Military Technical College
Kobry El-Kobbah
Cairo, Egypt

10th International Conference
On Aerospace Sciences&
Aviation Technology

Adaptive Weighted Median Filter For effective Impulse Noise Filtering

Mohiy M. Hadhoud*, Mohamed A El-Latif **,
Ehsan M Sabek, Hossam A El-Salam Diab****
*Faculty of Computers and Information, Menufia University.
**Faculty of Science, Menufia University.

Abstract:

This paper proposes a gradient–based adaptive weighted median (AWM) algorithm. The AWM is a modification of normal LMS, obtained by applying a proposed adaptive weighted median filter on the input signal instead of using (FIR) filter. The (AWM) algorithm is designed to facilitate adaptive filter performance close to the least squares optimum across a wide range of inputs. Also, we present an analysis of the (AWM) algorithm, we using the threshold decomposition technique that admitting real-valued signals. This threshold decomposition is used to develop LMS algorithm to optimally design the filter's coefficients to obtain our proposed filter, and Results are presented to illustrate the performance of the proposed algorithm and its application in noisy image filtering.

I- Introduction

To overcome the limitation of linear filters, various nonlinear filtering techniques have been proposed. Among those, the filters based on order statistics have found considerable attention due to their ability to reject outliers, closely track signal discontinuities, and effectively preserve signal details [1]. Adaptive signal processing, and particularly adaptive filtering, provides powerful approach to many signal processing problems [2]. The capacity of adaptive algorithms to operate when limited prior information is available makes them ideally matched to many practical applications. The most common adaptive filtering algorithm is the least mean squares (LMS) algorithm.

This paper proposes a combined approach that combines both the adaptive LMS algorithm and nonlinear weighted median techniques to obtain an algorithm named AWM algorithm. In order to design the proposed filter the threshold decomposition property is used [7],(this property exploits the weak superposition property described in appendix(1))that have the following advantage, "the analysis of median smoothing binary signals is much easier than the analysis of median smoothing real-valued signals" (see section III),this method in designing the proposed filter achieves the best results in filtering noisy images if it compared with other algorithms in case of salt & pepper noise and mixed of both Gaussian and salt & pepper noise. In this paper we first present a review of some non-linear image filtering methods based on median and other order statistics filters. Second, a review of the common LMS algorithm. And then, a review of threshold decomposition of real valued signals. The proposed combination is presented in section (v) and results are presented in section (vi) followed by conclusions and comments in section (vii).

II- Preliminaries

A- Running median smoothers:

The running median was first suggested as a non-linear smoother for time series data by TUKEY in 1974. To define the running median smoother :- Let $\{x(\,.\,)\}$ be a discrete time sequence, the running median passes a window over the sequence $\{x(\,.\,)\}$ that selects, at each instant n, a set of samples to comprise the observation vector $X(n)$ such that:

$$X(n) = [x(n-N_L),\ldots, x(n),\ldots, x(n+N_R)] \tag{1}$$

Where N_L and N_R may range in value over the non-negative integers and $(N = N_L + N_R + 1)$ is the window size. The median operating on the input sequence $\{x(\,.\,)\}$ produces the output sequence $\{Y(\,.\,)\}$, where at time index n

$$Y(n) = Median\ [x(n-N_L),\ldots, x(n),\ldots, x(n+N_R)] = Median\ [x_1(n),\ldots, x_N(n)] \tag{2}$$

where $x_i(n) = x(n - N_L - 1 + i)$ *for* $i = 1, 2, \ldots, N$ that is, the samples in the observation window are sorted and the middle or median value is taken as the output. If $x_{(1)},x_{(2)},\ldots,x_{(N)}$ are the sorted samples in the observation window, the median smoother outputs, is given by:

$$Y(n)=\begin{cases} x\left(\frac{N+1}{2}\right) & \text{if } N \text{ is odd} \\ \dfrac{x\left(\frac{N}{2}\right)+x\left(\frac{N}{2}+1\right)}{2} & \text{othewise} \end{cases} \tag{3}$$

The performance of running median is limited by the fact that it is temporally blind that is all observation samples are treated equally regardless of their location within the observation window.

