Optimal Weighted Median Filtering Under Structural Constraints

Ruikang Yang, *Student Member, IEEE*, Lin Yin, Moncef Gabbouj, *Member, IEEE*, Jaakko Astola, and Yrjö Neuvo, *Fellow, IEEE*

***Abstract*— A new expression for the output moments of weighted median filtered data is derived in this paper. The noise attenuation capability of a weighted median filter can now be assessed using the L-vector and M-vector parameters in the new expression.**

The second major contribution of the paper is the development of a new optimality theory for weighted median filters. This theory is based on the new expression for the output moments, and combines the noise attenuation and some structural constraints on the filter's behavior.

In certain special cases, the optimal weighted median filter can be obtained by merely solving a set of linear inequalities. This leads in some cases to closed form solutions for optimal weighted median filters. Some applications of the theory developed in this paper, in 1-D signal processing and image processing are discussed.

Throughout the analysis, some striking similarities are pointed out between linear FIR filters and weighted median filters.

I. Introduction

THE weighted median (WM) filter is a natural extension of the median filter and has the same advantages as the median filter: edge preservation and efficient suppression of impulsive noise. Having a set of weights, however, the WM filter is much more flexible in preserving desired signal structures than a median filter. This is explained as follows. A window width $2K+1$ median filter can only preserve details lasting more than $K+1$ points. To preserve smaller details in signals, a smaller window width median filter must be used. Unfortunately, the smaller the filter window is, the poorer its noise reduction capability becomes [1]. Weighted medians present an elegant solution to such a dilemma. That is, one can select to use a weighted median filter with long enough window, say $2K+1$, to suppress noise effectively and at the same time preserve details lasting less than $K+1$ points. This property of weighted median filters has been the major reason for their successful applications in speech, image, and image sequence processing [2]–[12].

Weighted median filters were introduced by Justusson [1] and further studied by Brownrigg [8], [13]. Using the threshold decomposition, Wendt *et al.* showed that WM filters belong to the class of stack filter and can be represented by threshold functions (also called neurons in neural networks [14]) in the binary domain [15]. Based on the threshold function representation, Yli-Harja *et al.* [16] analyzed some deterministic and statistical properties of WM filters; and Wendt and Haavisto *et al.* studied the root signal properties of WM filters [10, 17]. Without reference to the threshold decomposition, Prasad *et al.* derived a number of deterministic properties of WM filters [18], [19], which, e.g., show the equivalence between two WM filters. Recently, several adaptive algorithms have been developed for finding optimal WM and weighted order statistic filters under the MSE and the MAE criteria, [5], [12], [20], [21]. These algorithms were motivated by the training algorithms used in neural networks and therefore require training sequences of both the desired and corrupted signals.

Although a remarkable development has been achieved in the understanding of the behavior of WM filters, the theory is still far from being mature. One of the many open problems in this area is the lack of an intuitive and efficient procedure for evaluating the noise reduction of WM filters. The first attempt to quantize the noise reduction of weighted medians was made by Yli-Harja *et al.* [16] but fell short of being intuitive or efficient. Their result was derived based on the positive Boolean function representation of WM filters in the binary domain. As a result, all connections to the median operation and the weights of the filter disappear.

In this paper, we propose a new efficient and intuitive procedure to analyze the statistical behavior of weighted median filters. Explicit expressions are derived to quantize the noise reduction capability of weighted median filters as a function of the weights of the filter and without resorting to the stack filter representation of weighted median filters. These expressions consist of two parts: one depends on the input distribution and the filter window width only, and the other depends on the weights of the weighted median filter. The latter can be obtained using tables, included in the paper; while the former represents the contribution of an equivalently sized standard median filter, which can easily be evaluated using a standard procedure, [1].

Another open problem in weighted median filtering is how to design optimal weighted median filters which have "best" noise reduction capability and at the same time preserve desired signal structures. This is referred to as optimal filtering under structural constraints, in the literature. The problem has been solved for the general class of stack filters [22] and weighted order statistic filters [23]. However, none of

Manuscript received December 15, 1992; revised June 13, 1994. This work was supported by the European Communities Program ESPRIT BRA III, Project NAT 7130. The associate editor coordinating the review of this paper and approving it for publication was Prof. Gonzalo Arce.

R. Yang is with the Nokia Research Center, Tampere, Finland.

M. Gabbouj and J. Astola are with the Signal Processing Laboratory, Tampere University of Technology, Tampere, Finland.

L. Yin and Y. Neuvo are with Nokia Mobile Phones R&D Center, Nokia Corporation, Helsinki, Finland.

IEEE Log Number 9408229.

the above algorithms may be used to easily generate optimal weighted median filters under a given simple set of structural constraints, e.g., pulses of given widths.

Recall that a weighted median filter can preserve details, e.g., 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. Based on the new expression for the output moments of weighted medians presented in this paper, a new optimality theory for weighted median filters is developed. This theory combines the noise attenuation and some structural constraints on the filter's behavior. It is shown that in some cases the optimal weighted median filter can be obtained by merely solving a set of linear inequalities. This leads to closed form solutions for the optimal WM filter. Applications in 1-D signal processing and image processing are presented in the paper. It is found that the noise attenuation behavior of optimal WM filters which preserve the same length signal detail improves as the filter window width increases, as expected. More interestingly, it is found that there exists a subclass of optimal WM filters which is independent of the underlying noise distribution and the given error criterion.

This paper is organized as follows. In Section II, we briefly review the basic concepts of weighted median filters. The statistical behavior of weighted median filters is studied in Section III based on a partition of the space of positive subsets (to be defined later) of weights. The partition is based on the cardinality of the subsets, called M_i for the ith subspace of weight subsets, each containing i weights. Analytical expressions of the output distribution and output moments of WM filters are derived using these M_i's. The theory of optimal WM filtering under structural constraints is developed in Section IV. Based on this theory, applications in 1-D signal processing and image processing are addressed in Section V. Section VI contains some conclusions.

II. Weighted Median Filters

The WM filter can be defined in two different but equivalent ways. The most commonly used one assumes positive integer weights with odd sum.

Definition 1: For the discrete-time continuous-valued input vector $\underline{X}$,

$$\underline{X} = (X_1, X_2, \cdots, X_N)$$

the output Y of the WM filter of span N associated with the integer weights

$$\underline{W} = (W_1, W_2, \cdots, W_N)$$

is given by

$$Y = \mathrm{MED}\{W_1 \diamond X_1, W_2 \diamond X_2, \cdots, W_N \diamond X_N\} \quad (1)$$

where MED$\{\cdot\}$ denotes the median operation and $\diamond$ denotes duplication, i.e.

$$n \diamond X = \underbrace{X, \cdots, X}_{n\ times}.$$

The filtering procedure goes as follows: The samples inside the filter window are duplicated to their corresponding weight, and the median of the expanded list is selected.

The second definition of the weighted median allows positive noninteger weights to be used.

Definition 2: The weighted median with positive set of weights W of a discrete-time continuous-valued input $\underline{X}$ is the value β minimizing the expression

$$\Phi(\beta) = \sum_{i=1}^{N} W_i |X_i - \beta|. \quad (2)$$

Here, β is guaranteed to be one of the samples X_i because $\Phi(\cdot)$ is piecewise linear and convex if $W_i \geq 0$ for all i.

The output of the WM filter with real positive weights can be calculated as follows: sort the samples inside the filter window (in ascending order), add up the corresponding weights starting from the upper end of the sorted list until the partial sum $\geq \frac{1}{2}\sum_{i=1}^{N} W_i$, the output of the WM filter is the sample corresponding to last weight added.

Since the median is a threshold operation, weighted medians can be defined in general, for all positive weights, as follows.

Definition 3: The output Y of a weighted median filter with positive weights W is given by

$$\begin{aligned} Y = &\ T\text{th largest value of the set} \\ &\{W_1 \diamond X_1, W_2 \diamond X_2, \cdots, W_N \diamond X_N\} \end{aligned} \quad (3)$$

where the threshold

$$T = \begin{cases} \frac{1}{2}(1 + \sum_{i=1}^{N} W_i) & \text{for integer weights} \\ \frac{1}{2}\sum_{i=1}^{N} W_i & \text{for real weights.} \end{cases} \quad (4)$$

The procedure of picking the Tth largest value of a set is the same as was described above. Interested readers are referred to [24] for additional details concerning the definition of weighted medians (with real-valued weights) and particularly their relations to positive Boolean functions.

III. Statistical Properties of WM Filters

Intuitive analytical expressions are sought in this section to compute the output distribution and the output moments of weighted median filters. The noise attenuation capability of a given weighted median filter can thus be evaluated using these expressions.

The analysis in this section rests on the concept of some powerful and yet simple parameters called M_i's related to the weights of a weighted median filter. We shall define these parameters and show their relation to the weights of a weighted median and to its corresponding positive Boolean function. Some multisets of weights are next defined leading to the M_i's.

A. Definition and Properties of M_i

Definition 4:

Consider a WM filter with weight vector $\underline{W} = (W_1, W_2, \cdots, W_N)$. Denote by $\mathcal{W}$ the multiset of weights, i.e.,

$$\mathcal{W} = \{W_1, W_2, \cdots, W_N\}.$$

Denote by $\Upsilon^{[i]}$ the set of all submultisets of $\mathcal{W}$ having cardinality i, i.e.,

$$\Upsilon^{[i]} = \{A \mid A \subseteq \mathcal{W}, |A| = i\}, \quad i = 0, 1, \cdots, N, \tag{5}$$

and by $\Omega^{[i]}$ the set of those subsets of cardinality i whose sum of elements is at least the threshold T (where T is as defined in (4)), i.e

$$\Omega^{[i]} = \left\{ A \mid A \in \Upsilon^{[i]}, \sum_{W_j \in A} W_j \geq T \right\}, \quad i = 0, 1, \cdots, N. \tag{6}$$

Such subsets are called *positive subsets*.

