \right]$$

To determine the structure of F, we first rewrite the magnitude of the first difference by noting that

$$|x(n+1)-x(n)| =$$

$$\max\{x(n+1),x(n)\} - \min\{x(n+1),x(n)\}$$

The magnitude of the first difference is then seen to be the difference between the windowed maximum and minimum, which are equivalent to the well-known dilation and erosion operators in morphology. This suggests that we find two nonlinear filters which estimate the dilated and eroded signals, and then form the difference between these estimates. This difference is then thresholded to produce an estimate of the edge map of the target signal. This approach to edge detection is called the *Difference of Estimates* (DoE) approach.

We thus write

$$F(W_{\tilde{x}}(n)) = S_d(W_{\tilde{x}}(n)) - S_e(W_{\tilde{x}}(n))$$

where $S_d(\cdot)$ and $S_e(\cdot)$ are the stack filters used to estimate the noise free dilation and erosion, respectively. The error measure can then be upper bounded by the sum of the following two error terms

$$E\left[\ | S_d(W_{\tilde{x}}(n)) - \max(x(n),x(n+1))\ |\ \right]$$

$$E\left[\ | S_e(W_{\tilde{x}}(n)) - \min(x(n),x(n+1))\ |\ \right]$$

Our strategy is to minimize the overall error by choosing $S_d(\cdot)$ and $S_e(\cdot)$ to separately minimize each of the above two error expressions.

We train stack filters to minimize these two error expressions. The training is carried out on some image that is "similar" to the target image since the noise-free target image is rarely available. The image we used for training is shown in Figure 7. The image was corrupted with light noise that was a mixture of Gaussian noise and salt and pepper noise.

Figure 7: Training image.

The resulting stack filters were then used to estimate the noise-free dilated and eroded version of the image in Figure 8 which is lightly corrupted by a mixture of Gaussian noise and salt and pepper noise. The difference of the outputs of these filtering operations was then thresholded to obtain the edge map. In Figures 9 through 12 we show the edge map obtained by the DOE operator trained on the couple image with those obtained by a 7×7 Canny operator [Can], the improved morphological operator developed in [LHS], and by the dispersion operator [PiV2].

Figure 8: Noisy image.

Figure 9: DoE edge map.

Figure 10: 7×7 Canny edge map.

Figure 11: Morphological edge operator.

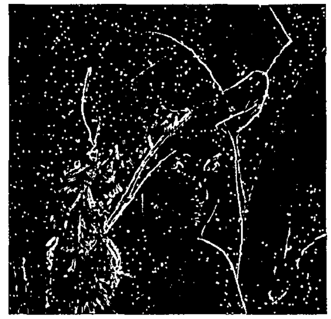

Figure 12: Dispersion edge operator.

From these images it is clear that the edge map produced by the DoE operator is significantly less corrupted by noise. The DoE operator has also retained more of the fine edges than the other edge detection operators. Recall that the DoE operator was trained with a different image and different noise sample functions. The robustness that stack filters bring to any application is thus clearly evident in this example.

3. Weighted Median Filters

As a major subclass of stack filters and an immediate generalization of the median filter, weighted median filters have attracted the attention of many researchers. Several papers have recently dealt with the deterministic and statistical properties of these filters [YAN2], [KoL], [PrL1], [PrL2], [HGN], [YAN1], [SGN1], [SGN2]. The most important result obtained for WM filters is perhaps the optimality theory built around these filters. It combines the deterministic and statistical properties of WM filters [YYG2].

The aim of this optimality theory is to select a WM filter which has the best noise attenuation capability (under a certain error criterion) and which, at the same time, preserves a set of pre-specified signal structures. In the literature, this is referred to as optimal filtering under structural constraints. A similar theory has been developed for stack filters [GHB], [CLG] and weighted order statistics filters [NYG].