B- Weighted median smoothers

Weighted median (WM) smoothers have received considerable attention in signal processing research over the last two decades [4]-[6].It is often stated that there are many analogies between weighted median smoothers and linear FIR filters, however, it was shown that WM smoothers are, highly constrained, having significantly less-powerful characteristics than linear FIR filters. In fact, WM smoothers are equivalent to normalized weighted mean filters admitting only positive weights. Weighted mean are, in essence, restricted to "low pass" type filtering characteristics.

In order to define the WM smoother, it is best to first recast the similarities between linear FIR filters and WM smoother. Given an observation set $x_1,x_2,\ldots,x_N$, the sample mean $\bar{\beta}=Mean(x_1,x_2,\ldots,x_N)$ can be generalized to linear FIR filters as :

$$\bar{\beta}=Mean(w_1.x_1,w_2.x_2,\ldots,w_N.x_N) \text{ Where } w_i \in R \tag{4}$$

It was shown in [7] that the sample median $\hat{\beta}=Median\ (x_1,x_2,\ldots,x_N)$ that plays an analogous role to the sample mean in location estimation can be extended to: -

$$\hat{\beta}=Median\ (w_1\Diamond x_1,w_2\Diamond x_2,\ldots,w_N\Diamond x_N), \tag{5}$$

And $\Diamond$ is the replication operator defined as :- $w_i\Diamond x_i = \underbrace{x_i,x_i,\ldots,x_i}_{w_i \text{ times}}$

The WM smoother operation can be schematically described as shown in figure (1).

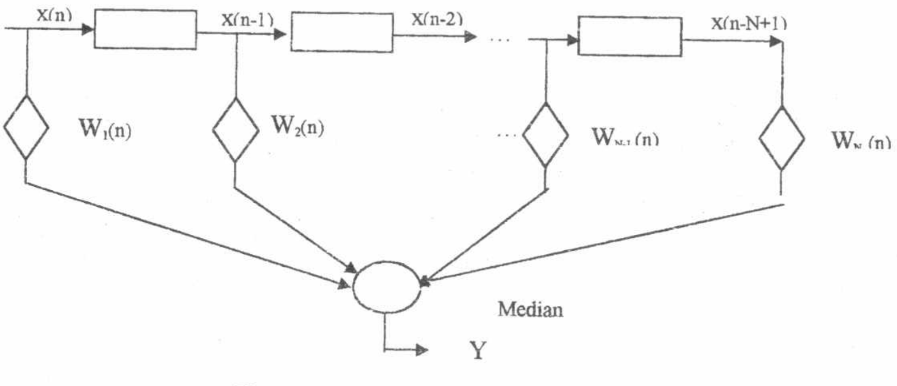


Figure (1) The weighted median smoothing operation

C- The center weighted median smoother

The weighting mechanism of WM smoothers allows for great flexibility in emphasizing or de-emphasizing specific input samples, in most applications, not all samples are equally important because of the symmetric nature of the observation window, the sample most correlated with the desired estimate is, in general, the center observation sample. This observation leads to the center-weighted median (CWM) smoother which is a relatively simple subset of WM smoother that has proven useful in many applications [8]. The CWM smoother is realized by allowing the center observation sample to be weighted, thus, the output of the CWM smoother is given:

$$Y(n) = Median\,[x_1,\ldots,x_{c-1}, w_c \Diamond x_c, x_{c+1},\ldots,x_N] \qquad (6)$$

Where w_c is an odd positive integer and $c = \frac{N+1}{2}$ is the index of the center sample, when $w_c=1$ the operator is a median smoother.

