Tampereen teknillinen korkeakoulu
Julkaisuja 75

Tampere University of Technology
Publication 75

Lin Yin

Adaptive Stack Filtering

Thesis for the degree of Doctor of Technology to be presented with due permission for public examination and criticism in Auditorium HB116 at Tampere University of Technology on the 14th of November, 1991, at 12.00 o'clock noon.

Tampere 1991

PREFACE

This study was carried out during the period 1989-1991 in the Signal Processing Laboratory, Department of Electrical Engineering, Tampere University of Technology.

I wish to express my sincere thanks to my thesis adviser, Professor Yrjö Neuvo, for his invaluable guidance and encouragement during the course of this study. I am also very grateful indeed to Professor Jaakko Astola for his help and advice.

In addition, I would also like to thank Dr. Olli Vainio for his friendly cooperation as well as all the members of the laboratory for the valuable support I received from them in an excellent working atmosphere.

Thanks are also due to Mr. John Leo for giving advice on the English, and to Professor Erkki Oja and Dr. Olli Yli-Harja for their technical review.

Finally I wish to thank my dear wife Qin Liu for all her support and patience in the course of my research work.

Tampere, Oct. 1991

Lin Yin

CONTENTS

LIST OF PUBLICATIONS

[1] O. Vainio, L. Yin, and Y. Neuvo, " On the properties and VLSI implementation of generalized median operations," *Proc. of 1990 IEEE Workshop on VLSI Signal Processing*, San Diego, CA, USA, November 7-9, 1990.

[2] L. Yin, J. Astola, and Y. Neuvo, "Optimal weighted order statistic filter under the mean absolute error criterion," *Proc. of ICASSP 91, The Int. Conferences on Acoustics, Speech, and Signal Processing*, 1991, Toronto, Canada, pp.2529-2532.

[3] L. Yin, J. Astola, and Y. Neuvo, " Adaptive weighted median filtering under the mean absolute error criterion," *Proc. of IEEE Workshop on Visual Signal Processing and Communications*, June 6-7, 1991, Hsinchu, Taiwan, pp.184-187.

[4] L. Yin, J. Astola, and Y. Neuvo, " Adaptive stack filtering with application to image processing," accepted by *IEEE Trans. on Signal Processing*, March 1991.

[5] L. Yin, J. Astola, and Y. Neuvo, " Adaptive multistage weighted order statistic filtering based on the back-propagation algorithm," submitted to *IEEE Trans. on Signal Processing*, April 1991.

[6] L. Yin, J. Astola, and Y. Neuvo, " Adaptive Boolean filters for noise reduction," *Proc. of Int. Conference on Artificial Neural Networks*, Espoo, Finland, June 24-28, 1991, pp. 759-764.

[7] L. Yin, J. Astola, and Y. Neuvo, "Adaptive neural filters," *First IEEE-SP Workshop on Neural Networks for Signal Processing*, Sept. 29-Oct. 2, 1991, Nassau Inn, Princeton, New Jersey, USA, pp.503-512.

[8] L. Yin, J. Astola, and Y. Neuvo, " A new class of nonlinear filters - neural filters," submitted to *IEEE Trans. on Signal Processing*, May 1991.

ABSTRACT

In this study, several efficient adaptive and nonadaptive algorithms have been developed for determining optimal stack filters including weighted median filters and weighted order statistic filters under the mean square error (MSE) criterion and the mean absolute error (MAE) criterion. The algorithms are based on the relationship between stack filters and threshold logic gates.

A new class of nonlinear filters called neural filters is also introduced. The neural filters defined by neural networks include all filters defined either by continuous functions such as linear FIR filters and microstatistic filters, or by Boolean functions such as median type filters and generalized stack filters. Adaptive least mean absolute error (LMA) and adaptive least mean square error (LMS) algorithms have been derived to determine optimal neural filters. In addition, as special cases, adaptive generalized stack and adaptive generalized weighted order statistic filtering algorithms under both error criteria have also been derived.

1. INTRODUCTION

Linear filters have been the dominating filter class throughout the history of signal processing. They are easy to implement and analyze and linear filters minimizing the mean square error (MSE) can usually be found in closed form. Furthermore, linear filters minimizing the MSE are optimal among the class of all filtering operations when the noise is additive and Gaussian. It is indeed unfortunate that, despite the elegant linear system theory, many signal processing problems cannot be satisfactorily addressed through the use of linear filters. In image processing applications, linear filters tend to blur the edges, fail to remove heavy tailed noise effectively, and do not perform well in the presence of signal dependent noise.

In order to overcome the shortcomings of linear filters, several new classes of nonlinear digital filters have been developed [Pit1, Coy2]. Perhaps the most popular class of nonlinear filters is the class of rank order-based filters. These filters perform well in many situations where linear filters fail. Various types of rank order-based filters have been proposed during the last decade and these filters usually fall into one of two categories. The first category is the set of hybrid order statistic filters which combine rank order filters with linear filters to exploit the desirable properties of both filter classes. Examples of this set would be FMH filters, Ll-filters, C-filters, and L-filters [Hei, Bov1, Gha, Pet, Pal1, Pal2]. There are techniques for finding the filter which is best suited for the application at hand and the algorithms are based on either the LMS algorithm or the back-propagation algorithm [Pal1, Pal2].

The second category is the set of generalized stack (GS) filters which exploit the generalized order statistic property, such as weighted median (WM) [Bro1, Yli1, Yli2, Pra], weighted order statistic (WOS) [Wen1, Yli1], morphological, multilevel max/median [Nie, Arc, Bon], stack [Wen1], and generalized stack (GS) filters [Lin1]. The fundamental difference between GS filters and other nonlinear filter classes is that the signal space is decomposed before filtering operations. This is called threshold decomposition and it maps the M-level valued input into $M-1$ binary signals by thresholding the original signal at each of the allowable levels. The set of $M-1$ signals is then filtered by $M-1$ Boolean operators which are constrained to have stacking property attributes. The multilevel output is finally obtained as the sum of the $M-1$ binary output signals.

The significance of GS filters is that optimal GS filters under the mean absolute error (MAE) criterion can be found via linear programming [Coy1, Lin1]. However, the computational complexity of this algorithm is very high since, for an N long filter, it requires the formulation of a linear program with $O(M(2I+3)^N)$ variables when $(2I+1)$ threshold levels are fed into the Boolean function at each level [Lin1]. This computational complexity renders the use of this approach extremely impractical. In order to alleviate this problem an adaptive approach has been developed where linear programming is avoided [Lin2]. However, the number of variables in the adaptive algorithm still increases exponentially with the window width. Thus, this algorithm is constrained to low values of I and small window sizes N.

In this study, we have developed several efficient stack filtering algorithms under the MAE and MSE criteria. These algorithms, similar in spirit to linear filtering algorithms, are based on the connection between stack filters and threshold logic gates. In addition, in order to unify linear FIR filters and median type filters, a new class of nonlinear filters called neural filters is introduced by combining in a specified manner threshold decomposition and neural networks. Neural filters include all linear FIR filters, microstatistic filters [Arc3], and GS filters. Adaptive algorithms have also been developed to determine optimal neural filters under the MAE criterion or MSE criterion.

The basic concepts of linear and median signal processing are reviewed briefly in Chapter 1, the emphasis being on issues related to this study. Special attention is paid to the threshold decomposition architecture of stack filters. The publications included in the thesis are summarized in Chapter 2.