Recall that a weighted median filter can preserve details, such as pulses of any desired length, by selecting appropriate weights. On the other hand, for a given pulse length, there may be many weighted median filters which can preserve the given pulse. The question is how to select one of them which suppresses noise the best. A solution is based on the statistical analysis of weighted medians developed in recent work [YYG2], [YYG1].

Consider a WM filter with weight vector $\underline{W} = (W_1, W_2, \ldots, W_N)$, N=2K+1. Denote by *scriptS* the multiset of weights, i.e. *scriptS* $= \{W_1, W_2, \ldots, W_N\}$. A parameter M_i, i=1, ..., K is introduced to denote the cardinality of the following set for $i = 0, 1, \cdots, N$

$$\Omega^{[i]} = \{A \mid A \subset W\ ,\ \mid A \mid = i,\ and \sum_{W_j \in A} W_j \geq T\}$$

(such subsets are called *positive subsets),* i.e.,

$$M_i = \mid \Omega^{[i]} \mid,\ i = 0,\ 1,\ ...,\ N.$$

Using these M_i's, the output moments of weighted medians are given by the following theorem.

Theorem 1: Consider a weighted median filter with window width N=2K+1, weights $\underline{W}$, and i.i.d. input from a common distribution function $\Phi(t)$ and density function $\phi(t)$. The l-order output moment of the weighted median filter can be expressed as

$$\mu_{wm}^{l} = \mu_s^{l} + \sum_{i=1}^{K} M_i L_i(N, \Phi, l),$$

where μ_s^l is the l-order output moment of the standard median with the same window width and

$$L_i(N, \Phi, l) = \int_{-\infty}^{+\infty} V_i(\Phi(y)) \mid y \mid^{l} \phi(y) dy \geq 0,$$

for $l \geq 0$ and i=1,..., K and

$$V_i(\Phi) = \frac{d}{d\Phi} (\Phi^i (1 - \Phi)^{N-i} + \Phi^{N-i} (1 - \Phi)^i).$$

Theorem 1 says that each output moment of WM filters is composed of two terms. The first term is the output moment of the standard median filter with the same window size, and has nothing to do with the weights. The second term is a function of the M_i's and is thus a function the weights; it can be referred to as the contribution of the weights. Note that this second term is equal to zero if all the weights are equal. According to the definition of the M_i's, this second term is always positive, i.e.

$$\sum_{i=1}^{K} M_i L_i \geq 0.$$

This means $\mu_{wm}^{l} \geq \mu_s{}^{l}$ for $l \geq 0$.

One application of Theorem 1 is the efficient evaluation of the noise attenuation performance of weighted medians [YGN]. It would be more interesting though if Theorem 1 could be used to design optimal WM filters with structural constraints. It has been shown that given a set of structural constraints, an optimal WM filter can be found through nonlinear programming. In some cases, the optimality problem is reduced to the solution of a group of linear inequalities.

Consider a one-dimensional signal processing application. If pulses with length 2 are supposed to be preserved, then optimal symmetric WM filters for arbitrary window size N are given by the following theorem.

Theorem 2: Given a WM filter with window size N=2K+1,

$$\underline{W} = (W_{-K}, W_{-K+1}, ..., W_0, ..., W_{K-1}, W_K),$$

optimal WM filters which preserve pulses of length 2 can be expressed as follows:

$$\underline{W} = (1,\ 1,\ ...,\ 1,\ a{+}1,\ 2a{+}1,\ a{+}1,\ 1,\ ...,\ 1,\ 1),$$

where $a \geq K - 1$.

4. Center Weighted Median Filters

Center weighted median (CWM) filters are a subclass of WM filters which combines the simplicity of median filters with some of the design freedom of WM filters. In CWM filters only the center sample in the window has a weight larger than or equal to one; all other weights are equal to one. They have simpler implementations than more general weighted medians since the output is always the median of three samples from the filter window [KoL].