D- Weighted median filters

Admitting only positive weights, WM smoothers are severely constrained as they are, in essence, smoothers having Low-pass type filtering characteristics, linear FIR equalizers admitting only positive filter weights, for instance, would lead to completely unacceptable results, thus, it is not surprising that weighted median smoothers admitting only positive weights lead to unacceptable results in a number of applications [9]. Much like the sample mean can be generalized to the rich class of linear filters, there is a logical way to generalize the median to an equivalently rich class of weighted median filters that admit both positive and negative weights [7]. The sample mean as explained before, can be generalized to the class of linear FIR filters given in equation (3). In order for the analogy to be applied to the median filter structure, the equation (3) must be written as :-

$$\overline{\beta} = Mean(|w_1|\operatorname{sgn}(w_1)x_1, |w_2|\operatorname{sgn}(w_2)x_2,\ldots,|w_N|\operatorname{sgn}(w_N)x_N) \qquad (7)$$

where sgn (.) denotes the signum function defined as: -

$$\operatorname{sgn}(w_i) = \begin{cases} 1 & if\ w_i \geq 0 \\ -1 & otherwise \end{cases} \qquad (8)$$

The sgn (.) of the weight affects the corresponding input sample and the weighting is constrained to be non-negative. By analogy, the class of weighted median filters admitting real-valued weights emerges as [7]:

$$\hat{\beta} = Median\,(|w_1|\Diamond\operatorname{sgn}(w_1)x_1, |w_2|\Diamond\operatorname{sgn}(w_2)x_2,\ldots,|w_N|\Diamond\operatorname{sgn}(w_N)x_N)\ ,w_i \in R \qquad (9)$$

The weight signs are uncoupled with the weight magnitude values and are merged with the observation samples. The weight magnitudes play the equivalent role of positive weights in the framework of weighted median smoothers [6]. Although the filter weights may seem restricted to integer values, the WM filter clearly allows for real-valued weights.

E- Adaptive FIR filters: -

An adaptive filter is essentially a digital filter with self-adjusting characteristics. An adaptive filter has the property that its frequency response is adjustable or modifiable automatically to improve its performance in accordance with some criterion, allowing the filter to adapt with change in the input signal characteristics.

An adaptive filter consists of two distinct parts: -

- A digital filter with adjustable coefficients. In most adaptive systems, the digital filter is realized using transversal or finite impulse response (FIR) structure, other forms are sometimes used, and in this paper we applied the weighted median as a filter.
- An adaptive algorithm that is used to adjust or modify the coefficients of the filter. Adaptive algorithms are used to adjust the coefficients of the filter being used such that the error term (between the desired output and the filter's output) is minimized according to some criterion. In this paper we applied the LMS algorithm as an adaptive algorithm. (See section v)

III- Threshold decomposition for real valued signals

An important tool for the analysis and design of weighted median filter is the threshold decomposition property [9]. The threshold decomposition was originally formulated to admit signals having only a finite number of positive valued quantization levels. Threshold decomposition was later extended to admit continuous level real-valued signals [7]. Given real-valued samples $X_1, X_2, \ldots, X_N$ forming the vector $X=[X_1, X_2, \ldots, X_N]^T$, where $X_i \in R$. Threshold decomposition maps this real valued vector to an infinite set of binary vectors such as:

$x^q \in \{-1,1\}^N$, $q \in (-\infty,\infty)$ [for every value of q there exists a binary vector], where

$$x^q = [\,\mathrm{sgn}(X_1 - q), \mathrm{sgn}(X_2 - q), \ldots, \mathrm{sgn}(X_N - q)\,]^T = [x_1^q, x_2^q, \ldots, x_N^q]^T \tag{10}$$

where $\mathrm{sgn}(.)$ denotes the signum function defined in (8). Threshold decomposition has several important properties. First, it is reversible. Given a set of threshold signals, each of the samples X_i in X can be reconstructed from its binary representation as

$$X_i = \frac{1}{2}\int_{-\infty}^{\infty} x_i^q \, dq \quad , i = 1, 2, \ldots, N \tag{11}$$

Thus, a real-valued signal has a unique threshold signal representation and vice versa. Second, threshold decomposition is of a particular importance in weighted median filtering, since they are commutable operations, that is, applying a weighted median operator to real-valued signals is equivalent to decomposing the real-valued signal using threshold decomposition into several binary threshold signals, applying the median operator to each binary signal separately, and then adding the binary outputs to obtain the real-valued output [3]. This property is important because the effects of the median on binary signals are much easier to analyze than those on multi-level signals. In fact, the weighted median operation on binary samples reduced to a simple Boolean operation [9]. The median of three binary samples x_1, x_2, x_3 , for example, is equivalent to $x_1 x_2 + x_2 x_3 + x_1 x_3$, where the + (OR) and $x_i x_j$ (AND) Boolean operators in $\{-1,1\}$ domain are defined as: $x_i + x_j = \max(x_i, x_j)$, $x_i x_j = \min(x_i, x_j)$. This opens new possibilities for the analysis because in binary domain, regular Boolean algebra can be applied. Now, since q can take any real-value, the infinite set of binary vectors $\{x^q\}$ contain repeated vectors in representing the real-valued vector X . Thus, according to [3,7], there are at most $(N+1)$ different binary vectors $\{x^q\}$ for each observation vector X , given by :

