L_p NORM DESIGN OF WEIGHTED ORDER STATISTIC FILTERS

C. E. Savin *M. O. Ahmad* *M. N. S. Swamy*

Centre for Signal Processing and Communications
Department of Electrical and Computer Engineering
Concordia University
1455 de Maisonneuve Blvd. West
Montreal, Quebec, Canada H3G 1M8

E-mails: cristian@ece.concordia.ca, omair@ece.concordia.ca & swamy@ece.concordia.ca

ABSTRACT

This paper addresses the problem of designing weighted order statistic (WOS) filters by employing an objective function given as the L_p norm of the error between the desired signal and the estimated one. The conventional design of WOS filters uses a mean absolute error (MAE) objective function, and as such, it is a special case of the general, L_p norm based design, developed here. In this paper, it is shown that in stack filtering, the L_p norm can be expressed as a linear combination of the decision errors incurred by the Boolean operators at each level of the stack filter architecture. Based on this formulation of the L_p norm, both nonadaptive and adaptive algorithms for the design of L_p WOS filters are developed. A design example is considered, to illustrate the performance of the designed L_p WOS filters with different values of p. The simulation results show that the L_p WOS filters with $p \geq 2$ are capable of removing more impulsive noise compared with the conventional MAE WOS filters.

1. INTRODUCTION

In recent years, considerable attention has been given to the study of rank order based filters, since these filters yield improved performance over linear filters in restoring signals corrupted with impulsive noise when sharp edges are to be preserved. A very important part of the research in this field has been concerned with the study of stack filters. A stack filter has been defined in [1] as being a threshold decomposition configuration in which the filtering operation is performed via a positive Boolean function (PBF). It has been shown in [2] that a stack filter which is optimal for signal estimation in the mean absolute error (MAE) sense can be determined as the solution of a linear program (LP). Unfortunately, the complexity of this LP increases faster than exponentially with the size of the filter. Therefore, an alternative approach of reducing the computational burden has been devised in [3],[4], by restricting the analysis to the subclasses of stack filters of well-known practical significance. One such example is the subclass of stack filters defined by linearly separable PBFs (LSPBFs), which perform the operation of weighted order statistic (WOS) filtering in the multilevel domain [6]. The problem of designing a WOS filter which is approximatley optimal in the MAE sense has been solved in [3].

The attractive feature of using the MAE criterion for the design of stack filters is that it allows the decomposition of the estimation error of the filter into a sum of decision errors incurred by the Boolean operators at each level of the stack filter architecture [2]. Although the MAE criterion of designing stack filters has been extensively used for image processing applications, higher order errors have been used only in restricted designs of stack filters, where the input signal is assumed to be a constant signal embedded in white noise [4],[5].

In spite of some expected benefits that could possibly be achieved by employing an objective function given as the pth order error between the desired and estimated signals in stack filter design, no attempt seems to have been made to develop a mathematical framework needed for this design problem. In this paper, we investigate the possibility of formulating the design problem of pth order error optimal stack filters as a linear optimization problem. Since the design of stack filters is generally carried out under the assumption of ergodicity (i.e., it is assumed that sample averages are equal to time averages) and by using "training" sequences [7], in this paper, the pth order error is simply referred to as an L_p norm. It is shown that in the case of signal estimation using stack filters, the L_p norm can be expressed as a linear function of the decision errors incurred by the Boolean operators at each level of the stack filter architecture. Based on this formulation of the L_p norm, we develop algorithms for the nonadaptive and adaptive design of L_p WOS filters.

2. STACK FILTER DESIGN USING AN L_p NORM OBJECTIVE FUNCTION

Let $\mathrm{X}(n)$ denote an L-level process which is received at the input of a stack filter $\mathcal{S}_f$ characterized by a positive Boolean function (PBF) f of window size M. The input process $\mathrm{X}(n)$ is assumed to be a corrupted version of some desired process $\mathrm{S}(n)$. At each instant n, the output of the stack filter is an estimate of $\mathrm{S}(n)$. This estimate is based on the sequence $\mathbf{X}(n)$ that appears in the input window of the stack filter, and it is denoted by $\mathcal{S}_f(\mathbf{X}(n))$. The pth order error between the desired signal and the filter's output is given by

$$J_p(\mathcal{S}_f) = E\left[\, |\, \mathrm{S}(n) - \mathcal{S}_f(\mathbf{X}(n))\, |^p\, \right] . \tag{1}$$

The minimum L_p norm design of stack filters can be formulated as an optimization problem, in which the cost function given by (1) is minimized. This minimization is subjected to the constraint that the Boolean operator at the binary levels of the filter satisfy the stacking property.