Example 1: Given a WM filter with $\underline{W} = (1, 4, \underline{5}, 3, 2)$. According to (4), its threshold is $T = 8$. When $i = 1$, it is obvious that there is no such set which satisfies (6). When $i = 2$, there are two *positive subsets*, i.e.

$$\Omega^{[2]} = \{\{3, 5\}, \{4, 5\}\}.$$

Similarly, one can find other positive subsets:

$$\begin{aligned}\Omega^{[3]} &= \{\{1,3,4\}, \{2,3,4\}, \{1,2,5\}, \{1,3,5\}, \\ &\quad \{2,3,5\}, \{1,4,5\}, \{2,4,5\}, \{3,4,5\}\} \\ \Omega^{[4]} &= \{\{1,2,3,4\}, \{1,2,3,5\}, \{1,2,4,5\}, \{1,3,4,5\}, \\ &\quad \{2,3,4,5\}\}, \\ \Omega^{[5]} &= \{\{1,2,3,4,5\}\}.\end{aligned}$$

Definition 5: Denote by M_i the cardinality of $\Omega^{[i]}$, i.e.

$$M_i = |\Omega^{[i]}|, \quad i = 0, 1, \cdots, N. \tag{7}$$

Example 2: Consider the same WM filter in *Example* 1, we have $M_1 = 0$, $M_2 = 2$, $M_3 = 8$, $M_4 = 5$, and $M_5 = 1$.

Three properties of the M_i's are next listed. The first establishes an interesting and useful interrelation between the M_i's.

Property 1: For a WM filter with window width $N = 2K + 1$, we have

$$M_i + M_{N-i} = \binom{N}{i}, \quad i = 0, 1, \cdots, N. \tag{8}$$

Proof: Let $0 \leq i \leq N$. Notice that

$$A \in \Upsilon^{[i]} \leftrightarrow [\mathcal{W} \setminus A] \in \Upsilon^{[N-i]}$$

and

$$A \notin \Omega^{[i]} \leftrightarrow [\mathcal{W} \setminus A] \in \Omega^{[N-i]}.$$

Thus

$$|\Omega^{[N-i]}| = |\Upsilon^{[i]} \setminus \Omega^{[i]}|$$

implying

$$|\Omega^{[i]}| + |\Omega^{[N-i]}| = |\Upsilon^{[i]}|$$

which is equivalent to (8). □

The second property is more intuitive and essential stating the monotonicity of the M_i's.

Property 2: The sequence $M_1, M_2, \cdots, M_K$ is nondecreasing with respect to i, i.e.

$$M_{i+1} \geq M_i, \quad i = 1, \cdots, K. \tag{9}$$

This monotonicity property of the M_i's should be intuitive since by increasing the number of weights, more subsets are likely to become positive.

The third and final property of the M_i's establishes a direct link to the positive Boolean function (PBF) corresponding to the given weighted median. It simply states that M_i is just the cardinality of the subset of the on-set, of the PBF representing the WM, containing all true vectors with Hamming weight i.

Property 3: Suppose the binary inputs of a WM filter are vectors

$$\underline{X} = (X_1, \cdots, X_N) \in \{0, 1\}^N.$$

Denote by $f(\underline{X})$, the positive Boolean function corresponding to this WM. Then, the M_i's are related to $f(\cdot)$ as follows

$$M_i = |\{\underline{X} \in \{0,1\}^N \mid f(\underline{X}) = 1, \omega(\underline{X}) = i\}| \tag{10}$$

where $\omega(\underline{X})$ denotes the Hamming weight of $\underline{X}$ (i.e., the number of 1's in $\underline{X}$) and $|\cdot|$ denotes the cardinality operation. It is easy to see that M_i can also be expressed as follows:

$$M_i = \sum_{\underline{X} \in S_i} U(\underline{W}\,\underline{X}^t - T) \tag{11}$$

where $U(\cdot)$ is the unit step function, and S_i is the set of all N-dimensional binary vectors with Hamming weight i. i.e.,

$$S_i = \{\underline{X} \mid \underline{X} \in \{0,1\}^N; \omega(\underline{X}) = i\}. \tag{12}$$

B. Output Distribution of WM Filters

Using the above tools, we shall derive the output distribution of WM filters in an intuitive and useful form. The goal is not just to develop another expression but to come up with one which is easy to understand and would be useful for optimal design purposes.

The next theorem states such an expression for the output distribution of WM filters in terms of the input distribution and the M_i's defined in the previous subsection.

Theorem 1: Let the inputs of a WM filter with window width $N = 2K + 1$ be independent and identically distributed with a common distribution function $\Phi(t)$. The output distribution of the WM filter $\Psi_{wm}(t)$ has the following form

$$\begin{aligned}\Psi_{wm}(t) = \Psi_s(t) + \sum_{i=1}^{K} M_i(\Phi(t)^i(1-\Phi(t))^{N-i} \\ - \Phi(t)^{N-i}(1-\Phi(t))^i)\end{aligned} \tag{13}$$

where $\Psi_s(t)$ is the output distribution of the standard median filter with the same window width, i.e.,

$$\Psi_s(t) = \sum_{i=K+1}^{N} \binom{N}{i} \Phi(t)^i (1-\Phi(t))^{N-i}.$$

Proof: Let ξ denote the random variable corresponding to the output of a WM filter with weight vector $(W_1, W_2, \cdots, W_N)$ and i.i.d. inputs $\xi_1, \cdots, \xi_N$ having the same distribution function $\Phi(t)$. Now

$$\Psi_{\text{wm}}(t) = P\{\xi \le t\}.$$

The input sample space $\mathbf{R}^N$ can be partitioned into 2^N mutually exclusive events:

$$\begin{array}{c}(-\infty, t] \times (-\infty, t] \times \cdots \times (-\infty, t],\\ (t, \infty) \times (-\infty, t] \times \cdots \times (-\infty, t],\\ (-\infty, t] \times (t, \infty) \times \cdots \times (-\infty, t],\\ \cdots \\ (t, \infty) \times (t, \infty) \times \cdots \times (t, \infty).\end{array} \tag{14}$$

A typical event having i terms of type $(-\infty, t]$ and $(N-i)$ terms of type $(t, +\infty)$ has probability $\Phi(t)^i(1-\Phi(t))^{N-i}$. It is easy to see that event $\{\xi \le t\}$ is the union of some m events in (14), say event $i_1, i_2, \cdots, i_m$, where terms $(-\infty, t)$ in event i_j $(j = 1, \cdots, m)$ match some subset of $\mathcal{W}$ belonging to the set of positive subsets $\Omega^{[i_j]}$. As the events are mutually exclusive and $|\Omega^{[i]}| = M_i$, we have

$$\Psi_{\text{wm}}(t) = \sum_{i=1}^{N} M_i \Phi(t)^i (1-\Phi(t))^{N-i}$$

which may be split in two terms as follows,

$$\Psi_{\text{wm}} = \sum_{i=1}^{K} M_i \Phi^i (1-\Phi)^{N-i} + \sum_{i=K+1}^{N} M_i \Phi^i (1-\Phi)^{N-i}. \tag{15}$$

By Property 1,

$$\Psi_{\text{wm}} = \sum_{i=1}^{K} M_i \Phi^i (1-\Phi)^{N-i} + \sum_{i=K+1}^{N} \left(\binom{N}{i} - M_{N-i} \right) \Phi^i (1-\Phi)^{N-i}$$

which is rewritten as

$$\Psi_{\text{wm}} = \sum_{i=K+1}^{N} \binom{N}{i} \Phi^i (1-\Phi)^{N-i} + \sum_{i=1}^{K} M_i \Phi^i (1-\Phi)^{N-i} - \sum_{i=K+1}^{N} M_{N-i} \Phi^i (1-\Phi)^{N-i}. \tag{16}$$

Let $j = N - i$, then the last term in (16) becomes

$$\sum_{i=K+1}^{N} M_{N-i} \Phi^i (1-\Phi)^{N-i} = \sum_{i=1}^{K} M_i \Phi^{N-i} (1-\Phi)^i. \tag{17}$$

Substituting (17) into (16), we have

$$\Psi_{\text{wm}} = \sum_{i=K+1}^{N} \binom{N}{i} \Phi^i (1-\Phi)^{N-i} + \sum_{i=1}^{K} M_i (\Phi^i (1-\Phi)^{N-i} - \Phi^{N-i} (1-\Phi)^i). \quad \square$$

Note that the first sum in the last expression does not depend on M_i and thus does not depend on the weights of the WM filter. Furthermore, familiar readers would immediately recognize this term as being the output distribution of the standard median filter! The second sum quantifies the contribution of the weights in the output distribution and equals zero if all weights are equal (standard median filter).

As an immediate consequence of the above theorem, it is easy to show that WM filters are unbiased estimators of the mean.

Note also that the shape of the window of the WM filter in Theorem 1 is irrelevant, that is Theorem 1 applies to multidimensional WM filtering as well.

C. Output Moments of WM Filters

Having obtained an intuitive expression for the output distribution of WM filters, we shall now extract a similar expression for the output moments, which is easy to assimilate and useful in designing optimal WM filters. The expression is stated in the next theorem and is again composed of two terms, one is weight-independent; while the other depends on the weights.

Theorem 2: Given a WM filter with window width $N = 2K + 1$, for i.i.d. inputs with common distribution function $\Phi(t)$ and density function $\phi(t)$, the γ-order output central moment, denoted by μ_{wm}^{γ}, of the WM filter can be expressed as

$$\mu_{\text{wm}}^{\gamma} = \mu_s^{\gamma} + \sum_{i=1}^{K} M_i L_i(N, \Phi, \gamma) \tag{18}$$

where μ_s^{γ} is the γ-order central moment of the standard median with the same window size

$$L_i(N, \Phi, \gamma) = \int_{-\infty}^{+\infty} U_i(\Phi(y)) |y - m_y|^{\gamma} \phi(y) dy \ge 0 \tag{19}$$

for symmetric ϕ, $\gamma \ge 0$, $i = 1, \cdots, K$, m_y is the output mean and

$$U_i(\Phi) = (i - N\Phi)\Phi^{i-1}(1-\Phi)^{N-i-1} + (i - N(1-\Phi))\Phi^{N-i-1}(1-\Phi)^{i-1}.$$

Proof: See the Appendix. $\square$

The first term μ_s^{γ} of μ_{wm}^{γ} is the γ-order output moment of the standard median filter which is independent of the weights; while the second term quantifies the contribution of the weights to the output moment. Of course, this second term would be zero if all weights are equal.