$$\{x^q\} = \begin{cases} [1, 1, \ldots, 1] & \text{if } -\infty \prec q \le X_{(1)} \\ \left[x_1^{X^+_{(i)}}, x_2^{X^+_{(i)}}, \ldots, x_N^{X^+_{(i)}} \right] & \text{if } X_{(i)} \prec q \le X_{(i+1)}, \; 1 \le i \le N-1 \\ [-1, -1, \ldots, -1] & \text{if } X_{(N)} \prec q \prec \infty \end{cases} \tag{12}$$

Where $X_{(i)}$ is the ith smallest signed sample and $X^+_{(i)}$ denotes a value of the real line approaching $X_{(i)}$ from the right.

IV- The weighted median filter by using threshold decomposition

By using the threshold decomposition property, the weighted median filter[7] in equation(9) can be expressed as:

$$\begin{aligned} \hat{\beta} &= Median\left(|w_i| \lozenge \mathrm{sgn}(w_i) X_i \Big|_{i=1}^{N} \right) \\ &= Median\left(|w_i| \lozenge 1/2 \int_{-\infty}^{\infty} \mathrm{sgn}\left(\mathrm{sgn}(w_i) X_i - q\right) dq \Big|_{i=1}^{N} \right) \end{aligned} \tag{13}$$

Now, let the signed sample vector S is

$$S = [\,\mathrm{sgn}(w_1) X_1, \mathrm{sgn}(w_2) X_2, \ldots, \mathrm{sgn}(w_N) X_N\,]^T \tag{14}$$

(the signed samples $\mathrm{sgn}(w_i)x_i$ is denoted as S_i). The sorted signed samples are then denoted as: $S_{(i)}$ where $S_{(1)} \le S_{(2)} \le .. \le S_{(N)}$. Let W_a be the vector whose elements are the magnitude weights, $W_a = [|w_1|, |w_2|, \ldots, |w_N|]^T$. Then according to [7], the WM filter operation can be expressed as: -

$$\hat{\beta} = 1/2 \int_{-\infty}^{\infty} \operatorname{sgn}(W_a^T s^q)\, dq \tag{15}$$

By using the definition of threshold decomposition in equation (12), equation (15) takes the following form And the derivation of the following relation in (16) is shown in **APPENDIX (1)**

$$\hat{\beta} = \frac{S_{(1)} + S_{(N)}}{2} + \frac{1}{2}\sum_{i=1}^{N-1}(S_{(i+1)} - S_{(i)}) \cdot \operatorname{sgn}(W_a^T s^{S^+_{(i)}}) \tag{16}$$

Where S_i is the ith smallest sample of the set of sorted signed samples defined above, and the value S_i^+ denotes a value on the real line approaching $S_{(i)}$ from the right. From this equation, it is clear that The computation of weighted median filter with the threshold decomposition is efficient, requiring only N-1 threshold logic (sign) operators, allowing the input signals to be arbitrary real-valued signals and allowing positive and negative weights [7], also, we notice that the output is computed by the sum of the midrange of the signed samples $V = \frac{(S_{(1)} + S_{(N)})}{2}$ and by a linear combination of the (i, i+1)th spacing $V_i = S_{(i)} - S_{(i+1)}$ for $i = 1,2,\ldots,N$. In section (vi) we show the effect of applying this filter on a noisy image (figures 3,4 and 10).