In order to obtain a binary-level expression for the pth order error between the desired and estimated signals, we make use of the following two properties of stack filters, which are revealed here for the first time in the literature of stack filtering.

Theorem 1 In stack filtering, at the binary level, the pth power of the multilevel error,

$$e^p(n) \triangleq (\,\mathrm{S}(n) - \mathcal{S}_f(\mathbf{X}(n))\,)^p\,, \tag{2}$$

can be expressed as

$$e^p(n) = \sum_{\ell=1}^{\mathrm{L}-1} \mathrm{A}_p(\ell,n)\cdot(\mathrm{s}_\ell(n) - f(\mathbf{x}_\ell(n)))\,, \tag{3}$$

where

$$\mathrm{A}_p(\ell,n) \triangleq (\mathrm{S}(n)-\ell+1)^p - (\mathrm{S}(n)-\ell)^p\,, \tag{4}$$

with $\mathrm{s}_\ell(n)$ and $f(\mathbf{x}_\ell(n)))$ designating, respectively, the binary sequences at the level ℓ in the threshold decompositions of $\mathrm{S}(n)$ and $\mathbf{X}(n)$.

Theorem 2 In stack filtering, the absolute value of the pth order error between the desired signal and the estimated one can be determined as a summation of the absolute values of the weighted errors, $e_{\mathrm{b}}(\ell,n) \triangleq |\mathrm{A}_p(\ell,n)||\mathrm{s}_\ell(n) - f(\mathbf{x}_\ell(n))|$, appearing at the binary levels of the filter, i.e.,

$$|\,e(n)\,|^p = \sum_{\ell=1}^{\mathrm{L}-1} |\,\mathrm{A}_p(\ell,n)\,|\cdot|\,\mathrm{s}_\ell(n) - f(\mathbf{x}_\ell(n))\,|\;. \tag{5}$$

Due to the lack of space in the present paper, the proofs of these theorems will be published elsewhere. However, it is observed that the results of (3) and (5) can be easily verified numerically, by choosing arbitrary values for p, $\mathrm{S}(n)$, and $\mathcal{S}_f(\mathbf{X}(n))$. As an example, to verify the relation (3) for the case of $p=2$, let us assume that $\mathrm{S}(n)=9$ and $\mathcal{S}_f(\mathbf{X}(n)) = 6$. Then, $e(n) = \mathrm{S}(n) - \mathcal{S}_f(\mathbf{X}(n)) = 3$. Now, $e^p(n) = \sum_{\ell=1}^{\mathrm{L}-1}\mathrm{A}_p(\ell,n) = 9$, as expected.

As an immediate consequence of Theorem 2, the pth order error given by (1) can be expressed as a linear function of the decision errors at the binary levels of the filter, i.e.,

$$J_p(\mathcal{S}_f) = \sum_{\ell=1}^{\mathrm{L}-1} E\,[\,|\,\mathrm{A}_p(\ell,n)\,|\cdot|\,\mathrm{s}_\ell(n) - f(\mathbf{x}_\ell(n))\,|\,]\;. \tag{6}$$

Based on the above formulation of the objective function for stack filter design, in the subsequent, we investigate the problem of L_p norm design of weighted order statistic filters.

3. L_p NORM DESIGN OF WOS FILTERS

As shown in [6], the output of a WOS filter at an instant n can be obtained by the following procedure:

(a) replicate each input sample, $\mathrm{X}(n-j)$ $(j=0,1,\ldots,\mathrm{M}-1)$, appearing in the filter's input window (which is denoted by $\mathbf{X}(n) = \{\mathrm{X}(n), \mathrm{X}(n-1), \ldots, \mathrm{X}(n-\mathrm{M}+1)\}$) at time n, by a given positive integer w_j called the weight;

(b) sort the resulting vector of $\sum_{j=0}^{\mathrm{M}-1}\mathrm{w}_j$ elements;

(c) choose the w_T-th largest value (w_T denotes a positive integer called the threshold) from the sorted vector.