This expression for the output moment will be studied further in the next subsection where it will be used to evaluate the noise attenuation capability of WM filters. But first, note that weights can only increase the output moment of WM filters, that is, the second term is always positive, i.e.,

$$\sum_{i=1}^{K} M_i L_i(N, \Phi, \gamma) \ge 0$$

implying that

$$\mu_{\text{wm}}^{\gamma} \ge \mu_s^{\gamma}$$

for $\gamma \ge 0$.

TABLE I
L-TABLE FOR $N = 3$ TO $N = 23$ WITH UNIFORM DISTRIBUTION

N	L_1	L_2	L_3	L_4	L_5	L_6	L_7	L_8	L_9	L_{10}	L_{11}
3	4.014-E1										
5	3.433-E1	5.714-E2									
7	2.395-E1	4.762-E2	9.524-E3								
9	1.711-E1	3.030-E2	7.792-E3	1.732-E3							
11	1.273-E1	1.958-E2	4.662-E3	1.399-E3	3.330-E4						
13	9.812-E2	1.319-E2	2.797-E3	7.992-E4	2.664-E4	6.660-E5					
15	7.789-E2	9.243-E3	1.745-E3	4.525-E4	1.469-E4	5.289-E5	1.371-E5				
17	6.333-E2	6.708-E3	1.135-E3	2.654-E4	7.938-E5	2.835-E5	1.083-E5	2.887-E6			
19	5.254-E2	5.012-E3	7.666-E4	1.622-E4	4.423-E5	1.474-E5	5.670-E6	2.268-E6	6.186-E7		
21	4.433-E2	3.839-E3	5.349-E4	1.030-E4	2.564-E5	7.867-E6	2.855-E6	1.165-E6	4.841-E7	1.345-E7	
23	3.794-E2	3.004-E3	3.834-E4	6.776-E5	1.545-E5	4.359-E6	1.468-E6	5.711-E7	2.447-E7	1.049-E7	2.958-E8

TABLE II
L-TABLE FOR $N = 3$ TO $N = 23$ WITH GAUSSIAN DISTRIBUTION

N	L_1	L_2	L_3	L_4	L_5	L_6	L_7	L_8	L_9	L_{10}	L_{11}
3	5.449-E1										
5	4.909-E1	5.395-E2									
7	3.907-E1	4.631-E2	7.639-E3								
9	3.197-E1	3.230-E2	6.372-E3	1.267-E3							
11	2.694-E1	2.312-E2	4.067-E3	1.037-E3	2.297-E4						
13	2.323-E1	1.723-E2	2.638-E3	6.221-E4	1.856-E4	4.408-E5					
15	2.039-E1	1.329-E2	1.783-E3	3.743-E4	1.062-E4	3.528-E5	8.803-E6				
17	1.816-E1	1.055-E2	1.254-E3	2.342-E4	6.029-E5	1.948-E5	6.992-E6	1.810-E6			
19	1.635-E1	8.567-E3	9.129-E4	1.528-E4	3.545-E5	1.055-E5	3.753-E6	1.429-E6	3.808-E7		
21	1.487-E1	7.095-E3	6.843-E4	1.035-E4	2.171-E5	5.888-E6	1.954-E6	7.495-E7	2.992-E7	8.153-E8	
23	1.362-E1	5.971-E3	5.256-E4	7.240-E5	1.383-E5	3.421-E6	1.045-E6	3.779-E7	1.539-E7	6.382-E8	1.771-E8

TABLE III
L-TABLE FOR $N = 3$ TO $N = 23$ WITH LAPLACIAN DISTRIBUTION

N	L_1	L_2	L_3	L_4	L_5	L_6	L_7	L_8	L_9	L_{10}	L_{11}
3	4.746-E1										
5	4.407-E1	3.391-E2									
7	3.770-E1	2.982-E2	4.085-E3								
9	3.290-E1	2.226-E2	3.473-E3	6.119-E4							
11	2.925-E1	1.705-E2	2.352-E3	5.087-E4	1.032-E4						
13	2.639-E1	1.348-E2	1.626-E3	3.210-E4	8.447-E5	1.876-E5					
15	2.409-E1	1.094-E2	1.166-E3	2.048-E4	5.053-E5	1.517-E5	3.588-E6				
17	2.219-E1	9.080-E3	8.640-E4	1.356-E4	3.024-E5	8.708-E6	2.876-E6	7.124-E7			
19	2.059-E1	7.668-E3	6.590-E4	9.315-E5	1.875-E5	4.947-E6	1.597-E6	5.669-E6	1.455-E7		
21	1.922-E1	6.573-E3	5.147-E4	6.609-E5	1.208-E5	2.903-E6	8.690-E7	3.063-E7	1.151-E7	3.037-E8	
23	1.804-E1	5.704-E3	4.102-E4	4.822-E5	8.051-E6	1.769-E6	4.866-E7	1.608-E7	6.081-E8	2.392-E8	6.456-E9

D. Noise Attenuation of WM Filters

The second order output moment is quite often used to measure the noise attenuation capability of a filter. It quantifies the spread of the input samples with respect to their mean value. It is desirable, however, in this case to derive an expression for the second order moment which is easy to compute and explicitly shows the effects of the weights of the filter. This is obtained by simply rewriting (18) with $\gamma = 2$,

$$\sigma^2_{\rm wm} = \sigma^2_s + \sum_{i=1}^{K} M_i L_i(N, \Phi, 2) \tag{20}$$

where σ^2_s denotes the variance of the standard median filter. $L_i(\cdot)$ is a function of the input distribution Φ, the window size $N = 2K+1$ and index i. Thus σ^2_s and $L_i(\cdot)$ are independent of the filter weights. M_i, of course, is a function of the weights. Let

$$\underline{L} = (L_1, L_2, \cdots, L_K)$$

and

$$\underline{M} = (M_1, M_2, \cdots, M_K)$$

denote the L-vector and M-vector, respectively. Then, (20) can be written as

$$\sigma^2_{\rm wm} = \sigma^2_s + \underline{M}\,\underline{L}^t \tag{21}$$

where t denotes the transpose.

For input signals having the same noise distribution function Φ, one can calculate the M-vectors and L-vectors and compare the variances of different WM filters. Since the L-vectors are independent of the weights, one can calculate all L-vectors for different noise distributions and different filter window widths. Tables I–III, called L-tables, show the values of the L-vectors for window sizes $N = 3$ to $N = 23$ for three commonly used (normalized) noise distributions, uniform, Gaussian and Laplacian. With the aid of these tables, one can easily compare the noise reduction behavior of different WM filters.

Example 3: Compare the noise reduction capability of the following two WM filters:

$$\text{WM}_1 = (2, 2, 2, 7, 12, \underline{13}, 12, 7, 2, 2, 2)$$
$$\text{WM}_2 = (4, 4, 4, 9, 14, \underline{25}, 14, 9, 4, 4, 4),$$

for the uniform distribution.

It is easy to obtain the M-vectors of WM_1 and WM_2 as follows:

$$\underline{M}(\mathrm{WM}_1) = (0, 0, 5, 47, 136)$$
$$\underline{M}(\mathrm{WM}_2) = (0, 0, 5, 34, 161).$$

Also, from L-table I we have

$$\underline{L} = (127.3 \quad 19.58 \quad 4.662 \quad 1.399 \quad 0.3330) \times 10^{-3}.$$

According to Table I, we have

$$\underline{L}\,\underline{M}^t(\mathrm{WM}_1) > \underline{L}\,\underline{M}^t(\mathrm{WM}_2) \tag{22}$$

which indicates that WM_2 has better noise attenuation than WM_1 for uniformly distributed noise. In fact, one can easily check that WM_2 has better noise performance than WM_1 for Gaussian and Laplacian distributed noise, as well. By increasing the variance of the distribution, the comparison results, e.g., (22) in Example 3, are not affected.

What makes this expression for the output moment even more powerful is that in many real applications of interest, the noise distribution is not known and yet, it may still be possible to compare the noise attenuation of different WM filters with the same window size.

When two WM filters WM_1 and WM_2 have the same window size, then the corresponding L-vectors are identical. Thus, according to (21), if the two M-vectors are such that

$$\underline{M}(\mathrm{WM}_1) \geq \underline{M}(\mathrm{WM}_2)$$

(partial ordering of the elements of the vectors), then

$$\sigma^2_{\mathrm{wm}}(\mathrm{WM}_1) \geq \sigma^2_{\mathrm{wm}}(\mathrm{WM}_2)$$

for any noise distribution.

Example 4: Given two WM filters with window width 7:

$$\mathrm{WM}_1 = (1, 1, 2, \underline{5}, 2, 1, 1)$$

and

$$\mathrm{WM}_2 = (1, 1, 4, \underline{5}, 4, 1, 1).$$

Which WM filter gives better noise reduction?

It is easy to compute the M-vectors for the two WM filters:

$$\underline{M}(\mathrm{WM}_1) = (0, 2, 15)$$
$$\underline{M}(\mathrm{WM}_2) = (0, 2, 13).$$

Clearly,

$$\sigma^2_{\mathrm{wm}}(\mathrm{WM}_1) > \sigma^2_{\mathrm{wm}}(\mathrm{WM}_2).$$

That is, filter $\mathrm{WM}_2 = (1, 1, 4, \underline{5}, 4, 1, 1)$ has better noise reduction than filter $\mathrm{WM}_1 = (1, 1, 2, \underline{5}, 2, 1, 1)$ for any symmetric noise distribution.

On the other hand, if two M-vectors are *incomparable* (the elements of one vector are neither all $\geq$ nor $\leq$ to the corresponding elements of the other vector), then the two WM filters cannot be compared without knowing the corresponding L-vectors.

IV. Optimal Weighted Median Filters

When deriving the statistical properties of WM filters in the previous section, we had several goals in mind. The most important is to develop an optimality theory for WM filters which allows the designer to pick an optimal WM filter in some specified sense. A theory with a similar goal has been developed for stack filters, see [22], [25], [26]. Other adaptive algorithms have been developed, with the same goal, for WM, WOS and stack filters, see [5], [21], [23], [27].