1.1. Optimal Linear Filtering

Assume the process $X(n)$ at the input of a linear filter to be a corrupted version of the desired process $S(n)$. A window of width N slides across the input process $X(n)$. Let $\underline{X}(n)$ be the vector containing the N samples in the window at time n,

$$\underline{X}(n) = [X(n - N_1), X(n - N_1 + 1), \ldots, X(n + N_2)]^T \tag{1}$$

where ^{T} denotes the transpose, N_1, N_2 are integers, and $N_1 + N_2 + 1 = N$. The linear filter output is an estimate, called $\hat{S}(n)$, of the desired process $S(n)$

$$\hat{S}(n) = \underline{W}^T \underline{X}(n) \tag{2}$$

where $\underline{W}$ is the coefficient vector of the linear filter

$$\underline{W} = [W_1, W_2, ..., W_N]^T \tag{3}$$

The goal of optimal linear filtering is to find a linear filter that would minimize an error measure, typically the mean square error (MSE) or mean absolute error (MAE), between the filter's output and the desired signal. If $S(n)$ and $X(n)$ are joint stationary, then the MSE or MAE to be minimized is

$$\begin{aligned} J_\gamma(\underline{W}) &= E[|S(n) - \hat{S}(n)|^\gamma] \\ &= E[|S(n) - \underline{W}^T \underline{X}(n)|^\gamma] \end{aligned} \tag{4}$$

where $\gamma = 1, 2$ and $E[\cdot]$ denotes the expectation operation. $\gamma = 2$ corresponds to the MSE criterion while $\gamma = 1$ means the MAE criterion.

It is well known that the optimal linear filter under the MSE criterion can be found in closed form [Wid1, Hay], i.e.,

$$\underline{W}^* = R^{-1}\underline{R}_s \tag{5}$$

where R is an $N \times N$ autocorrelation matrix,

$$R = E[\underline{X}(n)\underline{X}^T(n)] \tag{6}$$

and $\underline{R}_s$ is an N-dimensional correlation vector

$$\underline{R}_s = E[\underline{X}(n)S(n)] \tag{7}$$

When the second-order statistics of the observed and desired signals are unavailable, optimal linear filters under the MSE criterion can be estimated by nonadaptive and adaptive algorithms based on a finite observation of the input signal and the desired output signal. Although the nonadaptive algorithm gives good results when the observed and desired signals are jointly stationary, adaptive algorithms save memory space, are easy to implement, and can track the time-varying statistics of the observed and desired signals.

A widely used and simple adaptive algorithm is the Widrow's LMS algorithm [Wid1, Hay],

$$\underline{W}(n) = \underline{W}(n-1) + \mu(S(n) - \underline{W}^T(n-1)\underline{X}(n))\underline{X}(n) \tag{8}$$

where μ is an adaptation step size. It has been proven that $\underline{W}(n)$ will converge to the optimal solution $\underline{W}^*$ under certain conditions [Abu, Feu, Jon, Kim, Kri1, Kri2].

Optimal linear filters under the MAE criterion cannot be found in closed form and the following adaptive sign algorithm has been proposed to find the solution [Ger]:

$$\underline{W}(n) = \underline{W}(n-1) + \mu \mathrm{sgn}(S(n) - \underline{W}^T(n-1)\underline{X}(n))\underline{X}(n) \tag{9}$$

where sgn(.) is a function defined as

$$\mathrm{sgn}(X) = \begin{cases} 1, & \text{if } X \geq 0; \\ -1 & \text{else.} \end{cases} \tag{10}$$

The adaptive sign algorithm is easier to implement than the adaptive LMS algorithm. However, its convergence speed is much slower. The convergence of the adaptive sign algorithm has also been proven [Ger, Cho].

1.2. Median Filters and Threshold Decomposition

Median filtering is a nonlinear technique that is commonly used for signal smoothing. The best known characteristic of median filters is their ability to remove noise while retaining sudden changes in the signal. Median filters are robust. For instance the magnitude of error in only a few of the input samples has only a small or no effect on the output signal.

1.2.1. Median Filters in the Real Domain

The median filter is implemented by sliding an odd-sized window over the input signal one sample at a time. At each window position the samples inside the filter window are sorted by magnitude, and the mid value, the median, is the filter output. This filtering procedure is denoted as

$$Y(n) = \mathrm{MED}[\underline{X}(n)] = \mathrm{MED}[X(n-N_1), ..., X(n), ..., X(n+N_2)] \tag{11}$$

where $X(n)$ and $Y(n)$ are the input and output sequences.

One major difference between linear filters and median filters is that median filters have root signals that are invariant for further filtering. This property is important from the signal processing point of view and it gives the possibility of characterizing median type filters [Gal, Arc4, Fit3, Wen3, Ast2]. Another way of analyzing median type filters is to study their deterministic properties [Mal2, Tya, Wen1, Fit4, Mar1, Mar2, Arc6] and their statistical properties [Jus, Vel, Ata1, Yan, Nod1, Bov2, Bov3, Bov4]. It has been shown that the median is the maximum likelihood estimate of signal level in the presence of independent and identically distributed additive Laplacian noise [Jus, Ast1]. To implement median filters, several efficient methods have been developed [Knu p.216, 218, Hua1, Ata2, Rao, Che, Ast3]. Median type filters have also been applied in 1-D, 2-D, and 3-D signal processing [Rab, Jay, Won, Hua2, Nar, Sch, Kir, Bov5, Gro, Sec, Car, Lou, Alp1, Alp2, Arc7, Arc8, Vie1, Vie2].

1.2.2. Threshold Decomposition and Median Filters in the Binary Domain

The superposition property of linear filters cannot be applied to nonlinear filters and this is the main reason why the theoretical analysis of median filters is somewhat difficult. However, the output of a median filter is decided solely on the basis of the ranks of the samples. This property makes it possible to use a powerful technique called *threshold decomposition* to divide the analysis problem into smaller parts. Using threshold decomposition we can derive all properties of median filters by just studying their effect on binary signals.

The threshold decomposition of an M-valued signal $X(n)$, where the samples are nonnegative integers, $0 \leq X(n) < M$, means decomposing $X(n)$ into $M-1$ binary signals, $x^1(n)$, $x^2(n)$, ..., $x^{M-1}(n)$, according to the following rule:

$$x^m(n) = T^m(X(n)) = \begin{cases} 1, & \text{if } X(n) \geq m \\ 0 & \text{else.} \end{cases} \tag{12}$$

It is important to note that this thresholding scheme can also be applied to non-integer signals or, in general, to all signals that are quantized to a finite number of arbitrary levels.

Fig. 1 *Median Filtering by Threshold Decomposition*

The original multi-valued signal can be reconstructed from its binary signals

$$X(n) = \sum_{m=1}^{M-1} x^m(n) \tag{13}$$

A very important property of median filters was observed by Fitch *et al* [Fit1]. They found that applying a median filter to an M-valued signal is equivalent to decomposing the signal to $M-1$ binary components, then filtering these components separately with the corresponding binary median filter using the algorithm in Eq. (11), and finally adding the binary signals together using Eq. (13). That is,

$$Y(n) = \mathrm{MED}[\underline{X}(n)] = \sum_{m=1}^{M-1} \mathrm{MED}[\underline{x}^m(n)] \tag{14}$$

where

$$\underline{x}^m(n) = [x^m(n-N_1), ..., x^m(n), ..., x^m(n+N_2)]^T \tag{15}$$

An example of median filtering by threshold decomposition is shown in Figure 1. The importance of this property arises from the fact that binary signals are much easier to analyze than multi-valued signals. The median operation on binary samples reduces to a simple Boolean operation. For example, the threevariable median filter corresponds to the Boolean operation $x_{-1}x_0+x_0x_1+x_{-1}x_1$ in the binary domain.

1.3. Weighted Median Filters

A natural extension of the median filter is the weighted median filter [Jus, Bro1, Bro2]. It has a set of weights which allow the control of filter behavior and has much greater flexibility in specification than the median filter.

1.3.1. Weighted Median Filters in the Real Domain

For the same input as for the linear filter in Eq. (1), the output $Y(n)$ of the WM filter of span N associated with the integer weights $\underline{W} = [W_1, W_2, ..., W_N]^T$ is given by

$$Y(n) = \mathrm{MED}[W_1 \diamond X(n-N_1), W_2 \diamond X(n-N_1+1), ..., W_N \diamond X(n+N_2)] \quad (16)$$

where $\diamond$ denotes the duplication

$$W \diamond X = \overbrace{X, ..., X}^{W \text{ times}} \quad (17)$$

The output of the WM filter for real positive weights can be calculated as follows: sort the sample inside the filter window, add up the corresponding weights from the upper end of the sorted set until the sum $\overline{W} \geq \frac{1}{2}\sum_{i=1}^{N} W_i$, then the output of the WM filter is the sample corresponding to the last weight.