Therefore, a WOS filter is completely determined by a set of positive integer weights, $\mathrm{w}_0, \mathrm{w}_1, \ldots, \mathrm{w}_{\mathrm{M}-1}$, and w_T, and it is defined by the following input-output relation

$$\mathrm{WOS}(\mathbf{X}(n)) \;=\; \mathrm{w}_\mathrm{T}\text{-th largest value in the set}$$

$$\left\{ \overbrace{\mathrm{X}(n),\ldots,\mathrm{X}(n)}^{\mathrm{w}_0\ \text{times}},\ldots,\overbrace{\mathrm{X}(n-\mathrm{M}+1),\ldots,\mathrm{X}(n-\mathrm{M}+1)}^{\mathrm{w}_{\mathrm{M}-1}\ \text{times}} \right\}.$$

A WOS filter is a special type of stack filter, which is characterized by a linearly separable PBF [3],[6]. A PBF $f(\mathbf{x})$ is said to be linearly separable if it can be expressed in the form

$$f(\mathrm{x}_0,\ldots,\mathrm{x}_{\mathrm{M}-1}) = \begin{cases} 1 & \text{if} \quad \displaystyle\sum_{j=0}^{\mathrm{M}-1}\mathrm{w}_j\cdot\mathrm{x}_j \geq \mathrm{w}_\mathrm{T} \\ 0 & \text{otherwise}\,, \end{cases} \tag{7}$$

where w_T and all w_j's are positive real numbers.

The L_p norm design of a WOS filter requires to determine the weights $\mathbf{w} = [\mathrm{w}_0\mathrm{w}_1\ldots\mathrm{w}_{\mathrm{M}-1}]$, such that the pth order error in estimating a signal, $\mathrm{S}(n)$, from a noise corrupted observation of the same, $\mathrm{X}(n)$, is minimized. In order to derive algorithms for this design, it is noted that $J_p(\mathcal{S}_f)$ given by (6) can be equivalently expressed as

$$J_p(\mathcal{S}_f) = \sum_{\ell=1}^{\mathrm{L}-1} E\left[\,|\,\mathrm{A}_p(\ell,n)\,|\cdot(\,\mathrm{s}_\ell(n) - f(\mathbf{x}_\ell(n))\,)^2\,\right]\;. \tag{8}$$

Thus, the L_p norm design of WOS filters involves finding a PBF $f(\mathbf{x})$ which minimizes (8), subject to $f(\mathbf{x})$ being linearly separable. In order to overcome the difficulty of imposing the constraint of linear separability, we follow an approach commonly used in WOS filter design [3],[4], in which a linear approximation of $f(\mathbf{x})$ is employed. With this approximation, the L_p norm design of WOS filters is carried out by minimizing the cost function given by

$$\tilde{J}_p(\mathbf{w}) = \sum_{\ell=1}^{\mathrm{L}-1} E\left[\,|\,\mathrm{A}_p(\ell,n)\,|\cdot(\mathrm{s}_\ell(n) - \mathbf{w}\cdot\mathbf{x}_\ell^{\mathrm{T}}(n))^2\right]\;. \tag{9}$$

One may notice that after replacing the function $f(\mathbf{x})$ appearing in the expression for $J_p(\mathcal{S}_f)$ given by (8) with a linear function, WOS filters become, in fact, linear FIR filters. That is, the design problem reduces to that of finding an optimal linear FIR filter with nonnegative weights. However, in contrast to the traditional LMS linear filtering [8], this optimal FIR filter minimizes a weighted sum of squared errors incurred at the levels of the threshold decomposition architecture.

The positive weights $\mathbf{w}^* = [\mathrm{w}_0^*\mathrm{w}_1^*\ldots\mathrm{w}_{\mathrm{M}-1}^*]$, which minimize (9), determine an LSPBF f^* given by

$$f^*(\mathrm{x}_0,\ldots,\mathrm{x}_{\mathrm{M}-1}) = \begin{cases} 1 & \text{if} \quad \displaystyle\sum_{j=0}^{\mathrm{M}-1}\mathrm{w}_j^*\cdot\mathrm{x}_j \geq 0.5 \\ 0 & \text{otherwise}\,. \end{cases} \tag{10}$$

Following an approach similar to that given in [3] for the case of approximately optimal MAE WOS filters, it can be shown that $\tilde{J}_p(\mathbf{w}^*)$ is close to the L_p norm achieved by an L_p-optimal stack filter, in spite of the fact that $\tilde{J}_p(\mathbf{w})$ of (9) is not identical to $J_p(\mathcal{S}_f)$ given by (8).