There are several optimality criteria usually used in filtering. Some are classic, e.g., the mean square error, the mean absolute error and the minimax error; while, some are relatively newer, e.g., a set of structural constraints on the filter's behavior and associative memory [22], [28], [29]. The last two are intimately related to the theory of root signal sets, which, in simple terms, define the "passband" of nonlinear filters.

In the next subsection, we shall first formulate the optimization problem in a mathematical setting. Optimal solutions are then discussed in the following subsection.

A. Problem Formulation

Assume the input x_i, of the WM filter with weight vector $\underline{W}$, is a constant signal s plus additive white noise n_i, i.e.

$$x_i = s + n_i \tag{23}$$

where i stands for the ith sample. Call $x_{i-N}, x_{i-N+1}, \cdots, x_{i-1}$ (the N samples inside the filter window) $X_1, X_2, \cdots, X_N$. The output $\hat{s}$ of the WM filter, which is an estimate of s, can be written as

$$\hat{s} = \mathrm{MED}\{W_1 \diamond X_1, \cdots, W_N \diamond X_N\}. \tag{24}$$

One of the optimization criteria of this theory is an error criterion defined as follows,

$$J_\gamma = E[|s - \hat{s}|^\gamma] \tag{25}$$

where γ is some positive constant. The values of $\gamma = 1, 2$ correspond to the mean absolute error (MAE) and the mean square error (MSE), respectively.

Since s in (23) is a constant signal, by Theorem 2, we can recast J_γ as

$$J_\gamma = \mu^\gamma_{\mathrm{wm}} = \mu^\gamma_s + \sum_{i=1}^{K} L_i(N, \Phi, \gamma) M_i \tag{26}$$

where, as stated previously, μ^γ_s and $L_i(N, \Phi, \gamma)$ are independent of the weights, while the M_i's are functions of the weights.

The other optimization criterion consists of a pre-specified set of structural constraints on the filter's behavior. The goal of the structural constraints is to preserve some desired signal details, e.g., pulses with certain width in 1-D signals, or lines and corners in images, and to remove undesired signal patterns. Usually, structural constraints originate in the multilevel domain and can be transformed to the binary domain via threshold decomposition. A WM filter which satisfies the binary constraints obtained by thresholding will automatically satisfy the multi-level constraints. A detailed discussion on this topic can be found in [22]. The transformation brings about a clear and intuitive link between the constraints and the filter weights. We shall, therefore, proceed directly with the structural constraints as given in the binary domain.

Suppose that as a result of the above transformation of the given set of structural constraints from the multilevel domain to the binary domain, we get r N-dimensional binary vectors $\underline{x}_1, \cdots, \underline{x}_r$ for which the desired WM filter should have the following output:

$$\text{WM}(\underline{x}_i) = \begin{cases} 1, & \text{for } i = 1, \cdots, q \\ 0, & \text{for } i = q+1, \cdots, r \end{cases}. \quad (27)$$

We assume here, without loss of generality, that the middle entry in each of the above vectors is unity (due to self duality of the weighted median filter). Accordingly, (27) states that the first q vectors represent structural constraints which are to be *preserved*, while the last $r - q$ vectors represent those undesired signal patters which are to be *removed*. The former are called type 1 constraints, while the latter are type 0.

Define a matrix $\mathbf{C}$

$$\mathbf{C} = \begin{pmatrix} \underline{c}_1 \\ \vdots \\ \underline{c}_r \end{pmatrix} \quad (28)$$

whose rows are given in terms of the row vectors x_i's

$$\underline{c}_i = \begin{cases} (2\underline{x}_i - \underline{1}), & i = 1, \cdots, q \\ \\ (\underline{1} - 2\underline{x}_i), & i = q+1, \cdots, r \end{cases}$$

where $\underline{1} = (1, 1, \cdots, 1)$. Then, the structural constraints can be expressed in a matrix form as follows,

$$\mathbf{C}\underline{W}^t \geq 0. \quad (29)$$

$\mathbf{C}$ is called the *characteristic matrix* of the structural constraints. It is easy to see that any set of structural constraints is uniquely defined by its characteristic matrix $\mathbf{C}$. We call a set of structural constraints, which consists of both type 1 and type 0 constraints *feasible* if the inequalities in (29) are consistent. The corresponding solution space is called the solution space of the characteristic matrix $\mathbf{C}$. If, by dropping any row from $\mathbf{C}$, the solution space changes, the set of structural constraints is called *irredundant*. Denote the *irredundant* set by $\mathcal{E}$, then

$$\mathcal{E} = \{e_1, \cdots, e_r\}. \quad (30)$$

Recall that WM filters are stack filters defined by self-dual PBF's [30]. Therefore, if for any binary vector $\underline{x}_1$, $\text{WM}(\underline{x}_1) = 1$, then $\text{WM}(\underline{x}_2) = 1$ for all $\underline{x}_2 \geq \underline{x}_1$ (by the stacking property) and $\text{WM}(\underline{x}_3) = 0$ for all $\underline{x}_3 \leq \underline{1} - \underline{x}_1$ (by the self-duality property). Consequently, if a set of structural constraints is *irredundant*, then none of its elements (rows of the $\mathbf{C}$ matrix) can be reproduced by other elements in the set, using either the *stacking* property or *self-duality*. On the other hand, a set of structural constraints, if *feasible*, can always be reduced to its *irredundant* form. Obviously, a set of structural constraints is uniquely defined by its *irredundant* form, assumed in all future reference to structural constraints, unless specified otherwise. Note that $\mathbf{C}$ and $\mathcal{E}$ are equivalent in the sense that a set of structural constraints can be uniquely defined by either $\mathbf{C}$ or $\mathcal{E}$. Usually there are many elements in an *irredundant* set, each element may have different number of weights.

Example 5: For a WM filter with window width $N = 7$ and $\underline{W} = (W_1, W_2, W_3, \underline{W_4}, W_5, W_6, W_7)$, if preserving pulses of length 2 is the only structural constraint, then

$$W_3 + W_4 \geq T$$
$$W_4 + W_5 \geq T.$$

Its *characteristic matrix* is

$$\mathbf{C} = \begin{pmatrix} -1 & -1 & 1 & 1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 1 & 1 & -1 & -1 \end{pmatrix}.$$

Its irredundant set is

$$\mathcal{E} = \{\{W_3, W_4\}, \{W_4, W_5\}\}.$$

Similarly, for 3×3 WM filters

$$\underline{W} = \begin{pmatrix} W_1 & W_2 & W_3 \\ W_4 & W_5 & W_6 \\ W_7 & W_8 & W_9 \end{pmatrix}$$

in order to preserve the horizontal, vertical and diagonal lines in images, the weights have to satisfy the following

$$W_4 + W_5 + W_6 \geq T \text{ (horizontal)}$$
$$W_2 + W_5 + W_8 \geq T \text{ (vertical)}$$
$$W_1 + W_5 + W_9 \geq T \text{ (diagonal)}$$
$$W_3 + W_5 + W_7 \geq T \text{ (diagonal)}.$$

Its *characteristic matrix* is

$$\mathbf{C} = \begin{pmatrix} -1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 \\ 1 & -1 & -1 & -1 & 1 & -1 & -1 & -1 & 1 \\ -1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 \end{pmatrix}.$$

with weights $\underline{W} = (W_1, W_2, \cdots, W_9)$. Its irredundant set is

$$\mathcal{E} = \{\{W_4, W_5, W_6\}, \{W_2, W_5, W_8\}, \{W_1, W_5, W_9\}, \{W_3, W_5, W_7\}\}.$$

Now, we shall return to the optimal problem. Using (11), the problem of finding optimal WM filters under structural constraints is stated as follows:

$$\begin{aligned} \text{Minimize } & \sum_{i=1}^{K} L_i M_i = \sum_{i=1}^{K} \sum_{\underline{x} \in S_i} L_i U(\underline{W}\,\underline{x}^t - T) \\ & \text{subject to} \\ \underline{W} &\geq 0, \quad \text{(positivity constraints)} \\ \mathbf{C}\underline{W}^t &\geq 0, \quad \text{(structural constraints)} \end{aligned} \quad (31)$$

where $\mathbf{C}$ is the *characteristic matrix*, $U(\cdot)$ is the unit-step function, and

$$S_i = \{\underline{x} \mid \underline{x} \in \{0,1\}^N; \omega(\underline{x}) = i\}$$

where $\omega(\underline{x})$ denotes the Hamming weight of $\underline{x}$.

B. Optimal Solutions

Although the constraint set in (31) is of a linear type, the optimization problem is nonlinear because of its nonlinear objective function. The objective function is not differentiable since it involves the unit-step function $U(\cdot)$. Such optimization problems can be quite difficult to solve. The unit-step function can, however, be approximated by a sigmoidal function $U_s(\cdot)$,

$$U_s(z) = \frac{1}{1+e^{-\beta z}}. \quad (32)$$

$U_s(\cdot)$ is a continuous differentiable, monotonically increasing, step-like function. Its steepness is controlled by a gain term $\beta > 0$. When the gain β is large, the unit-step function can be approximated well by the sigmoidal function. Thus the optimal problem in (31) can be restated as:

$$\begin{aligned} &\text{Minimize } \sum_{i=1}^{K} L_i M_i = \sum_{i=1}^{K}\sum_{\underline{x}\in S_i} L_i U_s(\underline{W}\,\underline{x}^t - T) \\ &\qquad\qquad \text{subject to} \\ &\underline{W} \geq 0, \quad \text{(positivity constraints)} \\ &\mathbf{C}\underline{W}^t \geq 0, \quad \text{(structural constraints).} \end{aligned} \quad (33)$$

This is a nonlinear programming problem in which the objective function has first and second derivatives. It can be solved by successive quadratic programming method [31].

Although nonlinear programming produces optimal solutions, the optimization problem has N (window size) decision variables; and it will undoubtedly be computationally very complex for large window sizes. Therefore, it is desirable to find a more efficient method. A second method is presented, as we shall see later, to simplify this problem in some cases. It reduces the nonlinear programming problem with N decision variables to the solution of a set of linear inequalities.