WM filters can be easily extended to weighted order statistic (WOS) filters in the same manner as median filters are extended to rank order filters. The output $Y(n)$ of a WOS filter of span N associated with the integer weights $\underline{W}$ and integer threshold W_{N+1} is given by

$$Y(n) = W_{N+1} : th \text{ largest value of the set}$$

$$[W_1 \diamond X(n-N_1), W_2 \diamond X(n-N_1+1), ..., W_N \diamond X(n+N_2)] \quad (18)$$

WOS filters can also be computed for weights and thresholds having non-negative real values.

Fig. 2 *Linear separability of Boolean functions*

1.3.2. Weighted Median Filters in the Binary Doamin

WM filters and WOS filters can also be implemented using the threshold decomposition architecture method in the binary domain. To be more exact, WOS filters correspond to linearly separable positive Boolean functions (PBF's) in the binary domain. As shown in Fig. 2, a PBF $f(\underline{x})$ is said to be linearly separable if and only if it can be expressed in the form

$$f(\underline{x}) = U(F(\underline{x})) = \begin{cases} 1, & \text{if } F(\underline{x}) \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

where $U(.)$ is a unit step function,

$$U(X) = \begin{cases} 1, & \text{if } X \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{20}$$

and

$$F(\underline{x}) = \underline{W}^T \underline{x} - W_{N+1} \tag{21}$$

is the *separating function* of the Boolean function $f(\underline{x})$. $\underline{W}$ and $\underline{x}$ are N dimensional vectors

$$\underline{W} = [W_1, W_2, ..., W_N]^T \tag{22}$$

$$\underline{x} = [x_1, x_2, ..., x_N]^T \tag{23}$$

The components of $\underline{x}$ are understood here as real 1's and 0's.

It has been proven that the separating function corresponding to a PBF has all coefficients positive [Mur].

In the binary domain, weighted median filters correspond to linearly separable selfdual PBF's. A PBF $f(\underline{x})$ is said to be selfdual if and only if

$$f(\underline{x}) = \overline{f}(\overline{\underline{x}}). \tag{24}$$

1.3.3. Some Properties of Weighted Median Filters

In recent years, both the deterministic and statistical properties of WM filters have been extensively studied. For example, studies by Yli-Harja *et al* based on the threshold function representation of WM filters showed that the WM filters give an unbiased estimate of a constant in the same sense as the median does [Yli2]. Other studies by Wendt and Haavisto *et al* examined the root signal properties of WOS filters and WM filters [Wen2, Haa]. Yet another study by Prasad *et al*, without reference to threshold decomposition, discovered certain deterministic properties of WM filters which can be used to determine the equivalence between two WM filters [Pra]. In this subsection, however, we will only review two important properties of WM filters: the analogy between WM filters and FIR filters and the output distribution of WM filters.

A. Analogy between WM Filters and FIR Filters

There exists an interesting analogy between linear FIR filters and WM filters [Yli2]. In fact, the WM filter can be defined most naturally by replacing the "average" operator of a FIR filter with the "median" operator.

Let $H_i, i = 1, 2, ..., N$, be the impulse response coefficients of a 1-D FIR filter of size N. If $H_i \geq 0$ for all i, $1 \leq i \leq N$ and $\sum_{i=1}^{N} H_i = 1$ then the output Y_F of the filter is given by

$$Y_F = \sum_{i=1}^{N} H_i X_i = \text{average}\{W_1 \diamond X_1, W_2 \diamond X_2, ..., W_N \diamond X_N\} \tag{25}$$

where W_i, $1 \leq i \leq N$ are positive integers such that $H_i = W_i / \sum_{i=1}^{N} W_i$. If we replace the average in Eq. (25) with the median then the definition of the WM filter is obtained.

As is well known, the median is the maximum likelihood estimate of the signal level in the presence of uncorrelated additive biexponentially distributed noise, while the arithmetic mean is the same as for Gaussian distribution. This means that the median of the samples X_i is the value β that minimizes the expression

$$\Phi(\beta) = \sum_{i=1}^{N} |X_i - \beta|^{\gamma} \tag{26}$$

when $\gamma = 1$ and the arithmetic mean is the value β that minimizes (26) when $\gamma = 2$. If we introduce weights W_i to the expression (26) as follows

$$\Phi(\beta) = \sum_{i=1}^{N} W_i |X_i - \beta|^{\gamma} \tag{27}$$

then, in the case $\gamma = 2$, the minimizing value β can be expressed as

$$\beta = \sum_{i=1}^{N} W_i X_i / \sum_{i=1}^{N} W_i \tag{28}$$

which can be taken as a normalized FIR filter, i.e., a weighted average. Likewise, if $\gamma = 1$, the value β minimizing Eq. (27) is

$$\beta = \mathrm{MED}[W_1 \diamond X_1, ..., W_N \diamond X_N]$$

i.e., a weighted median.

The similarity between the WM filters and FIR filters suggests that WM filters may play as fundamental a role among nonlinear filters as FIR filters play among linear filters.

B. Statistical Properties of WM Filters

Since WM filters are highly nonlinear and thus have no frequency domain description, the most important description on the behavior of WM filters is provided by the output distribution function expressed in terms of input distributions. The output distribution of WM filters is given by the following theorem [Yli2]:

Theorem: Let the inputs of the WM filter be independent and identically distributed with the common distribution function $\Phi(t)$. Then the distribution function of the output is

$$\Psi(t) = \sum_{\underline{x} \in f^{-1}(0)} (1 - \Phi(t))^{H(\underline{x})} \Phi(t)^{N - H(\underline{x})} \tag{29}$$

where $f^{-1}(0) = \{\underline{x} \in \{0,1\}^N | f(\underline{x}) = 0\}$ and $H(\underline{x})$ denotes the Hamming weight of $\underline{x}$, i.e. the number of 1's in $\underline{x}$. For t_1 and t_2 such that

$$\Phi(t_1) = 1 - \Phi(t_2) \tag{30}$$

we also have

$$\Psi(t_1) = 1 - \Psi(t_2) \tag{31}$$

According to Eq. (31), it can be concluded that WM filters are statistically unbiased in the sense of the median, i.e., the median of the input is also the median of the output, which is to say, the WM filters also behave consistently for asymmetrical noise. In fact, it has been shown that all stack filters which are defined by selfdual PBF's are statistically unbiased in the sense of the median [Yli2]. In practice, particularly in image processing, the unbiasedness of stack filters is a typical constraint and is satisfied automatically. Therefore, optimal stack filters under an error criterion should be either WM filters or stack filters defined by nonlinearly separable selfdual PBF's.

1.4. Stack Filters

The concept of threshold decomposition and binary filtering has led to a general class of lowpass type filters, called stack filters [Wen1].

$$S_f(\underline{X}(n)) = \sum_{m=1}^{M-1} f(\underline{x}^m(n)) \tag{32}$$

where $\underline{X}(n)$ and $\underline{x}^m(n)$ are defined by Eq. (1) and Eq. (15), and $f(.)$ is the Boolean function used at each level of the architecture.