A. Nonadaptive Design

The gradient vector of the cost function $\tilde{J}_p(\mathbf{w})$ of (9) is given by

$$\nabla \tilde{J}_p = -2 \sum_{\ell=1}^{\mathrm{L}-1} E\left[\,|\,\mathrm{A}_p(\ell,n)\,|\cdot\left(\mathrm{s}_\ell(n) - \mathbf{w}\cdot\mathbf{x}_\ell^{\mathrm{T}}(n)\right)\mathbf{x}_\ell^{\mathrm{T}}(n)\,\right]. \tag{11}$$

Thus, when the positivity constraints are not imposed on $\mathbf{w}$, a WOS filter which is approximately optimal in the sense of the L_p norm is obtained as the solution of the following system of linear equations

$$\mathbf{R}\cdot\mathbf{w}^* = \mathbf{c}\;, \tag{12}$$

where $\mathbf{R}$ is an autocorrelation matrix given by

$$\mathbf{R} = \sum_{\ell=1}^{\mathrm{L}-1} E\left[\,|\,\mathrm{A}_p(\ell,n)\,|\cdot\mathbf{x}_\ell^{\mathrm{T}}(n)\cdot\mathbf{x}_\ell(n)\,\right], \tag{13}$$

and $\mathbf{c}$ is a cross-correlation vector given as

$$\mathbf{c} = \sum_{\ell=1}^{\mathrm{L}-1} E\left[\,|\,\mathrm{A}_p(\ell,n)\,|\cdot\mathrm{s}_\ell(n)\cdot\mathbf{x}_\ell^{\mathrm{T}}(n)\,\right]. \tag{14}$$

The entries of the autocorrelation matrix can be easily evaluated at the multilevel domain as

$$\mathbf{R}(i,j) = E\left[\sum_{\ell=1}^{\mathbf{L}_{\mathbf{R}}(i,j)} |\,(\mathrm{S}(n)-\ell+1)^p - (\mathrm{S}(n)-\ell)^p\,|\right], \tag{15}$$

with

$$\mathbf{L}_{\mathbf{R}}(i,j) = \min\{\mathrm{X}(n-i), \mathrm{X}(n-j)\}, \tag{16}$$

and $i,j = 0,1,\ldots,\mathrm{M}-1$. Similarly, the entries of the cross-correlation vector can be determined as

$$\mathbf{c}(i) = E\left[\sum_{\ell=1}^{\mathbf{L}_{\mathbf{C}}(i)} |\,(\mathrm{S}(n)-\ell+1)^p - (\mathrm{S}(n)-\ell)^p\,|\right], \tag{17}$$

with

$$\mathbf{L}_{\mathbf{C}}(i) = \min\{\mathrm{X}(n-i), \mathrm{S}(n)\}, \tag{18}$$

and $i = 0,1,\ldots,\mathrm{M}-1$. Note that for the special case of $p = 1$, $\mathbf{R}$ and $\mathbf{c}$ become equal to the morphological correlation matrices appearing in the conventional MAE design of WOS filters developed in [3], i.e.,

$$\mathbf{R}(i,j) = E\left[\min\{\mathrm{X}(n-i), \mathrm{X}(n-j)\}\right], \tag{19}$$

and

$$\mathbf{c}(i) = E\left[\min\{\mathrm{X}(n-i), \mathrm{S}(n)\}\right]. \tag{20}$$

With linear inequality constraints imposing the positivity of the weights, the L_p optimization problem can be solved by a gradient projection method similar to the one used for the design of MAE WOS filters [3].

B. Adaptive Design

When the statistics of the observed and desired signals are not available, the following adaptive algorithm can be used to estimate the weights of an L_p-optimal WOS filter:

$$\mathbf{w}(n+1) \;=\; \mathbf{w}(n) + {} \\ +\mu \sum_{\ell=1}^{\mathrm{L}-1} \left[|\,\mathrm{A}_p(\ell,n)\,|\left(\mathrm{s}_\ell(n) - \mathbf{w}(n)\cdot\mathbf{x}_\ell^{\mathrm{T}}(n)\right)\mathbf{x}_\ell^{\mathrm{T}}(n)\right]. \tag{19}$$

An alternative adaptive design algorithm can be derived by using a sigmoidal approximation for the linearly separable PBF $f(\mathbf{x})$ appearing in the expression of the cost function $J_p(\mathcal{S}_f)$ given by (8), instead of the linear approximation which has been used for the derivation of (19).