V- The adaptive weighted median (AWM) algorithm

In most previous adaptive methods, we would estimate the gradient of $\xi(w) = E(J^2(w))$ (the gradient of the mean square error) where $J(w)$ is the error term between the desired output and the filter's output. In developing the (AWM) algorithm, instead, we take $J^2(w)$ itself as an estimate of $\xi(w)$, and then, during each iteration in the adaptive process we compute an estimate of the gradient. Let $D(n)$ be the desired signal and $Y(n)$ be the filter's output, then the error term is $J(w) = D(n) - Y(n)$. Now, the goal is to determine the weight values in $W = [w_1, w_2, \ldots, w_N]$ which will minimize the estimation error under the mean square error (MSE) and the steepest descent algorithm:

$$w_j(n+1) = w_j(n) + 2u(-\hat{\nabla}) \tag{17}$$

Where $\hat{\nabla} = \frac{\partial}{\partial w}(J^2(w))$, $J(w) = D(n) - Y(n)$. Then, by computing the gradient and substituting in equation (17) we can derive a formula for the adaptive weighted median, now by using threshold decomposition the error term can be expressed as

$$J(w) = \frac{1}{2}\int_{-\infty}^{\infty}(\operatorname{sgn}(D - q) - \operatorname{sgn}(W_a^T s^q))\, dq \tag{18}$$

Since $\hat{\nabla} = \frac{\partial}{\partial W}(J^2(W)) = 2J\frac{\partial J(W)}{\partial W}$, thus,

$$\hat{\nabla} = 2 \cdot \frac{1}{2}\int_{-\infty}^{\infty}\left[\operatorname{sgn}(D - q) - \operatorname{sgn}(W_a^T s^q)\right] dq \cdot \frac{-1}{2}\int_{-\infty}^{\infty}\frac{\partial}{\partial W}(\operatorname{sgn}(W_a^T s^q))\, dq$$

$$\hat{\nabla} = \frac{-1}{2}\int_{-\infty}^{\infty} e^q(n)\, dq \cdot \int_{-\infty}^{\infty}\frac{\partial}{\partial W}(\operatorname{sgn}(W_a^T s^q))\, dq \tag{19}$$

Where $e^q(n) = \operatorname{sgn}(D - q) - \operatorname{sgn}(W_a^T s^q)$ can be thought of as the threshold decomposition of the error function $e(n) = D(n) - Y(n)$. At this point our task is to evaluate the value of $\hat{\nabla}$, this expression is derived in the **APPENDIX (2),** and is given by

$$\hat{\nabla} = \frac{-1}{2}\sum_{i=Min(m,j)}^{Max(m,j)-1} (S_{(i+1)} - S_{(i)})\, e^{S^{+}_{(i)}} \cdot \left\{ \begin{array}{l} (S_{(1)}+S_{(N)}).\mathrm{sgn}(w_j).\sec h^{2}(\sum_{l=1}^{N} |w_l|)+ \\ \sum_{k=1}^{N-1} [(S_{(k+1)}-S_{(K)}).\ \mathrm{sgn}\ (w_j).s_j^{S^{+}_{(k)}} \\ \sec h^{2}(W_a^{T} s^{S^{+}_{(k)}}) \end{array} \right\} \quad (20)$$

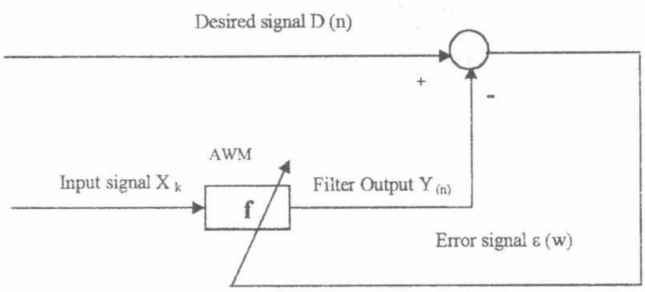
By substitute (20) into (17)(the steepest descent algorithm), we get the following recursive expression that used in adjusting the filter coefficients under the MSE.