Because of the nonlinearity of median-based filters, deterministic properties involving root set structures and convergence behavior have been analyzed to represent the filtering behavior. The deterministic properties of one-dimensional CWM filters have been investigated. To be specific, CWM filters have been shown to have the convergence property and their root structures have been specified [SGN1], [SGN2]. It has been shown that an arbitrary finite length signal will be reduced to a root signal after a finite number of passes of any CWM filter. A necessary and sufficient condition for an arbitrary signal to be a root signal has been derived. Roots of CWM filters include constant neighborhoods, edges and some other irregular structures not found in roots of the median filter. It has also been shown that the root signal sets of CWM filters with the same window size and different center weights are nested.

Analysis of the root structures of CWM filters has also been extended to the two-dimensional case [Sun93a]. The required form for a signal to be a root of a 2-D CWM filter has been derived in [Sun93a]. These results can be used to evaluate the detail-preserving properties of these filters.

Time invariant filtering is not appropriate for nonstationary images due to the compromise between noise reduction and resolution. Therefore, time-variant filters may show their advantage in this case. Based on the simple structure of CWM filters, adaptive CWM filters based on the local statistics have been designed. In particular, we have proposed a cascade of several one-dimensional adaptive CWM filters oriented in different important correlation directions of the image for image restoration [SGN2][SGN3]. This filtering scheme is better than the corresponding two-dimensional adaptive CWM filters in many cases. Each pixel is processed independently, so a parallel processor can be used in real-time signal processing applications. The most important advantage of one-dimensional adaptive CWM filters is their ability to remove noise along image details while preserving these details.

5. References

[Can] J. Canny, "A computational approach to edge detection," *IEEE Trans. on Pattern Anal. and Machine Intell,"* vol. PAMI-8, pp 679-698, Nov. 1986.

[CLG] E.J. Coyle, J.-H. Lin, and M. Gabbouj, "Optimal stack filtering and the estimation and structural approaches to image processing," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-37, Dec. 1989, pp. 2037-2066.

[CoL] E.J. Coyle and J.-H. Lin, "Stack filters and the mean absolute error criterion," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. ASSP-36, no. 8, pp. 1244-1254, August 1988.

[FCG1] J.P. Fitch, E.J. Coyle and N.C. Gallagher, Jr., "Median filtering by threshold decomposition," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-32, no. 6, pp. 1183-1188, December 1984.

[FCG2] ____, "Threshold decomposition of multidimensional ranked order operations," *IEEE Trans. on Circuits and Systems*, vol. CAS-32, no. 5, pp. 445-450, May 1985.

[GCG] M. Gabbouj, E.J. Coyle, and N.C. Gallagher, "An overview of median and stack filtering," *Circuits, Systems, and Signal Processing*,

[GHB] M. Gabbouj, P. Haavisto, and Y. Neuvo, "Recent Advances in Median Filtering," (invited), *Proceedings of the 1990 Bilkent Conf. on New Trends in Communication, Control, and Signal Processing*, Bilkent University, Ankara, Turkey, July 2-5, 1990, Vol II, E. Arikan, Ed., Elsevier Science Publishers B.V., 1990, pp. 1080-1094.

[HGN] P. Haavisto, M. Gabbouj, and Y. Neuvo, "Median based idempotent filters," *Journal of Circuits, Systems, and Computers*, Vol.1, no.2, pp.125-148, June 1991

[KoL] S.-J. Ko, and Y.H. Lee, "Center weighted median filters and their application to image enhencement," *IEEE Trans. on Circuits and Systems*, vol.38, no.9, September 1991.

[LHS] J.S.J. Lee, R.M. Haralick, and L.G. Shapiro, "Morphological edge detection," *IEEE Trans. on Robotics and Automation*, vol. RA-3, pp. 142-156, April 1987.

[LSC] J.-H. Lin, T.M. Sellke, and E.J. Coyle, "Adaptive stack filtering under the mean absolute error criterion," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-38, no. 6, pp. 938-954, June 1990.