While nonadaptive algorithms are known to give good results when the observed and desired signals are jointly stationary, an adaptive algorithm can track the time-varying statistics of the signals and it is suitable for low-cost implementations.

4. DESIGN EXAMPLE

To assess the performance of the L_p WOS filters with different values of p, their application to the problem of restoring images corrupted with impulsive noise has been investigated. Figures 1(a) and 1(b) show, respectively, a typical test image and its noise-corrupted version which have been used in the experiment. The noise-corrupted image of Figure 1(b) has been generated by adding positive impulsive noise to the noise-free image of Figure 1(a). The impulsive noise has a probability of occurrence of 0.35, and a magnitude of 255. Figure 2 illustrates the processed images obtained by applying the L_p WOS filters of size 5×5 with $p = 1, 2, 5$ and 8, to the degraded image of Figure 1(b). The values of L_p^p ($p = 1, 2, 5$ and 8) error of the restored images are listed in Table 1, while the corresponding values of $\tilde{J}_p(\mathbf{w})$ are given in Table 2. As illustrated by Figure 2 of this example, the L_p WOS filters with $p \geq 2$ provide a better visual performance than that provided by the MAE WOS filter. Specifically, the L_2, L_5 and L_8 WOS filters are capable of removing more impulses compared with the conventional MAE WOS filter. Moreover, as shown by the results in Table 1, the $\tilde{J}_1$-optimal WOS filter (traditionally referred to as the MAE WOS filter [3],[4]) provides a larger value of the MAE cost function compared to the L_p WOS filters with $p = 2, 5$ and 8.

The results of Table 1 indicate that, in this experiment, the MAE, MSE, L_5-norm error and L_8-norm error continuously decrease for the L_p WOS filters with increasing values of p. Thus, in this example, the L_∞ WOS filter is expected to provide the best approximation to any L_p-optimal WOS filter, among all possible approximately optimal L_p WOS filters. It has been also observed that actually, the L_8 WOS filter is very close to the L_∞ WOS filter. Specifically, by increasing p beyond the value $p = 8$, the designed filters achieve almost insignificant reductions of the values of the L_p-norm errors compared to those of the L_8 WOS filter.

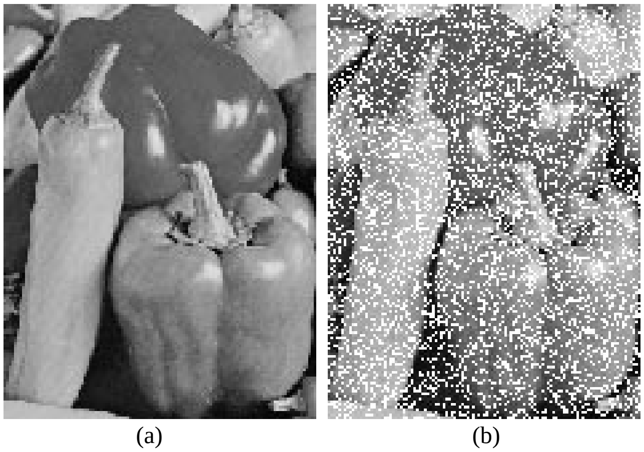

Figure 1. A 240x180 8-bit Peppers image. (a) The original image. (b) Noise-corrupted version of (a).