Only Boolean functions possessing the stacking property can be used in stack filtering. The stacking property ensures that the output is one of the input samples. The necessary and sufficient condition for a binary function to possess the stacking property [Gil] is that it can be expressed as a Boolean expression which contains no complements of input variables. Such functions are called PBF's. Thus, each PBF defines a stack filter. Stack filters can be implemented either using the threshold decomposition architecture in the binary domain or using a real domain architecture which is based on MAX/MIN operations. The real domain stack filter corresponding to a given PBF can be expressed by replacing Boolean AND- and OR- operations with MIN- and MAX-operations, respectively. For example, the three-point median over integer variables X_{-1}, X_0, X_1 is a stack filter defined by the PBF $f(x_{-1}, x_0, x_1) = x_{-1}x_0 + x_{-1}x_1 + x_0x_1$, corresponding to

$$\mathrm{MED}[X_{-1}, X_0, X_1] = \mathrm{MAX}[\mathrm{MIN}[X_{-1}, X_0], \mathrm{MIN}[X_{-1}, X_1], \mathrm{MIN}[X_0, X_1]]$$

However, for many simple filters, e.g. the median filter, the MAX-MIN representation becomes very complex as the number of variables increases. For the N-point median filter the number of MIN-operation is $\binom{N}{(N+1)/2}$.

One important property of stack filters is that any composition of stack filters again forms a stack filter. This is based on the fact that the composition of two PBF's is another PBF. Stack filters also possess close connection to mathematical morphology [Ser, Mar1, Mar2, Ste]. The basic operations of mathematical morphology, namely erosion and dilation (and their compositions), are in reality stack filter operations. The statistical properties of stack filters have been extensively studied by Wendt *et al* [Wen1]. Yli-Harja *et al* have proven that the output of a selfdual stack filter is statistically unbiased [Yli2].

Over the past several years, significant progress has been made in the theoretical analysis of stack filters. For instance, under the MAE criterion, optimal stack filters can be obtained as a solution to a linear program [Coy1, Coy2, Coy3, Gab]. In order to avoid modeling the signal and the noise and to reduce the amount of computation which is required by the linear programming algorithm, an adaptive stack filtering algorithm has been developed by Lin *et al* [Lin2]. However, all these optimal stack filtering algorithms are based on the estimation of the truth tables of Boolean functions and suffer from the common problem that the number of variables increases exponentially with the window width. This results in heavy computations and large memory requirements both during the filter training process and during the actual filtering process. In the filter training process, the memory required by stack filters increases with the number of elements in a window, N, as 2^N, and this has so far limited the size of the window to contain at most 18 elements. During the signal restoration process, stack filters are only implemented in the binary domain by looking up the truth tables of the Boolean functions since the number of elements in the truth tables is so large that their analytic expressions cannot be easily found. Even if we have the analytic expressions for the Boolean functions, implementation of stack filters is cumbersome since the complexity of the threshold decomposition architecture increases exponentially with the word length.

1.5 Generalized Stack Filters

The concept of threshold filtering can be generalized in two ways: (a) allowing different Boolean functions to operate at each thresholded level, and (b) allowing each of these Boolean functions to observe signals at several different levels. These generalizations lead to the concept of generalized stack filters.

Following the notation in [Lin1] we define the $(2I+1) \times N$ binary array at each level m,

$$\mathbf{x}^m(n) = \begin{pmatrix} \underline{x}^{m+I}(n) \\ \underline{x}^{m+I-1}(n) \\ \vdots \\ \underline{x}^{m-I}(n) \end{pmatrix} = \begin{pmatrix} x^{m+I}(n-N_1) & \dots & x^{m+I}(n+N_2) \\ x^{m+I-1}(n-N_1) & \dots & x^{m+I-1}(n+N_2) \\ \vdots & \vdots & \vdots \\ x^{m-I}(n-N_1) & \dots & x^{m-I}(n+N_2) \end{pmatrix} \tag{33}$$

where I is an integer and $m = 1, 2..., M-1$. In order to account for the upper I and lower I thresholded binary arrays we let $\underline{x}^m$ be a zero-valued vector for $m > M-1$ and a one-valued vector for $m < 1$.

The set of binary threshold arrays leads to the definition of GS filters, $g(.)$, as the stacking set of $M-1$ Boolean functions, where the input to the m'th Boolean function is the m'th binary threshold array of Eq. (33). The output is defined as

$$Y(n) = g(\underline{X}(n)) = \sum_{m=1}^{M-1} f^m(\mathbf{x}^m(n)) \tag{34}$$

where $\{f^1(.), ..., f^{M-1}(.)\}$ is called the stacking set of Boolean functions satisfying

$$f^{m+1}(\mathbf{x}^{m+1}) \leq f^m(\mathbf{x}^m), \quad m = 1, 2, ..., M-2. \tag{35}$$

The stacking constraint at the output ensures that the decisions at different levels are consistent. This means that if the filter on level m decides that the signal is less than m, then all the filters above it will make a compatible decision. Sufficient conditions for the set of Boolean functions in Eq. (35) to satisfy the stacking characteristics are simple to find. The necessary conditions are still unknown, except for the obvious case $I = 0$ [Wen1].

Optimal GS filters can also be found via linear programming when the goal is to minimize the MAE criterion [Lin1]. However, this truth table based optimal GS filtering algorithm suffers from a formidable computational complexity. In the study by Lin *et al* [Lin1], it has been shown that optimal GS filters under the MAE criterion can be found via a linear program with $O(M(2I+3)^N)$ variables. To illustrate this complexity, consider a signal with 256 levels of quantization and the design of a filter with window size 7. If $I = 1$, the number of variables is about $20,000,000$. Finding an optimal GS filter with a 3×3 window size (which is a small mask for image processing) is therefore quite impossible at the present time. Although the adaptive stack filtering algorithm can be extended to GS filtering, the number of variables in the adaptive GS filtering algorithm still increases exponentially with the window width.

2. SUMMARY OF PUBLICATIONS

This thesis consists of eight publications. They are divided into three parts.

Part one, which includes publications 1, 2, and 3, mainly considers the implementation and the optimal design of WOS filters. Part two, which includes publications 4 and 5, presents several efficient optimal stack filtering algorithms under the MAE and the MSE criteria. Part three, which includes publications 6, 7, and 8, introduces a new class of nonlinear filters called neural filters. Adaptive algorithms are therein developed for finding optimal neural filters under either the MAE criterion or the MSE criterion. As by-products, adaptive GS and generalized WOS (GWOS) filtering algorithms are also derived.

2.1. Weighted Order Statistic Filters: Implementation and Optimal Design

As with median filters, the traditional method of finding the output of WOS filters is sorting. The implementation of WOS filters requires an ordering operation to be applied to the samples inside the window. For large window sizes, the ordering procedure becomes cumbersome. In recent years, neural networks have attracted considerable attention as candidates for novel computational systems [Hop, Atk]. These types of large scale dynamical systems, analogous to biological structures, take advantage of distributed information processing and their inherent potential for parallel computation [Hop].

In Publication 1, a high-speed bit-sliced VLSI architecture and a neural network structure are presented for the implementation of WOS filters. The bit-sliced architecture is based on a theorem proved by Chen [Che]. In this architecture, the nonlinear operations are implemented at each bit level using positive Boolean functions operating on binary data. This is a considerable advantage as compared to many traditional approaches, e.g. the bubble sorter for median filtering. The resulting circuit architecture is modular, requires a relatively small number of transistors, and can be optimized for high throughput by pipelining.

The neural network structure is based on the deterministic property of WOS filters, which states that the output of WOS filters depends only on comparisons among the input samples rather than on their actual values. This neural network structure, consisting of three layers each of which operates locally, has the same properties as the MAXNET [Lip]. The synaptic weights of this structure are determined by the parameters of WOS filters. This structure implements weighted median, rank order, and median filters, and can be easily extended to a sorting neural network. In comparison to the other sorting neural network [Atk], this has the advantage of the input data being able to take arbitrary real values.

Among stack filters, WOS filters are quite attractive as they contain all extensively used median type filters: weighted median, rank order, and median filters. However, optimal WOS filters under the MAE criterion cannot be designed using the optimal stack filtering algorithms [Coy1, Coy2, Lin1, Lin2, Gab], as the necessary additional constraints are difficult to impose. In addition, these optimal stack filtering algorithms limit the window size to at most 18 due to their computational complexity and memory requirement.