[NaR] Y. Nakagawa and A. Rosenfeld, "A note on the use of local min and max operations in digital picture processing," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. SMC-8, no. 8, pp. 632-635, August 1978.

[MaS] P.A. Maragos and R.W. Schafer, "Morphological filters - Part I: their set theoretic analysis and relations to linear shift invariant filters,", and "Morphological filters - Part II: their relations to median, order statistic and stack filters," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-35, no. 8, pp. 1153-1184, August 1987.

[NYG] J. Nieweglowski, L. Yin, M. Gabbouj, and Y. Neuvo, "Optimal Weighted Order Statistic Filters under Structural Constraint s," *1992 IEEE Proc. International Symposium on Circuits and Systems*, Special Session on Nonlinear Filtering, May 10-13, 1992, Sheraton Harbor Island, San Diego, USA, pp. 2621-2624.

[PiV1] I. Pitas and A.N. Venetsanopoulos, *Nonlinear Digital Filters: Principles and Applications*, Boston: Kluwer Academic Publishers, 1990.

[PiV2] I. Pitas and A.N. Venetsanopoulos, "Edge detectors based on nonlinear filters," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. PAMI-8, No. 4, July 1986.

[PrL1] M.K. Prasad, and Y.H. Lee, "Weighted median filters: generation and properties," *Proc. IEEE Int. Symp. Circuits and Systems, ISCAS-89*, pp.425-428, 1989.

[PrL2] M.K. Prasad, and Y.H. Lee, "Analysis of weighted median filters based on inequalities relating the weights," *Circuits, Systems and Signal Processing*, vol.11, no.1, 1992.

[SGN1] T. Sun, M. Gabbouj, and Y. Neuvo, "Deterministic properties of center weighted median filters", *Proceedings of 1992 International Conference on Communication Technology*, Beijing, china, Sep. 16-18, pp. 06.09.1-06.09.4, 1992.

[SGN2] T. Sun, M. Gabbouj and Y. Neuvo, "Center weighted median filters: some properties and application in image processing", Submitted to *Signal processing*, June. 1992.

[SGN3] T. Sun, M. Gabbouj, and Y. Neuvo, "Image restoration by one-dimensional adaptive center weighted median filters," submitted to *ECCTD'93* Switzerland, August 1993.

[Tuk] J.W. Tukey, *Exploratory Data Analysis*, Addison-Wesley: Reading, MA, 1977.

[YYG1] R. Yang, L. Yin, M. Gabbouj, and Y. Neuvo, "2-D Optimal weighted Median Filters," *Proc. of 1992 IEEE Workshop on Visual Signal Processing and Communications*, North Carolina State University, Raleigh, North Carolina, September 2 -3, 1992, pp. 31-35.

[YYG2] R. Yang *et al.*, "Optimal weighted median filters under structural constraints," submitted to *IEEE Trans. on Signal Processing*, Dec. 1992.

[YGN] R. Yang, M. Gabbouj, and Y. Neuvo, "The output moments of weighted median filters," accepted by *J.T.E.A.'93*, Feb 5-7, 1993, Hammamet, Tunisia.

[YBD] J. Yoo, C.A. Bouman, E.J. Delp, and E.J. Coyle, "The nonlinear prefiltering and difference of estimates approaches to edge detection: applications of stack filters," *Graphical Models and Image Processing*, to appear.

[YAN1] L. Yin, J. Astola, and Y. Neuvo, "Optimal weighted order statistic filters under the mean absolute error criterion," *Proc. of Int. Conference on Acoustics, Speech, and Signal Processing*, vol.4, pp.2529-2532, Canada, May 1991.

[YAN2] O. Yli-Harja, J. Astola, and Y. Neuvo, "Analysis of the properties of median and weighted median filters using threshold logic and stack filter representation," *IEEE on Acoust. Speech and Signal Processing*, vol.39, no.2, pp-395-410, Feb. 1991.