$$w_j(n+1) = w_j(n) + u \left\{ \begin{array}{l} \left(\sum_{i=Min(m,j)}^{Max(m,j)-1} (S_{(i+1)} - S_{(i)})\ e^{S^{+}_{(i)}} \right). \\ \left\{ \begin{array}{l} (S_{(1)} + S_{(N)}).\ \mathrm{sgn}\ (w_j).\sec h^{2}(\sum_{l=1}^{N} |w_l|)+ \\ \sum_{k=1}^{N-1} (S_{(k+1)} - S_{(K)}).\ \mathrm{sgn}\ (w_j).s_j^{S^{+}_{(k)}}\ .\sec h^{2}(W_a^{T}\ s^{S^{+}_{(k)}}) \end{array} \right\} \end{array} \right\} \quad (21)$$

Since the MSE criterion was used in the derivation, the recursive in equation (20) is referred to as the least mean square (LMS) weighted median adaptive algorithm (AWM)

Figure (2)

Adaptive filter diagram



VI. Results

The following results show the effects of using median, weighted median, adaptive weighted median filters on a noisy image presented in figure (4). Table 1 compares different filtering methods using Mean square Error (MSE) measure. The results show that the proposed method using (AWM) produces the best results as compared to other filtering methods.

Table 1 Summary of filtering results (salt & pepper) noisy image

Method	MSE
Median	0.0015 fig (7)
Weighted median	0.0013 fig (5), 0.0017 fig (6)
Adaptive Weighted median(AWM)	0.0010 fig (8), 0.0011 fig (9)

Figure (3)

Figure (4)

Figure (5)

Figure (6)

Figure (3) original image, figure (4) a noisy image (salt & pepper), figure (5) and figure (6) weighted median Filter

Table 2 compares different filtering methods on (salt & pepper and Gaussian) noisy image presented in fig (10) using MSE measure. The results show that the proposed method using (AWM) produces the best results as compared to other filtering methods.

Table 2 Summary of filtering results (salt & pepper and Gaussian) noisy image

Method	MSE
Median	0.0027 fig (11)
Weighted median	0.0026 fig (12)
Adaptive Weighted median (AWM)	0.0024 fig (13)
Average filter	0.0038 fig (14)

Figure (7)

Figure (8)

Figure (9)

Figure (10)

Figure (7) median filter, figure (8), figure (9) adaptive weighted median filter(AWM) and figure (10), noisy image by (salt & pepper, Gaussian noise).

Figure (11)

Figure (12)

Figure (13)

Figure (14)

Figure (11) median filter, figure (12) weighted median filter, figure (13) adaptive weighted median filter(AWM)and figure (14) average filter on (salt & pepper and Gaussian) noisy image.

VI. Conclusions

This paper presented an adaptive algorithm that combines both LMS algorithm and the weighted median. The proposed algorithm is called AWM. This algorithm is used for impulse noise filtering in images. The results presented shows that the proposed algorithm gives the best results considering the MSE error measure and visually. Analysis of the proposed algorithm is also presented.

APPENDIX (1)

In this appendix, we derive the expression for the output of the weighted median given in equation (16), by using the threshold decomposition. The weighted median filter [7] defined in equation (9) can be expressed as

$$\hat{\beta} = Median\left(|w_i| \Diamond \operatorname{sgn}(w_i) X_i \Big|_{i=1}^{N} \right) = Median\left(|w_i| \Diamond 1/2 \int_{-\infty}^{\infty} \operatorname{sgn}(\operatorname{sgn}(w_i) X_i - q)\, dq \Big|_{i=1}^{N} \right)$$

The signed sample vector S is $S = [\operatorname{sgn}(W_1) X_1, \operatorname{sgn}(W_2) X_2, \ldots, \operatorname{sgn}(W_N) X_N]^T$, and let the sorted signed samples be denoted as $S_{(i)}$ where $S_{(1)} \le S_{(2)} \le \ldots \le S_{(N)}$. Now, we resort to the weak superposition property of the nonlinear median, which states that applying a weighted median operator to real-valued signals is equivalent to decomposing the real-valued signal using threshold decomposition into several binary threshold signals, applying the median operator to each binary signal separately, and then adding the binary outputs to obtain the real-valued output [3]. This property lead to interchanging the integral and median operators, and thus,

$$\hat{\beta} = 1/2 \int_{-\infty}^{\infty} Median\left(|W_i| \Diamond \operatorname{sgn}[\operatorname{sgn}(W_i) X_i - q] \Big|_{i=1}^{N} \right) dq \tag{22}$$