In publication 2, an iterative algorithm is developed for determining optimal WOS filters under the MAE criterion. This algorithm is based on the relationship between WOS filters and threshold logic gates in the binary domain. With the aid of threshold decomposition, it is shown that optimal WOS filters under the MAE criterion in the real domain are equivalent to optimal Bayesian linear classifiers minimizing the classification error in the binary domain and satisfying the stacking constraints. The optimal WOS filters can be found by using iterative algorithms. This optimal WOS filtering algorithm is much more efficient than the adaptive stack filtering algorithm [Lin2]. This is because (1) the number of parameters to be estimated is equal to the window width of the filter, (2) the algorithm can be implemented in the real domain instead of the binary domain, and (3) the stacking constraints can be easily guaranteed by forcing all parameters of the WOS filter to be non-negative.

Among WOS filters, only WM filters, which are defined by linearly separable selfdual PFB's, give an unbiased estimate in the same sense as the median does [Yli2]. However, in general the previous optimal WOS filtering algorithm cannot directly find the optimal WM filters, nor can the nonadaptive WOS filtering algorithm track time-varying signals.

In publication 3, an adaptive algorithm is derived for finding optimal WM filters under the MAE criterion. This algorithm is based on the threshold function representation of WM filters and on the linearization of the unit step function. Although derived in the binary domain, the adaptive algorithm can be implemented in the real domain. This saves in computations and enables us to look into the relationship between optimal linear FIR filters and optimal WM filters.

2.2. Optimal Stack Filtering Algorithms under the MAE and the MSE Criteria

As stated earlier, the exponential increase of the variables in the traditional optimal stack filtering algorithms results in a rapid increase of computations and memory requirements [Coy1, Coy2, Lin2, Gab] and this has tended to limit the window size in adaptive stack filtering [Lin2].

In publication 4, we present a new stack filtering algorithm which is based on threshold decomposition and Bayesian decision theory. Any PBF with N variables can be represented by an Nth-order polynomial separating function. The problem of finding an optimal stack filter is equivalent to optimally estimating the coefficients of the corresponding separating function. We propose constrained nonadaptive and adaptive algorithms to estimate the parameters of the polynomial separating functions, i.e. stack filters. These algorithms can be implemented both in the binary domain and in the real domain although the algorithms are derived in the binary domain. The convergence of the adaptive algorithm is proven. More importantly, these algorithms enable one to approximate an optimal stack filter which has a high order separating function with a stack filter which has a lower order polynomial separating function. Experimental results demonstrate that in some situations the stack filters with lower order separating functions obtained by using the new stack filtering algorithms outperform even Lin's adaptive stack filters [Lin2]. This may be due to the premature termination of Lin's adaptive algorithm.

In principle, the extension to higher order separating functions is natural. However, the number of variables in the separating functions increases rapidly with the order of the polynomial. For example, there are $N(N+1)/2$ variables in quadratic separating functions. The computational burden thus becomes heavy for large window widths. In threshold logic, another efficient way to represent an arbitrary Boolean function is to use multistage threshold logic (MTL) gates [Lew]. MTL gates are also called multilayer perceptrons in neural network literature [Lip].

In publication 5, we use MTL gates to represent stack filters. It is shown that an MTL gate with positive weights in the binary domain corresponds to a multistage weighted order statistic (MWOS) filter in the real domain. MWOS filters are, in fact, generalized multilevel median filters, proposed first by Heinonen *et al* [Hei] and then by Arce *et al* [Arc1, Arc2]. As another form of stack filters, MWOS filters have several distinct advantages: 1) the number of variables needed to describe a stack filter is much less than the number of the elements in the truth table of the corresponding PBF; 2) MWOS filters can be implemented using sorting operations in the real domain rather than using the threshold decomposition architecture; 3) the relationship between MWOS filters and multilayer perceptrons in neural networks [Lip] opens up the possibility of using the existing training algorithms of neural networks to estimate MWOS filters.

In publication 5, two adaptive algorithms, the constrained least mean absolute back-propagation (CLMA-BP) algorithm and the constrained least mean square back-propagation (CLMS-BP) algorithm, are developed in the binary domain for estimating optimal MWOS filters under the MAE and the MSE criteria, respectively. These algorithms are almost the same as the back-propagation algorithm used in multilayer perceptrons [Rum]. The primary difference is that the positivity of the weights of MTL gates must be imposed. This is necessary for the maintenance of the stacking property, that is, to ensure that the logical operator at each level in the architecture is a PBF. In this sense the MWOS filters are special cases of multilayer perceptrons in the binary domain and the CLMA-BP and the CLMS-BP algorithms are special cases of the back-propagation algorithm. These relationships between stack filters and perceptrons are significant not only in the optimal design of stack filters under the MAE and the MSE criteria but also in revealing the robustness of perceptrons from the adaptive filtering point of view.

2.3. Adaptive Neural Filtering

The stacking property implies certain attractive properties. Relatively fast algorithms which are based on binary search exist for realizing stack filters [Wen1, Che]. Stack filters have root signals, i.e., signals which are invariant to filtering. These root signals provide valuable information on the behavior of the filter. On the other hand, the stacking property imposes certain limitations on the operations that can be performed using stack filters. Any stack filter is basically a smoother with unit gain. In many applications where there are tasks in which stack filters behave well, there are also tasks where highpass or bandpass type characteristics are needed, notably in detection and enhancement. In publication 6, using a neural network architecture, we define Boolean filters by dropping the requirement of positivity of Boolean functions. Two adaptive Boolean filtering algorithms, the LMS algorithm and the LMA algorithm, are derived for determining optimal Boolean filters under the MSE and the MAE criteria, respectively.

Usually, Boolean filters do not suppress Gaussian noise as effectively as linear filters. In publication 7, a new class of nonlinear filters called neural filters is introduced based on threshold decomposition and on neural networks with sigmoidal activation functions. The neural filters can approximate all the filters defined by continuous functions and Boolean functions, such as linear FIR filters and WOS filters, since sigmoidal functions can approximate both linear functions and the unit step function. An adaptive algorithm is derived for determining optimal neural filters under the MSE criterion. This algorithm is derived in the binary domain and is almost the same as the linear LMS algorithm. Experimental results demonstrate that if the signal is corrupted by Gaussian noise adaptive neural filters converge to linear filters and if the noise is impulsive, the optimal neural filters become Boolean filters or WOS filters.

In publication 8, the neural filters defined in publication 7 are generalized by allowing different neural networks at each threshold level. Two classes of neural filters are defined, hard neural filters and soft neural filters. The hard neural filters which perform Boolean operations are defined by a set of neural networks whose activation functions are unit step functions. When they satisfy the stacking property, the hard neural filters reduce to GS filters. A sufficient condition under which hard neural filters satisfy the stacking property is derived. This condition can greatly simplify the procedure of checking the stacking property of GS filters. The relationship between hard neural filters and median type filters is also explored. Any homogeneous hard neural filter which has the same neuron or neural network at each threshold level becomes a generalized weighted order statistic (GWOS) filter or a multistage GWOS filter when its weights are non-negative. Soft neural filters are defined by neural networks whose activation functions are sigmoidal. The universal approximation property of neural networks suggests that soft neural filters can approximate all filters defined by continuous linear and nonlinear functions such as linear FIR filters and microstatistic filters [Arc]. Moreover, soft neural filters can also approximate the hard ones because sigmoidal functions tend to be hard at "high gain" state, i.e., for large incoming weights. These approximation abilities of the soft neural filters are significant since they indicate that the soft neural filters could be adapted to various noise environments.

In publication 8, we derive two adaptive neural filtering algorithms, the adaptive least mean absolute (LMA) algorithm for finding the optimal neural filters under the MAE criterion, and the adaptive least mean square (LMS) algorithm for finding the optimal neural filters under the MSE criterion. Using the two algorithms we can train both hard neural filters and soft neural filters. These algorithms are analogous to Widrow's LMS algorithm [Wid1] or the back-propagation algorithm [Rum]. As by-products, adaptive GS and GWOS filtering algorithms under both error criteria are derived. These algorithms are much more efficient than Lin's adaptive stack filtering algorithm [Lin2] because of the reduction of the parameters to represent a GS filter and the simplification of the procedure of enforcing the stacking constraints of GS filters and GWOS filters. In addition, using the neural network representation we can derive optimal GS and GWOS filtering algorithms under the MSE criterion. This cannot be done using methods that are based on the truth table representation of Boolean functions.