Consider again the optimization problem in (31). From Theorem 2, it is true that

$$L_i \geq 0, \quad \text{for} \quad i = 1, \cdots, K.$$

Thus, if M_i, for $i = 1, \cdots, K$, are minimized, then $\sum_{i=1}^{k} L_i M_i$ is minimized because of the positivity of the L_i's. Usually, minimizing M_i leads to a set of inequalities about W_i. Our second method is then to find these inequalities which minimize the M_i's. Together with the other set of inequalities due to the structural constraints, a solution of the optimization problem in (31), which satisfies the given set of structural constraints and minimizes the M_i's, can be found by solving these inequalities.

In the following, we consider the cases where only type 1 constraints are present.

Theorem 3: For a given window size N and a set of type 1 structural constraints $\mathcal{E} = \{e_1, \cdots, e_r\}$, let WMF($\mathcal{E}$, N) denote the set of all WM filters of window size N which satisfy $\mathcal{E}$. If a WM filter WM^*, WM* $\in$ WMF($\mathcal{E}, N$), satisfies the following:

$$\sum_{\substack{W_j \in A \\ A \in \Upsilon^{[i]} \setminus \Theta^{[i]}}} W_j < T, \qquad i = 1, \cdots, K \quad (34)$$

where

$$\Theta^{[i]} = \bigcup_{\substack{e_j \in B \\ B=\{e_m \mid |e_m| \leq i\}}} \{e_j \cup C \mid C \in \Upsilon^{[i-|e_j|]}, C \cap e_j = \phi\} \quad (35)$$

then, the M_i's, $i = 1, \cdots, K$ have reached their minimum simultaneously. Such WM filter is said to have reached the *global minimum* under $\mathcal{E}$.

Proof: According to (35), $\Theta^{[i]}$ contains all possible supersets, with cardinality i, of e_j, for $j = 1, \cdots, r$. Every element in $\Theta^{[i]}$ is a *positive subset* by the stacking property. Obviously, for any WM, WM $\in$ WMF($\mathcal{E}, N$),

$$M_i(\text{WM}) \geq |\Theta^{[i]}|, \quad i = 1, \cdots, K.$$

By (34),

$$M_i(\text{WM}^*) = |\Theta^{[i]}|, \quad i = 1, \cdots, K$$

which means that the M_i's, $i = 1, \cdots, K$, have reached their minimum simultaneously. □

Note that if WM* has reached the global minimum under $\mathcal{E}$, then

$$\underline{M}(\text{WM}^*) \leq \underline{M}(\text{WM}), \quad \text{for any WM} \in \text{WMF}(\mathcal{E}, N). \quad (36)$$

Theorem 4: WM* in Theorem 3, if it exists, is unique (weight permutations are excluded).

Proof: Let

$$X = \{\underline{x} \mid \underline{x} \in \{0,1\}^N\}.$$

Denote the positive Boolean function (PBF) of WM_1 and WM_2 by f_1 and f_2, respectively. The structural constraints of type 1 correspond to the true vectors in the binary domain. Denote these true vectors by X_t, i.e.

$$X_t(f) = \{\underline{x} \mid f(\underline{x}) = 1\}$$

and let X_f be the complement of the set X_t,

$$X_f = \overline{X_t} = X \setminus X_t.$$

If WM_1 and WM_2 both satisfy the same structural constraints, then

$$X_t(f_1) = X_t(f_2) = X_t$$

which implies, for any $\underline{x} \in X_t$

$$f_1(\underline{x}) = f_2(\underline{x}) = 1. \quad (37)$$

However, WM_1 and WM_2 both satisfy (34). Hence, for any binary vector $\underline{x} \in X_f$,

$$f_1(\underline{x}) = f_2(\underline{x}) = 0. \quad (38)$$

By (37) and (38) it is clear that the positive Boolean functions f_1 and f_2 are identical. □

Note, however, that two different WM filters may have the same $\underline{M}$-vector as shown in the next example.

Example 6: Consider these different WM filters with weight vectors

$$\text{WM}_1 = (7, 4, \underline{3}, 2, 1) \quad \text{and} \quad \text{WM}_2 = (4, 4, \underline{3}, 1, 1)$$

they have the same $\underline{M}$-vector

$$\underline{M}(\text{WM}_1) = \underline{M}(\text{WM}_2) = (0, 3).$$

Theorem 5: Given some structural constraints $\mathcal{E}$, producing a set of linear inequalities in the weights, i.e.

$$\mathbf{C}\underline{W}^t \geq 0. \tag{39}$$

The solution of the inequalities in (39) combined with the inequalities in (34), if it exists, is a unique optimal solution for WM filters under $\mathcal{E}$.

This theorem follows from Theorem 3, Theorem 4 and

$$L_i \geq 0, \quad \text{for } i = 1, \cdots, K.$$

Theorem 5 is a very important theorem. It states that if a WM filter has reached its global minimum, for a given set of structural constraints, then the optimal WM filter is unique and can be found by solving a group of linear inequalities.

Optimal WM filters satisfying the conditions in Theorem 5 have several interesting properties.

Corollary 1: The optimality property of the WM filter, which has reached its global minimum, is independent of the underlying noise distribution. That is, the optimal WM filter is optimal for any noise having a symmetric probability density.

Proof: This corollary follows since $L_i(N, \Phi, \gamma) \geq 0$, and the M_i's have reached their minimum simultaneously. □

Corollary 2: The optimal WM filter, which has reached its global minimum, in the mean square error (MSE) sense under a given set of structural constraints is also optimal in the mean absolute error (MAE) sense under the same structural constraints.

In the simplest case where there are no structural constraints, we have the following theorem.

Theorem 6: For a given window size $N = 2K + 1$, if there are no structural constraints on the weights, then the optimal WM filter is the standard median filter under both the MSE and MAE criteria.

Proof: Consider (26). Since $L_i \geq 0$ and $M_i \geq 0$, then by simply setting $M_i = 0$, for $i = 1, \cdots, K$. J_γ reaches its lower bound μ_s. This is simply the standard median filter. □

In fact, this result has been pointed out without proof in [16]. Although it is simple, it is very meaningful. Consider the case of linear FIR filters, if the input is a constant signal corrupted by additive white noise, then the optimal linear FIR filter, which is also called the FIR Wiener solution, is the arithmetic mean filter in which all weights are equal. Theorem 6 shows yet another striking analogy between WM filters and linear FIR filters.

V. Applications of Optimal WM Filters

The results of the previous section will be used here to design optimal WM filters in one and two dimensional applications.

A. One-Dimensional Signal Processing

A simple way to reduce the number of filter parameters when designing optimal weighted median filters is to require, e.g., symmetry in the weights. This may be a desirable constraint on the filter weights in the absence of prior information. We shall start first by defining two subclasses of weighted median filters, symmetric WM and bell-shaped WM filters.

Definition 6: A WM filter with $N = 2\ K + 1$ weights $\underline{W} = (W_1, W_2, \cdots, W_N)$ is said to be symmetric if

$$W_i = W_{N-i+1}, \quad i = 1, \cdots, N. \tag{40}$$

Note that the weights are symmetric with respect to the center weight W_{K+1}.

Definition 7: A bell-shaped WM (b-WM) filter is any WM filter whose weights are nonincreasing from the center.

In this subsection, we shall exclusively deal with symmetric bell-shaped WM filters, i.e., WM filters of this form:

$$\underline{W} = (W_K, \cdots, W_1, \underline{W}_0, W_1, \cdots, W_K)$$

where $W_i \geq W_{i+1}$, for $i = 1, \cdots, K-1$. These restrictions on the weights of WM filters are suitable for the type of structural constraints considered in this section. In addition, they reduce the design complexity by reducing the number of parameters.

B. Important Property of Symmetric b-WM Filters

The weights of any symmetric b-WM filter with window width $N = 2K + 1$ satisfy the following relation [32]:

$$2\sum_{i=p}^{K} W_i < W_0 < 2\sum_{i=p-1}^{K} W_i \tag{41}$$

for some $1 \leq p \leq K + 1$. This parameter p plays a crucial role in the optimization procedure in the sequel. A careful examination of the relation above reveals that this parameter corresponds to the minimum length of a pulse (a run of 0's or 1's of length p) which can be preserved by a symmetric b-WM filter.

In the binary domain, any signal detail can appropriately be characterized by pulses of certain length. Optimal design of WM filters, in this case, translates to the problem of finding a best WM filter which preserves those pulses and, at the same time, achieves best noise attenuation.

According to Theorem 5, this can be done by identifying the weight inequalities which minimize the M_i's (best noise attenuation) and combine them with the structural constraints stemming from the detail preservation requirement. The optimal solution must satisfy the above two requirements. Consider the following example.

Example 7: Find an optimal symmetric b-WM filter of window width $N = 7$ which would preserve pulses of length 2.

The weight vector $\underline{W}$ of a symmetric b-WM filter with window width 7 is given by

$$\underline{W} = (W_3, W_2, W_1, \underline{W}_0, W_1, W_2, W_3).$$

The structural constraints can be written as

$$W_0 > 2(W_2 + W_3) \tag{42}$$

which implies

$$W_0 + W_1 + W_v \geq T \quad \text{for } v = 1, 2, 3. \tag{43}$$

Next, with T equals half of the sum of weights, we identify those inequalities corresponding to (34). In this case, it is found that there are two weight combinations in which the sum of weights may be the largest except the weight combinations in (43). These are

$$(W_0, W_2, W_2) \text{ and } (W_1, W_1, W_2). \tag{44}$$

This is the case since $W_i \geq W_{i+1}$, $i = 0, \cdots, K-1$; in addition, we are only concerned with M_1, M_2, and M_3. Let the sum of weights in each combination in (44) be less than the threshold T. We then obtain the desired inequalities as follows,

$$\begin{cases} W_0 + W_2 + W_2 < W_1 + W_1 + W_3 + W_3 \\ W_1 + W_1 + W_2 < W_0 + W_2 + W_3 + W_3. \end{cases} \tag{45}$$

Solving (42) and (45), we obtain the solution, which, by Theorem 5, is the solution for the optimal symmetric b-WM filter of window width seven:

$$\underline{W} = (1, 1, 3, \underline{5}, 3, 1, 1).$$

Consider now a more general case. If pulses of length two are to be preserved, then what are the optimal symmetric b-WM filters of arbitrary (but fixed) window size $N = 2K+1$ that preserve such pulses and have best noise attenuation (among the class of b-WM filters with window size N)? The answer is given by the following theorem [33].