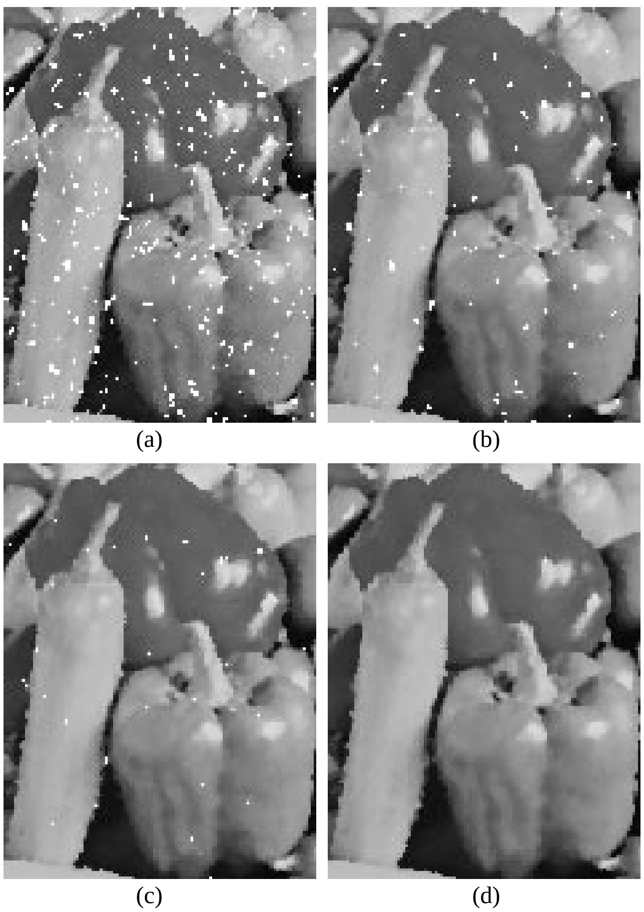

Figure 2. Processed images applying 5x5 WOS filters obtained by using the criteria of (a) MAE, (b) MSE, (c) L_5-norm error, and (d) L_8-norm error.

5. CONCLUSION

In this paper, a solution to the problem of designing WOS filters using an objective function given as the L_p norm of the error between the desired and estimated signals, has been formulated. It has been shown that in stack filtering, the L_p norm can be expressed as a linear combination of the decision errors incurred by the Boolean operators at each binary level of the filter. Based on this error formulation, nonadaptive and adaptive algorithms for the design of an L_p WOS filter have been developed. Simulation results show that an L_p WOS filter with $p \geq 2$ provides a better visual performance compared to that obtained by using an MAE WOS filter.

Table 1. Impulsive noise reduction with L_p WOS filters: MAE, MSE, L_5^5, and L_8^8 errors.

Filter	MAE	MSE	L_5^5 Error	L_8^8 Error
MAE WOS	10.74	9.9e+02	5.5e+09	5.2e+16
MSE WOS	9.28	5.7e+02	2.5e+09	2.2e+16
L_5 WOS	8.08	3.6e+02	1.0e+09	8.7e+15
L_8 WOS	6.82	2.1e+02	3.1e+08	2.1e+15

Table 2. Impulsive noise reduction with L_p WOS filters: $\tilde{J}_1(\mathbf{w})$, $\tilde{J}_2(\mathbf{w})$, $\tilde{J}_5(\mathbf{w})$, and $\tilde{J}_8(\mathbf{w})$ errors.

Filter	$\tilde{J}_1$	$\tilde{J}_2$	$\tilde{J}_5$	$\tilde{J}_8$
MAE WOS	19.6	2.66e+03	1.60e+10	1.46e+17
MSE WOS	19.8	2.62e+03	1.55e+10	1.40e+17
L_5 WOS	20.5	2.68e+03	1.52e+10	1.33e+17
L_8 WOS	22.5	2.71e+03	1.58e+10	1.29e+17

ACKNOWLEDGMENT

This work was partially supported by MICRONET, a Canadian Network of Centres of Excellence, under the project S1.MC1.

REFERENCES

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

[2] E. J. Coyle and J. H. Lin, "Stack filters and the mean absolute error criterion," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1244-1254, August 1988.

[3] L. Yin, J. T. Astola, and Y. A. Neuvo, "Adaptive stack filtering with application to image processing," *IEEE Trans. Signal Process.*, vol. 41, pp. 162-184, Jan. 1993.

[4] L. Yin, R. Yang, M. Gabbouj, and Y. Neuvo, "Weighted median filters: a tutorial," *IEEE Trans. Circuits Syst.—II*, vol. 43, pp. 157-192, March 1996.

[5] L. Yin, "Stack filter design: A structural approach," *IEEE Trans. Signal Process.*, vol. 43, pp. 831-840, April 1995.

[6] O. Y. Harja, J. T. Astola, and Y. A. Neuvo, "Analysis of the properties of median and weighted median filters using threshold logic and stack filter representation," *IEEE Trans. Signal Processing*, vol. 39, pp. 395-410, February 1991.

[7] I. Tabus, D. Petrescu, and M. Gabbouj, "A training framework for stack and Boolean filtering — fast optimal design procedures and robustness case study," *IEEE Trans. Image Processing*, vol. IP-5, pp. 809-825, June 1996.

[8] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1985.