2.4. The Author's Contribution to the Publications

In this thesis, the author's contribution to Publications 2 to 8 is essential in that he developed the theoretical framework, and performed the experiments. The main contribution of the author in each paper is as follows. In Publication 1, the author derives a neural network structure for implementing WOS filters in parallel. This structure can be used to implement median filters, rank order filters, and weighted median filters. In Publication 2, an iterative algorithm is developed by the author for determining optimal WOS filters under the MAE criterion and the convergence of this algorithm is also proven by the author. In Publication 3, the author derives an adaptive weighted median filtering algorithm under the MAE criterion.

In Publication 4, the author develops adaptive and nonadaptive stack filtering algorithms under the MAE criterion. The proof of the convergence of the adaptive algorithm is also provided by the author. In Publication 5, the author introduces multistage weighted order statistic filters. This class of filters can be used to represent stack filters efficiently. The author derives two adaptive algorithms for finding optimal multistage weighted order statistic filters under the MAE and the MSE criteria.

In Publication 6, the author introduces Boolean filters which are generalized from stack filters. According to the relationship between Boolean filters and neural networks, the author develops adaptive LMA and LMS Boolean filtering algorithms. In Publication 7, the author introduces a new class of nonlinear filters, referred to as neural filters. The neural filters, defined by neural networks in the threshold decomposition architecture, include all linear filters and Boolean filters. To find optimal neural filters under the MSE criterion, the author develops an adaptive algorithm. In Publication 8, the author generalizes the standard neural filters. The generalized neural filters include all linear filters, microstatistic filters, and generalized stack filters. The author also develops two adaptive neural filtering algorithms under the MAE criterion and the MSE criterion.

References

[Abu] S. Abu Ei Alta, " Asymptotic behavior of an adaptive estimation algorithm with application of M-dependent data," *IEEE Trans. Automata. Contr.*, vol. AC-27, pp. 1255-1257, Dec. 1982.

[Alp1] B. Alp, P. Haavisto, T. Jarske, K. Öistämö, and Y. Neuvo, "Median Based Algorithms for Image Sequence Processing," in *Visual Communications and Image Processing '90*, pp.122–134, Lausanne, Switzerland, October 1990.

[Alp2] B. Alp and Y. Neuvo, "3-Dimensional Median Filters for Image Sequence Processing," *Proc. IEEE Int. Conf. on ASSP*, pp. 2917–2920, Toronto, Canada, May 1991.

[Arc1] G.R. Arce and R.E. Foster, " Detail-preserving ranked-order based filter for image processing," *IEEE Trans. Acoust., Speech, and Signal Proc.* , vol. ASSP-37, Jan. 1989.

[Arc2] G.R. Arce and M.P. Mcloughlin, " Theoretical analysis of Max/Median filters," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol.ASSP-35, no.1, January, 1987.

[Arc3] G.R. Arce, " Micro statistics in signal decomposition and the optimal filtering problem," submitted to *IEEE Trans. Signal Processing*, Oct. 1990.

[Arc4] G.R. Arce and N.C. Gallager, Jr, " State description for the root-signal set of median filters," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-30, no.6, pp.894-902, Oct. 1982.

[Arc5] G.R. Arce, " Statistical threshold decomposition for recursive and nonrecursive median filters," *IEEE Trans., Information Theory*, vol. IT-32, no.2, pp. 243-253, March 1986.

[Arc6] G.R. Arce and R.L. Stevenson, " On the synthesis of median filter systems," *IEEE Trans. Circuits and Systems*, vol. CAS-34, no. 4, pp.420-429, April 1987.

[Arc7] G. R. Arce and E. Malaret, "Motion-Preserving Ranked-Order Filters for Image Sequence Processing," *Proc. IEEE Int. Symp. on Circuits and Systems*, pp. 983–986, Portland, Oregon, May 1989.

[Arc8] G. A. Arce, "Multistage Order Statistic Filters for Image Sequence Processing," *IEEE Trans. on Signal Processing*, vol. 39, no. 5, pp. 1146–1163, May 1991.

[Ast1] J. Astola and Y. Neuvo, " Optimal median type filters for exponential noise distributions " *Signal Processing*, vol. 17, no.2, June 1989, pp. 95-104.

[Ast2] J. Astola, P. Heinonen, and Y. Neuvo, " On root structures of median and median type filters," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-35, no.8, pp.1199-1201, Aug. 1987.

[Ast3] J. Astola and G. Campbell, " On the computation of running median," *IEEE Trans. Acoust., Speech, and Sig. Proc.*, vol.ASSP-37, pp.572-574, Apr. 1989.

[Ata1] E. Ataman, V.K. Aatre, and K.M. Wong, " Some statistical properties of median filters," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-29, no.5, pp. 1073-1075, Oct. 1981.

[Ata2] E. Ataman, V.K. Aatre, and K.M. Wong, " A fast method for real-time median filtering," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. ASSP-28, no.4, pp. 415-421, Aug. 1980.

[Atk] M. Atkins, " Sorting by Hopfield Net," *Proc. International Joint Conf. on Neural Network*, pp.II-65-68, 1989.

[Bov1] A.C. Bovik, T.S. Huang and D.C. Munson Jr., " A generalization of median filtering using linear combinations of order statistics," *IEEE Trans. on*

Acoustics, Speech and Signal Processing, vol.ASSP-31, no.6, pp.1342-1350, Dec. 1983.

[Bov2] A.C. Bovik, T.S. Huang, and D.C. Munson, Jr., "Nonparametric tests for edge detection in noise," *Pattern Recognition,* vol.19, pp.209-219, 1986.

[Bov3] A.C. Bovik, "Streaking in median filtered images," *IEEE Trans. Acoust., Speech, Signal Processing,* vol.ASSP-35, pp.493-503, Apr. 1987.

[Bov4] A.C. Bovik, T.S. Huang, and D.C. Munson, Jr., "The effect of median filtering on edge estimation and detection," *IEEE Trans. Pattern Anal. Machine Intell.,* vol.PAMI-9, pp.181-194, Mar. 1987.

[Bov5] A.C. Bovik, T.S. Huang, and D.C. Munson, Jr., "Edge-sensitive image restoration using order-constrained least squares methods," *IEEE Trans. Acoust., Speech, Signal Processing,* vol.ASSP-33, pp.1253-1263, Oct. 1985.

[Bon] C. G. Boncelet Jr, R. Hakami, and P. J. Warter, " Some theory of multistage order statistic filters," *Proc. Midwest Conference on Circuits and System*, pp.59-61, 1990.

[Bra] M.L. Brady, R. Raghavan, and J. Slawny, " Back propagation fail to separate where perceptrons succeed," *IEEE Trans. on Circuits and Systems*, vol.36, no.5, pp.665-674, May 1989.

[Bro1] D.R.K. Brownrigg, " The weighted median filter," *Commun. ACM*, vol.27, pp. 807-818, Aug. 1984.

[Bro2] D.R.K. Brownrigg, " Generation of representative members of an RrSst weighted median filter class," *IEE Proc. Part.F, Comm., Radar, and Signal Proc.*, vol.133, no.5, pp. 445-448, Aug. 1986.

[Car] R. Carla, V.M. Sacco, and S. Baronti, "Digital techniques for noise reduction in APT NOAA satellite images," in *Proc. IGARSS-86,* Zürich, Switzerland, pp.995-1000, Sept. 1986.

[Che] K. Chen, "Bit-serial realizations of a class of nonlinear filters based on positive boolean functions," IEEE Transactions on Circuits and Systems, Vol. 36, No. 6, pp. 785-794, June 1989.