Theorem 7: Given a WM filter with window size $N = 2K+1$

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

the optimal WM filter that preserves pulses of length two has the following explicit form:

$$\underline{W} = (1, 1, \cdots, 1, K, \underline{2K-1}, K, 1, \cdots, 1, 1,).$$

The proof is in the Appendix.

It is interesting to note that optimal WM filters preserving pulses of length two can be found in closed form. This, however, is not the case for pulses of length three. One has to resort to successive quadratic programming [31] to find solutions. This approach yields real-valued WM filters which can be converted to integer-valued, see [24]. Several optimal symmetric b-WM filters preserving pulses of length 3 are listed in Table IV.

We mentioned earlier that median filters with larger window widths have better noise attenuation capabilities. Furthermore, we established that for the same set of structural constraints, there is usually a number of weighted median filters (with fixed and arbitrary window widths) which can preserve these constraints. It remains to be shown, though, that the noise attenuation capability of WM filters, preserving the same set of structural constraints, increases as the filter window width increases. The following simulation results confirm this hypothesis. Fig. 1 shows the noise attenuation of those WM filters which preserve length two pulses under uniform, Gaussian and Laplacian distributions. Similar results are found (but not included) for optimal WM filters that preserve pulses of length three.

TABLE IV
OPTIMAL WM FILTERS PRESERVING LENGTH THREE PULSES

N=3	$M_1 = 0$	{ 1 $\underline{1}$ 1 }
N=5	$M_1 = 0$ $M_2 = 0$	{ 1 1 $\underline{1}$ 1 1 }
N=7	$M_1 = 0$ $M_2 = 0$ $M_3 = 5$	{ 3 4 5 $\underline{7}$ 5 4 3 }
N=9	$M_1 = 0$ $M_2 = 0$ $M_3 = 5$ $M_4 = 29$	{ 2 2 5 6 $\underline{9}$ 6 5 2 2 }
N=11	$M_1 = 0$ $M_2 = 0$ $M_3 = 5$ $M_4 = 34$ $M_5 = 161$	{ 3 3 3 7 11 $\underline{19}$ 11 7 3 3 3 }
N=13	$M_1 = 0$ $M_2 = 0$ $M_3 = 5$ $M_4 = 53$ $M_5 = 209$ $M_6 = 484$	{ 2 2 2 2 9 10 $\underline{17}$ 10 9 2 2 2 2 }
N=15	$M_1 = 0$ $M_2 = 0$ $M_3 = 5$ $M_4 = 65$ $M_5 = 321$ $M_6 = 955$ $M_7 = 1905$	{ 2 2 2 2 2 11 12 $\underline{21}$ 12 11 2 2 2 2 2 }
N=17	$M_1 = 0$ $M_2 = 0$ $M_3 = 5$ $M_4 = 77$ $M_5 = 457$ $M_6 = 1662$ $M_7 = 4136$ $M_8 = 7447$	{ 2 2 2 2 2 2 13 14 $\underline{25}$ 14 13 2 2 2 2 2 2 }

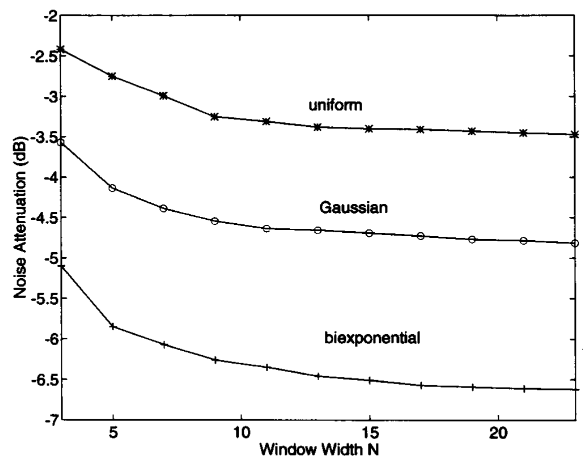


Fig. 1. Noise attenuation of optimal WM filters preserving length two pulses.

B. Optimal Weighted Median Image Filtering

An image consists of many signal structures, such as lines, corners, which are critical to perception. When filtering noisy images, care must be taken not to remove such important image details. The class of weighted median filters has proven to be a potential candidate for such tasks due to its capabilities in (impulsive) noise attenuation and detail preservation [2], [8], [9]. Several adaptive algorithms have been developed for weighted median and weighted order statistic filters, see [5], [20], [12], [21], all requiring training (ideal and noisy) signals. However, in many applications of interest, ideal signals may not be available. It is, therefore, necessary to develop an alternative scheme to design weighted median filters in these cases.

Important image structures, such as those listed in (47)–(50), constitute (part) of the structural constraints which must be preserved by the WM filter. The task is to select a WM filter, among those which preserve the given set of structural constraints, which achieves maximum noise attenuation. We shall accomplish this task using the theory developed in Section IV.

Consider WM filters with 3 × 3 square window

$$\underline{W} = \begin{pmatrix} W_1 & W_2 & W_3 \\ W_4 & W_5 & W_6 \\ W_7 & W_8 & W_9 \end{pmatrix}. \quad (46)$$

Four different sets of image structures are presented in the following. represented by a set of linear inequalities on the weights of WM filters.

Structural Constraints 1: Vertical and horizontal lines

$$\begin{cases} W_2 + W_5 + W_8 \geq T \\ W_4 + W_5 + W_6 \geq T. \end{cases} \quad (47)$$

Structural Constraints 2: Vertical, horizontal and diagonal lines

$$\begin{cases} W_2 + W_5 + W_8 \geq T \\ W_4 + W_5 + W_6 \geq T \\ W_1 + W_5 + W_9 \geq T \\ W_3 + W_5 + W_7 \geq T. \end{cases} \quad (48)$$

Structural Constraints 3: Corners

$$\begin{cases} W_5 + W_2 + W_3 + W_6 \geq T \\ W_5 + W_6 + W_8 + W_9 \geq T \\ W_5 + W_4 + W_7 + W_8 \geq T \\ W_5 + W_1 + W_2 + W_4 \geq T. \end{cases} \quad (49)$$

Structural Constraints 4: Compound structures

$$\begin{cases} W_2 + W_5 + W_8 \geq T & \text{vertical lines} \\ W_4 + W_5 + W_6 \geq T & \text{horizontal lines} \\ W_5 + W_1 + W_3 + W_7 + W_9 \geq T & \text{“}\times\text{”shapes} \\ W_5 + W_1 + W_2 + W_3 \geq T & \text{“T”shapes} \\ W_5 + W_1 + W_4 + W_7 \geq T & \text{“}\vdash\text{”shapes} \\ W_5 + W_3 + W_6 + W_9 \geq T & \text{“}\dashv\text{”shapes} \\ W_5 + W_7 + W_8 + W_9 \geq T & \text{“}\perp\text{”shapes.} \end{cases} \quad (50)$$

Optimal WM filters were found under these four sets of structural constraints [34]. It is worth to mention that these optimal WM filters are found by only solving a linear set of inequalities. They are labeled by WM1 through WM4, corresponding to *Structural Constraints* 1 through *Structural Constraints* 4, as follows:

$$\text{WM}_1 = \begin{pmatrix} 1 & 3 & 1 \\ 3 & 5 & 3 \\ 1 & 3 & 1 \end{pmatrix}, \quad \text{WM}_2 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 5 & 1 \\ 1 & 1 & 1 \end{pmatrix},$$
$$\text{WM}_3 = \begin{pmatrix} 2 & 3 & 2 \\ 3 & 5 & 3 \\ 2 & 3 & 2 \end{pmatrix}, \quad \text{WM}_4 = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 1 \end{pmatrix}. \quad (51)$$

In order to demonstrate the performance of the optimal WM filters obtained in (51), we conducted a series of simulations. In the first simulation, an original image, called "Bridge-over-stream," was corrupted by impulsive noise. The probability of impulses was 0.12 and the height of the impulses was set to ± 200. Using the original image and its corrupted version, the 3 × 3 FIR Wiener filter and the adaptive WM filters (which were trained under the MSE and MAE senses, respectively) [5] were applied to restore the corrupted image. The MSE and MAE values are listed in Table V. These filters were also applied to restore images "Lenna" and "Harbor," which were corrupted by the same noise in order to show the robustness of WM filters. Results are included in Table V. Note that the performance of the optimal WM filters (WM1 through WM4) is much better than that of the Wiener filter and the standard median filter, and is comparable to that of the adaptive WM filters. Remember that adaptive WM filters were trained using whole part of the ideal and noisy images. One should also remember that the optimality results in this paper were derived on the basis of constant signals corrupted by additive noise. The results obtained suggest that the assumed model is far from being discounted.

TABLE V
FILTERS' PERFORMANCE IN IMPULSIVE NOISE

	Bridge		Lenna		Harbor	
Filter	MSE	MAE	MSE	MAE	MSE	MAE
Wiener Filter	415	15.1	357	13.8	457	15.9
Standard Median	165	7.46	46.2	3.61	230	7.15
WM trained under MSE	141	6.39	46.4	3.88	178	5.84
WM trained under MAE	207	4.08	152	2.51	210	3.77
WM1	124	5.58	42.2	2.83	164	5.33
WM2	140	5.28	68.2	2.83	142	4.29
WM3	230	4.10	180	2.61	244	3.91
WM4	138	4.47	71.8	2.36	149	3.95

VI. CONCLUSION

Two major contributions have been reported in this paper. The first is the derivation of a new expression for the output moments of weighted median filtered data. This expression contains two parts, one is weight-independent; while, the second depends on the weights of the WM filter. The latter term is a function of a set of parameters, we called M_i's, which are the cardinalities of the positive subsets of the WM filter with a fixed Hamming weight. The minimization of these parameters leads to the second contribution of the paper which is the design of optimal WM filters under structural constraints and goals.