[Cho] S.H. Cho and V.J. Mathews, " Tracking analysis of the sign algorithm in nonstationary environments ," *IEEE Trans. Acoustics, Speech, and Signal*

Processing, vol.ASSP-38, no.12, pp.1244-1254, Dec. 1990.

[Coy1] E. J. Coyle and J.H. Lin, " Optimal stacking filters and mean absolute error nonlinear filtering," *IEEE Trans. Acoust., Speech, and Signal Proc.* , vol.ASSP-36, no.8, pp.1244-1254, Aug. 1988.

[Coy2] E. J. Coyle, J-H Lin, and M. Gabbouj, " Optimal stack filtering and the estimation and structural approaches to image processing," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-37, no.12, pp. 2037-2066, Dec. 1989.

[Coy3] E.J. Coyle and N.C.Jr. Gallagher, " Stack filter and neural networks," *Proc. IEEE Int. Symp. Circ., Syst., ISCAS-89*, pp.995-998, 1989.

[Eya] Eyal Yair and Allen Gersho, " The Boltzmann perceptron network: a soft classifier," *Neural Networks*, vol.3, pp.203-221, 1990.

[Feu] A.Feuer and E.Weinstein, " Convergence analysis of LMS filter with uncorrelated Gaussian data," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-33, pp. 220-230, Feb. 1985.

[Fit1] J.P.Fitch, E.J. Coyle, and N.C. Gallagher, Jr., " Median filtering by threshold decomposition," *IEEE Trans. Acoustic., Speech, Signal Processing,* vol.ASSP-32, pp.1183-1188, Dec. 1984

[Fit2] J.P.Fitch, E.J. Coyle, and N.C. Gallagher, Jr., " Threshold decomposition of multidimensional rank order operations," *IEEE Trans. Circuits and Systems*, vol.CAS-32, pp.445-450, May 1985.

[Fit3] J.P.Fitch, E.J.Coyle, and N.C.Gallagher,Jr., "Root properties and convergence rates of median filters," *IEEE Trans. Acoust., Speech, Signal Processing,* vol.ASSP-33, pp.230-240, Feb. 1985.

[Fit4] J.P. Fitch, E.J. Coyle, and N.C. Gallagher, Jr., "The analog median filter," *IEEE Trans. Circuits Systems,* vol.CAS-33, pp.94-102, Jan. 1986.

[Gab] M.Gabbouj and E.J.Coyle, " Minimum mean absolute error stack filtering with structural constraints and goals," *IEEE Trans. on Acoust., Speech and Signal Processing*, vol. ASSP-38, no.6, pp. 955-968, June, 1990.

[Gal] N.C. Gallager Jr. and G.L. Wise, " A theoretical analysis of the properties of median filters," *IEEE Trans. Acoustics, Speech, and Signal Processing,* vol. ASSP-29, no.6, pp.1136-1141, Dec. 1981.

[Ger] A. Gersho, " Adaptive filtering with binary reinforcement," *IEEE Trans. Information Theory*, Vol. IT-30, no.2, pp.191-199, March 1984.

[Gha] P.P. Ghandi, S.R. Peterson, and S.A. Kassam, " Frequency/selective signal restoration using nonlinear combination filters," in *Proc. 1988 Int. Conf. Acoust., Speech, Signal Processing*, New York, NY, pp.1522-1525, Apr. 1988.

[Gil] E. N. Gilbert, " Lattice theoretic properties of frontal switching functions," *J. Mathm. Physics*, vol. 33, pp.57-67, 1954.

[Gro] W.Grochulski, P.Mitraszewski, and P.Penczek, "Application of combined median-averaging filters to scintigraphic image processing," *Nucl. Med.,* No.24, pp.164-168, 1985.

[Haa] P. Haavisto, M. Gabbouj, and Y. Neuvo, " Median based idempotent filters," submitted to *Journal of Circuits, Systems, and Computers.*

[Hay] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, 1986.

[Hei] P. Heinonen and Y. Neuvo, " FIR median hybrid filter," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-25, pp. 832-838, June 1987.

[Hop] J. J. Hopfield, " Neural Networks and Physical Systems with Emergent Collective Computational Abilities," *Proc. Natl. Acad. Sci. USA, Biophysics*, vol.79, pp.2554-2558, April 1982.

[Hu] S.T.Hu, *Threshold Logic*, University of California Press, Berkeley 1965.

[Hua1] T.S. Huang, G.J. Yang, and G.Y. Tang, "A fast two-dimensional median filtering algorithm," *IEEE Trans. Acoust., Speech, Sig. Proc.,* vol.ASSP-27, pp.13-18, Feb. 1979.

[Hua2] T.S. Huang, and G.J. Yang, "Median filters and their applications to image processing," School Elec. Eng., Purdue Univ., West Lafayette, IN, Tech. Rep. EE80-1, Jan. 1980.

[Jay] N.S. Jayant, "Average- and median-based smoothing techniques for improving digital speech quality in the presence of transmission errors," *IEEE Trans. Commun.,* vol.COM-24, pp.1043-1045, Sept. 1976.

[Jon] S.K.Jones, R.K.Cavin, III, and W. M. Reed, " Analysis of error-gradient adaptive linear estimator for a class of stationary dependent processes," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 431-437, July 1975.

[Juh] J. Juhola, P. Haavisto, O. Vainio, T. Raita-aho, and Y. Neuvo, "On VLSI implementation of median based field rate up-conversion," *in Proc. IEEE 1990 International Symposium on Circuits and Systems*, New Orleans, LA, USA, May 1-3, 1990.

[Jus] B.I. Justusson, " Median filtering: statistical properties," *Two-Dimensional Digital Signal Processing II: Transforms and Median Filters*, Chap.4, vol.42, pp. 161-196. *Topics in Applied Physics*, T.S. Huang, Ed., New York: Springer Verlag, 1981.

[Kim] J.K.Kim and L.D.Davisson, " Adaptive linear estimation for stationary M-dependent processes," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 23-31, Jan. 1975.

[Kir] R.L. Kirlin, R. Hutchkins, B. Cudzilo, L.A. Dewey, and L.L. Hailey, " The enhancement of seismogram parameters using image processing techniques," *Geoexploration*, no.23, pp. 41-76, 1984/1985.

[Knu] D.E. Knuth, *Sorting and Searching*, vol.3 of *The Art of Computer Programming*, Addison-Wesley, Reading, Mass., 1985.

[Kri1] A.Krieger and E.Masry, " Convergence analysis of adaptive linear estimation for dependent stationary processes," *IEEE Trans. on Information Theory*, vol. 34, no.4, pp. 642-654, July 1988.

[Kri2] A.Krieger and E.Masry, " Constrained adaptive filtering algorithms: asymptotic convergence properties for dependent data," *IEEE Trans. Information Theory*, vol. 35, no.6, pp. 1166-1176, Nov. 1989.

[Lee] Y. H. Lee and A. T. Fam, " An edge gradient enhancing adaptive order statistic filter," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol.ASSP-35, no.5, May 1987.

[Lew] P. M. Lewis II. and C.L. Coates, *Threshold Logic*, John Wiley & Son, New York, 1967.

[Lin1] J.H. Lin and E.J.Coyle, " Generalized stack filters and mean absolute error nonlinear filtering," *IEEE Trans. on Acoust., Speech, and Signal Processing*

, vol. ASSP-38, no.6, pp. 938-954, June, 1990.

[Lin2] J.H.Lin, T.M.Sellke and E.J.Coyle, " Adaptive stack filtering under the mean absolute error criterion," *IEEE Trans. Acoust., Speech, and Signal Proc.*, vol. ASSP-38, no.6 , pp. 938-954, June 1990.

[Lip] R. P. Lippmann, " An introduction to computing with neural nets, " *IEEE ASSP magazine*, pp.4-22, April 1987.