In the absence of any structural constraints, the optimal WM filter is shown to be the standard median filter under both the MSE and MAE criteria. When a WM filter is desired to possess certain structural constraints on its behavior, an optimal WM filter is sought to produce the best noise attenuation and, at the same time, satisfy the given set of structural constraints. If this set satisfies a certain condition, the optimal WM filter can easily be obtained by merely solving a set of linear inequalities. If, on the other hand, the condition is not satisfied, i.e., the resulting set of linear inequalities is not consistent, nonlinear programming can be used to find an optimal solution.

Applications of optimal weighted median filters in one-dimensional signals and in image filtering included in the paper clearly show the potential of this class of nonlinear filters. Much more work is needed in this direction in order to fully exploit the properties of weighted median filters.

The optimality theory developed in this paper for weighted median filters reveal many salient features of the class of WM filters as well as several striking analogies between linear FIR filters and WM filters. Could it be that weighted median filters would be as important among the class of stack filters as FIR filters for linear filters?

Appendix A
Proof of Theorem 2

By Theorem 1, we obtain the output density function $\psi_{\rm wm}$

$$\psi_{\rm wm} = \psi_s + \sum_{i=1}^{K} M_i U_i(\Phi)\phi \qquad (52)$$

where ψ_s denotes the output density function of the standard median and ϕ denotes the input density function. $U_i(\Phi)$ is defined as:

$$\begin{aligned} U_i(\Phi) &= \frac{d}{dt}(\Phi^i(1-\Phi)^{N-i} - \Phi^{N-i}(1-\Phi)^i) \\ &= (i - N\Phi)\Phi^{i-1}(1-\Phi)^{N-i-1} \\ &\quad + (i - N(1-\Phi))\Phi^{N-i-1}(1-\Phi)^{i-1}. \end{aligned} \qquad (53)$$

From (52), it is easy to obtain the γ-order output central moments $\mu_{\rm wm}^{\gamma}$ as expressed in (18) with

$$L_i(N, \Phi, \gamma) = \int_{-\infty}^{+\infty} U_i(\Phi)|y - m_y|^{\gamma}\phi dy.$$

Now, we show that $L_i \geq 0$ for $i = 1, \cdots, K$. Without loss of generality, we assume $m_y = 0$. Obviously, L_i can be expressed as

$$L_i = \int_{-\infty}^{+\infty} |y|^{\gamma} \frac{d}{d\Phi}(u_i(\Phi))d\Phi \qquad (54)$$

where

$$u_i(\Phi) = \Phi^i(1-\Phi)^{N-i} - \Phi^{N-i}(1-\Phi)^i.$$

Let $t = \Phi(y)$, then $y = \Phi^{-1}(t) = \beta(t)$ is a strictly increasing function, i.e.,

$$\frac{d}{dt}(\Phi^{-1}(t)) = \frac{d\beta(t)}{dt} > 0. \qquad (55)$$

Equation (54) can be rewritten as

$$\begin{aligned} L_i &= \int_0^1 |\beta(t)|^{\gamma} du_i(t) \\ &= |\beta(t)|^{\gamma} u_i(t)|_{t=0}^1 - \int_0^1 u_i(t) d(|\beta(t)|^{\gamma}) \\ &= -\int_0^{\frac{1}{2}} u_i(t) d(|\beta(t)|^{\gamma}) - \int_{\frac{1}{2}}^1 u_i(t) d(|\beta(t)|^{\gamma}). \end{aligned}$$

Let $\frac{1}{2} \leq t \leq 1$. Because $N - i > i$ we have

$$t^{N-2i} > (1-t)^{N-2i}$$

which implies

$$u_i(t) = t^i(1-t)^{N-i} - t^{N-i}(1-t)^i \leq 0. \qquad (56)$$

Since $y = \beta(t) = \Phi^{-1}(t) > 0$ for $t \geq \frac{1}{2}$, we have, using (55)

$$d|\beta(t)|^{\gamma} = \gamma|\beta(t)|^{\gamma-1}\frac{d}{dt}(|\beta(t)|) = \gamma|\beta(t)|^{\gamma-1}\frac{d\beta(t)}{dt} > 0$$

which together with (56) implies

$$\int_{\frac{1}{2}}^1 u_i(t) d|\beta(t)|^{\gamma} \leq 0. \qquad (57)$$

Similarly, when $0 \leq t \leq \frac{1}{2}$ we have

$$u_i(t) \geq 0. \qquad (58)$$

Furthermore, $y = \beta(t) < 0$, implying

$$\begin{aligned} d|\beta(t)|^{\gamma} &= \gamma|\beta(t)|^{\gamma-1}\frac{d}{dt}(|\beta(t)|) \\ &= \gamma|\beta(t)|^{\gamma-1}\left(-\frac{d\beta(t)}{dt}\right) < 0. \end{aligned} \qquad (59)$$

From (58) and (59), it follows that

$$\int_0^{\frac{1}{2}} u_i(t) d|\beta(t)|^{\gamma} \leq 0$$

which together with (57) proves

$$L_i = \int_{-\infty}^{+\infty} |y|^l \frac{d}{dt}(u_i(t))dt \geq 0.$$

Appendix B
Proof of Theorem 7

From the property of symmetric b-WM filters, the structural constraint of preserving pulses of length 2 gives the following inequality

$$W_0 > 2\sum_{i=2}^{K} W_i \qquad (60)$$

which implies

$$\begin{aligned} &[\{W_0, W_1\} \cup A] \in \Omega^{[K]}, \\ &\text{for} \quad A \in \Upsilon^{[K-2]}, A \cap \{W_0, W_1\} = \phi. \end{aligned} \qquad (61)$$

Recall that

$$W_0 \geq W_1 \geq W_2, \geq \cdots \geq W_K.$$

We can list two sets of weights A_1 and A_2 which belong to the set

$$\begin{aligned} &\Upsilon^{[K]} \setminus [\{W_0, W_1\} \cup A], \\ &\text{for} \quad A \in \Upsilon^{[K-2]}, A \cap \{W_0, W_1\} = \phi \end{aligned} \qquad (62)$$

these are

$$A_1 = \{W_0, \underbrace{W_2, W_2, W_3, \cdots, W_L, W_{L+1}}_{(K-1)\text{weights}}\}$$

and

$$A_2 = \{\underbrace{W_1, W_1, W_2, \cdots, W_L, W_{L+1}}_{K\text{weights}}\}.$$

Note that there does not exist another set of K weights, belonging to the above set (62), whose sum of weights is larger than those of A_1 or A_2. Let the sum of A_1 and A_2 each be less than the threshold T, i.e.

$$\begin{cases} W_0 + 2\sum_{j=2}^{(\frac{K-1}{2})} W_j < T \\ 2\sum_{j=1}^{(\frac{K-1}{2})} W_j + W_{(\frac{K+1}{2})} < T \end{cases} \quad \text{when } K \text{ is odd} \qquad (63)$$

and

$$\begin{cases} W_0 + 2\sum_{j=2}^{(\frac{K-2}{2})} W_j + W_{(\frac{K}{2})} < T \\ 2\sum_{j=1}^{(\frac{K}{2})} W_j < T \end{cases} \quad \text{when } K \text{ is even} \quad (64)$$

then by Theorem 3, the M_i's have reached their global minimum. Combining inequality (60) with (63) or (64), one solution of these inequalities has the following explicit form:

$$\underline{W} = (1, 1, \cdots, 1, K, \underline{2K-1}, K, \cdots, 1, 1).$$

According to Theorem 5, the solutions are the optimal WM filters preserving pulses of length 2. It is easy to check the M_i's in this case can be expressed in the following closed form, for $i = 1, \cdots, K$

$$M_i = 2\binom{2K-2}{i-2} + \binom{2K-2}{i-3}.$$

REFERENCES

[1] B. J. Justusson, "Median filtering: Statistical properties," in *Two Dimensional Digital Signal Processing II* (T. S. Huang, Ed.). Berlin: Springer-Verlag, 1981.

[2] T. Loupas, W. N. McDicken, and P. L. Allan, "An adaptive weighted median filter for speckle suppression in medical ultrasonic images," *IEEE Trans. Circ. Syst.*, vol. 36, pp. 129–135, Jan. 1989.

[3] T. Jarske and Y. Neuvo, "Adaptive DPCM with median type predictors," *IEEE Trans. Consumer Elect.*, vol. 37, pp. 348–352, Aug. 1991.

[4] T. Jarske, K. Saarinen, J. Juhola, and Y. Neuvo, "Quincunx coding for picture memories," in *Proc. 131st SMPTE Tech. Conf.*, Los Angles, CA, Oct. 21–25, 1989.

[5] L. Yin, J. Astola, and Y. Neuvo, "Adaptive stack filtering with application to image processing," *IEEE Trans. Signal Processing*, vol. 41, pp. 162–184, Jan. 1993.

[6] B. Alp, J. Juhola, Y. Neuvo, and T. Jarske, "Multidimensional reconstruction of quincunx coded image sequences," in *Proc. Picture Coding Symp.*, Cambridge, MA, Mar. 1990, pp. 4.5-1–4.5-3.

[7] J. Salo, Y. Neuvo, and V. Hameenaho, "Improving TV picture quality with linear median type operation," *IEEE Trans. Consumer Elect.*, vol. CE-34, pp. 373–379, Aug. 1988.

[8] D. R. K. Brownrigg, "The weighted median filter," *Commun. ACM*, vol. 27, pp. 807–818, Aug. 1984.

[9] S.-J. Ko and Y. H. Lee, "Center weighted median filters and their application to image enhancement," *IEEE Trans. Circ. Syst.*, vol. 38, pp. 984–993, Sept. 1991.

[10] P. Haavisto, M. Gabbouj, and Y. Neuvo, "Median based idempotent filters," *J. Circ. Syst. Comp.*, vol. 1, no. 2, pp. 125–148, June 1991.

[11] P. Haavisto, J. Juhola, and Y. Neuvo, "Motion adaptive scan rate up-conversion," *Multidimensional Syst. Sig. Process.*, Special issue on Multidimensional Processing of Video Signals, vol. 3, no. 2/3, pp. 113–130, May 1992.

[12] I. Tabus, M. Gabbouj, and L. Yin, "Applications of neural networks training methods to optimal WOS image filtering," in *Proc. IEEE Workshop Visual Signal Processing Commun.*, Melbourne, Australia, Sept. 20–23, 1993.

[13] D. R. K. Brownrigg, "Generation of representative members of an RrSst weighted median filter class," *Proc. Inst. Elect. Eng.*, vol. 133, pt. F, no. 5, pp. 445–448, Aug. 1986.

[14] R. P. Lippmann, "An introduction to computing with neural nets," *IEEE ASSP Mag.*, pp. 4–22, Apr. 1987.

[15] P. D. Wendt, E. J. Coyle, and N. C. Gallagher, Jr., "Stack filters," *IEEE Trans. Acoust. Speech Signal Processing*, vol. 34, pp. 898–911, Aug. 1986.

[16] 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 Trans. Acoust. Speech Signal Processing*, vol. 39, pp. 395–410, Feb. 1991.

[17] P. D. Wendt, "Nonrecursive and recursive stack filters and their filtering behavior," *IEEE Trans. Acoust. Speech Signal Processing*, vol. 38, pp. 2099–2107, Dec. 1990.

[18] M. K. Prasad and Y. H. Lee, "Weighted median filters: generation and properties," in *Proc. IEEE Int. Symp. Circ. Syst., ISCAS-89*, Portland, OR, May 8–11, 1989, pp. 425–428.

[19] ——, "Analysis of weighted median filters based on inequalities relating the weights," *Circ. Syst. Signal Processing*, vol. 11, no. 1, pp. 115–136, 1992.

[20] L. Yin, J. Astola, and Y. Neuvo, "A new class of nonlinear filters—neural filters," *IEEE Trans. Signal Processing*, vol. 41, pp. 1201–1222, Mar. 1993.

[21] I. Tabus, M. Gabbouj, and L. Yin, "Real domain adaptive WOS filtering using neural network approximations," in *Proc. IEEE Winter Workshop on Nonlinear Digital Signal Processing*, Jan. 17–20, 1993, Tampere, Finland, pp. 7.2-1.1–7.2-1.6.

[22] M. Gabbouj and E. J. Coyle, "Minimum mean absolute error stack filtering with structural constraints and goals," *IEEE Trans. Acoust. Speech Signal Processing*, vol. 38, pp. 955–968, June 1990.

[23] J. Nieweglowski, L. Yin, M. Gabbouj, and Y. Neuvo, "Optimal weighted order statistic filters under structural constraints," in *Proc. IEEE Int. Symp. Circ. Syst.*, San Diego, CA, May 10–13, 1992., pp. 2621–2624.

[24] J. Nieweglowski, M. Gabbouj, and Y. Neuvo, "Weighted medians-positive Boolean functions conversion algorithms," *Signal Processing*, vol. 39, no. 2, 1994.

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

[26] E. Coyle and J.-H. Lin, "Stack Filters and the mean absolute error criterion," *IEEE Trans. Acoust. Speech Signal Processing*, vol. ASSP-36, pp. 63–76, Jan. 1988.

[27] J.-H. Lin, T. M. Sellke, and E. Coyle, "Adaptive stack filtering under the mean absolute error criterion," *IEEE Trans. Acoust. Speech Signal Processing*, vol. 38, pp. 938–954, June 1990.

[28] P.-T. Yu and E. Coyle, "Convergence behavior and N-roots of stack filters," *IEEE Trans. Acoust. Speech Signal Processing*, vol. 38, pp. 1529–1544, Sept. 1990.

[29] ——, "On the existence and design of the best stack filter based associative memory," *IEEE Trans. Circ. Syst.*, vol. 39, pp. 171–184, Mar. 1992.

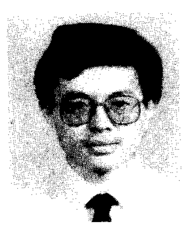
[30] S. Muroga, *Threshold Logic and Its Applications*. New York: Wiley, 1971.

[31] "User's Manual, Math/Library," Problem-Solving Software Systems, version 1.0, Apr. 1987.

[32] B. Zeng, "On the root structures of weighted median filters," *IEEE Trans. Circ. Syst.*, vol. 37, pp. 1402–1404, Nov. 1991.

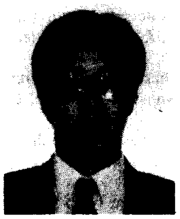
[33] R. Yang, L. Yin, M. Gabbouj, J. Astola, and Y. Neuvo, "Optimal weighted median filters under structural constraints," in *Proc. IEEE Int. Symp. Circ. Syst.*, Chicago, IL, May 3–6, 1993, pp. 942–945.

[34] R. Yang, L. Yin, M. Gabbouj, and Y. Neuvo, "2-D optimal weighted median filters," in *IEEE Workshop on Visual Signal Processing Commun.*, Sept. 2–3, 1992, pp. 31–35.



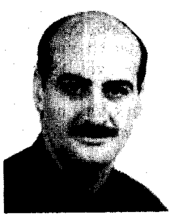
Ruikang Yang (S'93) received the B.S. and M.S. degrees from the University of Electronic Science and Technology of China (formerly Chengdu Institute of Radio Engineering), Chengdu, China, in 1983 and 1986, respectively, and the Ph.D. degree in electrical engineering from Tampere University of Technology, Finland, in 1994.

From 1986 to 1991, he was a Lecturer in the Department of Electrical Engineering, University of Electronic Science and Technology of China, Chengdu, China. Between 1991 and 1993, he was with the Signal Processing Laboratory, Tampere University of Technology, Finland. Since 1994, he has been a Senior Research Engineer at the Nokia Research Center, Tampere, Finland. His research interests include nonlinear filtering theory, image filtering and compression, neural networks, VLSI design, and speech recognition.



Lin Yin received the B.S.E.E. degree from the Shandong University of Technology, China, in 1983, the M.S.E.E. degree from the Xi'an Jiaotong University, China, in 1986, and the Doctor of Technology degree in electrical engineering from the Tampere University of Technology, Finland, in 1991.

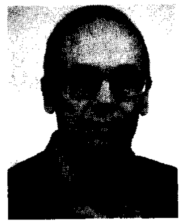
From 1986 to 1989, he was a Teacher in the Department of Information and Control Engineering at the Xi'an Jiaotong University. Between 1989 and 1992, he was with the Signal Processing Laboratory, Tampere University of Technology, as a Researcher. Currently, he is a DSP Engineer in the R&D Center of the Nokia Mobile Phones, Oulu, Finland. His research interests include nonlinear signal processing, adaptive algorithms, speech and channel coding, and neural networks.



Moncef Gabbouj (S'85–M'90) received the B.S. degree in electrical engineering in 1985 from Oklahoma State University, Stillwater. He received the M.S. and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, in 1986 and 1989, respectively.

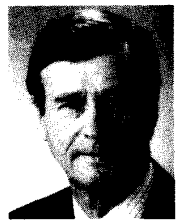
He is an Associate Professor with the Signal Processing Laboratory, Department of Information Technology, Tampere University of Technology, Finland. Since 1990, he has been a Senior Research Scientist with the Research Institute for Information Technology, Tampere, Finland. He is the Director of the International University Program in Digital Signal Processing in the Signal Processing Laboratory, Department of Information Technology, Tampere University of Technology. His research interests include nonlinear signal and imaging processing and analysis, mathematical morphology, and neural networks.

Dr. Gabbouj is an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING and Guest Editor of the European journal *Signal Processing*, Special Issue on Nonlinear Digital Signal Processing (August 1994). He is Chairman of the IEEE Circuits and Systems Society Technical Committee on Digital Signal Processing and Chairman of the IEEE SP/CAS Chapter of the IEEE Finland Section. He is a member of Eta Kappa Nu and Phi Kappa Phi. He was a co-recipient of the Myril B. Reed Best Paper Award from the 32nd Midwest Symposium on Circuits and Systems.



Jaakko Astola received the B.Sc., M.Sc., Licentiate, and Ph.D. degrees in mathematics from Turku University, Finland, in 1972, 1973, 1975, and 1978, respectively.

From 1976 to 1977, he was a Research Assistant at the Research Institute for Mathematical Sciences of Kyoto University, Japan. Between 1979 and 1987, he was with the Department of Information Technology, Lappeenranta University of Technology, Finland, holding various teaching positions in mathematics, applied mathematics, and computer science. From 1988 to 1993, he was an Associate Professor in Applied Mathematics, Tampere University, Finland. Currently, he is a Professor of Digital Signal Processing at the same institution. His research interests include image and signal processing, coding theory, and statistics.



Yrjö Neuvo (S'70–M'74–SM'82–F'89) received the Ph.D. degree in electrical engineering from Cornell University, Ithaca, NY, in 1974.

He held various research and teaching positions at Helsinki University of Technology, the Academy of Finland, and Cornell University from 1968 to 1976. Since 1976, he has been a Professor of Electrical Engineering at the Tampere University of Technology, Finland. In the academic year 1981–1982, he was with the University of California, Santa Barbara, as a Visiting Professor. In 1984–1992, he was a National Research Professor at the Academy of Finland. In 1993, he joined Nokia Corporation as Senior Vice President, Technology. Since February 1994, he has been Senior Vice President, R&D, Nokia Mobile Phones.

Dr. Neuvo is a member of EURASIP, Phi Kappa Phi, the Finnish Academy of Technical Sciences, Foreign Member of the Royal Swedish Academy of Technical Sciences, and Academiae Europae. He is an Honorary Doctor of Medicine from Tampere University. He is a member of the EC's Committee for Development and Science and Technology (CODEST) and board member of the Nodic Fund for Technology and Development. He is editor and member of the editorial boards of several professional journals. He serves as a member of the Signal Processing Committee of the IEEE Circuits and Systems Society. He was a Distinguished Lecturer of that Society during 1991–1992. He was the President of the Society of Electrical Engineers in Finland during 1978–1980. He was the General Chairman of the 1988 IEEE International Symposium on Circuits and Systems, Helsinki, Finland. He has published over 400 technical articles and holds several patents.