[Lou] T. Loupas, N. McDicken, and P.I. Allan, "Noise reduction in ultrasonic images by digital filtering," *Br.J.Radiology,* vol.60, pp.389-392, Apr. 1987.

[Mal1] C. L. Mallows, " Resistant Smoothing," in *Time Series*, O. D. Anderson Editor, North Holland, pp. 147-155, 1980.

[Mal2] C.L.Mallows, "Some theory of nonlinear smoothers," *The Annals of Statistics,* vol.8, No.4, pp.695-715, 1980.

[Mar1] P. Maragos and R.W. Schafer, " Morphological Filters - Part I: Their set theoretic analysis and relations to linear shift-invariant filters," *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. ASSP-35, No.8, pp. 1153-1169, Aug. 1987.

[Mar2] P. Maragos and R.W. Schafer, " Morphological Filters - Part II: Their relations to linear median, order-statistic, and stack filters," *IEEE Trans. Acoust., Speech, Signal Proc.*, vol. ASSP-35, No.8, pp. 1170-1184, Aug. 1987.

[Mur] S. Muroga, *Threshold Logic and Its Applications*, Wiler Interscience, New York 1971.

[Nar] P.M. Narendra, "A separable median filter for image noise smoothing," *IEEE Trans. Pattern Anal. Machine Intell.,* vol.PAMI-3, pp.20-29, Jan. 1981.

[Nie] A. Nieminen, P. Heinonen, and Y. Neuvo, " A new class of detail-preserving filters for image processing," *IEEE Trans. Pattern Anal. Machine Intell.*, vol.PAMI-9, Jan. 1987.

[Nod1] T.A. Nodes, " Theoretical results on the properties of median type operations," *Ph. D. Dissertation*, School of Electrical Engineering, Purdue University, West Lafayette, IN, Dec. 1982.

[Nod2] T.A. Nodes and N.C. Gallagher " Median filters: some modifications and their properties," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol.ASSP-30, no.5, pp. 739-746, Oct. 1982.

[Pal1] F. Palmieri and C.G. Boncelet Jr., " Ll-filter - a new class od order statistic filters," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol.ASSP-37, no.5, pp. 691-701, May 1989.

[Pal2] F. Palmieri, " A backpropagation algorithm for multilayer hybrid order statistic filters," in *Proc. 1989 IEEE Int. Conf. Acousti., Speech and Signal Processing*, Glasgow, Scoland, pp.1179-1182, 1989.

[Pet] S.R. Peterson and S.A. Kassam, " General order statistic-based structures for nonlinear signal processing," in *Proc. 1986 Conf. Inform. Sci. Syst.,* Princeton, NJ, pp.723-728, Mar. 1986

[Pit1] L. Pitas and A.N. Venetsanopoulos, *Nonlinear Digital Filters: Principles and Applications*, Kluwer Academic Publisher, 1990.

[Pit2] L. Pitas and A N. Venetsanopoulos, " Nonlinear mean filters in image processing," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol.ASSP-34, no.3, June 1986.

[Pra] Mohit K Prasad and Yong H Lee, " Weighted median filters: generation and properties," *Proc. IEEE Int. Symp. Circ. and Syst., ISCAS-89*, pp.425-428, 1989.

[Rab] L.R. Rabiner, M.R. Sambur, and C.E. Schmidt, "Applications of a nonlinear smoothing algorithm to speech processing," *IEEE Trans. Acoust., Speech, Signal Processing,* vol.ASSP-23, pp.552-557, Dec. 1975.

[Rao] V.B. Rao and K.S. Rao, "A new algorithm for real time median filtering," *IEEE Trans. Acoust., Speech, Signal Processing,* vol.ASSP-32, no.4, pp.1674-1675, Dec. 1986.

[Rum] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, " Learning internal representations by error propagation ", in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol.1, Ed. D. E. Rumelhart and J. L. McClelland, MIT Press, pp.318-362, 1986.

[She] C.L. Sheng, *Threshold Logic*, Academic Press, New York 1969.

[Sch] R.M. Schmitt, C.R. Meyer, P.L. Carson, and B.I. Samuels, "Median and spatial low-pass filtering in ultrasonic computed tomography," *Med. Phys.,* vol.11, pp.767-771, Dec. 1984.

[Sec] N. Secord and K.M. Wong, "Median filtering of AMBTC images for improved image compression," in *Proc. of IEEE Electronicom-85*, Toronto, Canada, pp.190-193, Oct. 1985.

[Ser] J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, London, 1982.

[Sny] D.L. Snyder, *Random Point Processes*, New York: Wiley, 1975.

[Ste] S.R. Sternberg, " Grayscale Morphology," in S.R. Sternberg and J. Serra ed.: " Special Section on Mathematical Morphology," *Computer Vision, Graphics and Image Processing*, 35, pp.333-355, 1986

[Tya] S.G. Tyan, "Median filtering: Deterministic properties," in *Topics in Applied Physics,* vol.43, *Two-Dimensional Digital Signal Processing 2,* T.S. Huang, Ed. Berlin, Germany: Springer-Verlag, 1981, vol.43, pp.197-217.

[Vel] P.F.Velleman, "Definition and comparison of robust nonlinear data smoothing algorithms," *J.Amer. Statist. Ass.,* pp.609-615, Sept. 1980.

[Vie1] T. Viero and Y.Neuvo, "Non-Moving Regions Preserving Median Filters for Image Sequence Filtering," *Proc. IEEE Int. Conf. on Systems Engineering*, pp. 245–248, Dayton, Ohio, August 1991.

[Vie2] T. Viero, K. Öistämö, and Y. Neuvo, "Median and Vector Median Filters for Image Sequence Filtering," *submitted to European Conference on Circuit Theory and Design*, Copenhagen, Denmark, Sept. 1991.

[Wen1] P.D. Wendt, E.J. Coyle and N.C. Jr. Gallagher, " Stack filter", *IEEE Trans. Acoust., Speech, Signal Proc.,* vol.ASSP-34, no.4, pp.898-911, Aug. 1986.

[Wen2] P.D. Wendt, " Convergence properties of multi-dimensional stack filters," *Proc. of the 1990 SPIE/SPSE Symp. on Electronic Imaging*, Feb. 1990, Santa Clara, CA.

[Wen3] P.D. Wendt, E.J. Coyle, and N.C. Gallagher, Jr., "Some convergence properties of median filters," *IEEE Trans. Circuits and Systems,* vol.CAS-33, pp.276-286, Mar. 1986.

[Wid1] B.Widrow and, S.D.Stearns, *Adaptive Signal Processing.* Englewood Cliffs, NJ: Prentice - Hall, 1985.

[Wid2] B. Widrow, R. G. Winter, and R. A. Baxter, " Layered neural nets for pattern recognition," *IEEE Trans. Acoust., Speech, Signal Processing*, vol.36, pp. 1109-1118, July 1988

[Wid3] B. Widrow and R. Winter, " Neural nets for adaptive filtering and adaptive pattern recognition," *Computer*, pp. 25-39, March 1988.

[Won] K.M. Wong and S. Chen, "Detection of narrow-band sonar signals using order statistical filters," *IEEE Trans. Acoust., Speech, Signal Processing,* vol.ASSP-35, pp.597-613, May 1987.

[Yan] G.J.Yang and T.S.Huang, "The effect of median filtering on edge location estimation," *Computer Graphics and Image Processing,* vol.15, pp.224-245, Mar. 1981.

[Yli1] O.Yli-Harja, J.Astola and Y.Neuvo," A response-based design method for weighted median filters," *IEEE Int. Symp. Information Theory, ISIT - 88*, Kobe, Japan, June 1988.

[Yli2] 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. on Signal Processing* , vol. 39, no.2 pp.395-410, Feb. 1991.

[Yu1] Pao-Ta Yu and E.J. Coyle, " On the existence and design of the best stack filter based on associated memory," *Proc. 1990 Int. Symp. on Circuits and Systems*, pp.2495-2499, New Orleans, LA, May 1990.