Now, let the vector s^q represents the threshold decomposed signed samples vector defined below: -

$s^q = [\operatorname{sgn}(\operatorname{sgn}(W_1) X_1 - q), \ldots, \operatorname{sgn}(\operatorname{sgn}(W_N) X_N - q)]^T = [s_1^q, s_2^q, \ldots, s_N^q]^T$, where $\operatorname{sgn}(.)$ denotes the signum function defined in (8). Let W_a be the vector whose elements are the magnitude weights

$W_a = [|W_1|, |W_2|, ..., |W_N|]^T$. Then, according to [7], the WM filter operation can be expressed as:

$$\hat{\beta} = 1/2 \int_{-\infty}^{\infty} Sgn(W_a^T s^q) dq \tag{23}$$

Let $S_{(i)}$ be the ith smallest signed sample. Then, by using the threshold decomposition in (12), we have

$$\begin{aligned} \hat{\beta} &= \frac{1}{2} \int_{-\infty}^{S_{(1)}} \operatorname{sgn}(W_a^T s^{S_{(1)}^-}) dq \\ &+ \frac{1}{2} \sum_{i=1}^{N-1} \int_{S_{(i)}}^{S_{(i+1)}} \operatorname{sgn}(W_a^T s^{S_{(i)}^+}) dq \\ &+ \frac{1}{2} \int_{S_{(N)}}^{\infty} \operatorname{sgn}(W_a^T s^{\infty}) dq \end{aligned} \tag{24}$$

Where $S_{(i)}^+$ denotes a value on the real line approaching $S_{(i)}$ from the right.

Since the first and the last integrals are improper then, by the definition of improper integral we can reduce the above equation to:

$$\hat{\beta} = \frac{1}{2} \lim_{Q \to \infty} (S_{(1)} + Q) + \frac{1}{2} \sum_{i=1}^{N-1} (S_{(i+1)} - S_{(i)}) . \operatorname{sgn}(W_a^T s^{S_{(i)}^+}) - \frac{1}{2} \lim_{Q \to \infty} (Q - S_{(N)})$$

$$\hat{\beta} = \frac{1}{2} \lim_{Q \to \infty} (S_{(1)} + Q - Q + S_{(N)}) + \frac{1}{2} \sum_{i=1}^{N-1} (S_{(i+1)} - S_{(i)}) . \operatorname{sgn}(W_a^T s^{S_{(i)}^+})$$

Then, the output of the weighted median can be expressed as the following: -

$$\hat{\beta} = \frac{S_{(1)} + S_{(N)}}{2} + \frac{1}{2} \sum_{i=1}^{N-1} (S_{(i+1)} - S_{(i)}) . \operatorname{sgn}(W_a^T s^{S_{(i)}^+}) \tag{25}$$

APPENDIX (2)

In this appendix, we derive the expression for computing an estimate of the gradient to use it in adjusting the coefficients of the weighted median by the steepest descent algorithm to get the adaptive weighted median (AWM) given in equation (21). Now, our task is to evaluate the following integral: -

$$\hat{\nabla} = \frac{-1}{2} \int_{-\infty}^{\infty} e^q(n)\, dq \;\; . \int_{-\infty}^{\infty} \frac{\partial}{\partial W} (\operatorname{sgn}(W_a^T s^q))\, dq \tag{26}$$

Since the signum function is discontinuous at the origin so it is approximated by a differentiable function $\operatorname{sgn}(X) \cong \tanh(X) = \frac{e^X - e^{-X}}{e^X + e^{-X}}$ Whose derivative is $\frac{\partial}{\partial X}(\tanh(X)) = \sec h^2(X)$, then

$$\begin{aligned} \frac{\partial}{\partial W} \operatorname{sgn}(W_a^T s^q) &\cong \sec h^2 (W_a^T s^q) \frac{\partial}{\partial W} (W_a^T s^q) \\ &= \sec h^2 (W_a^T s^q) \begin{bmatrix} \operatorname{sgn}(w_1) s_1^q \\ . \\ . \\ \operatorname{sgn}(w_N) s_N^q \end{bmatrix} \end{aligned} \tag{27}$$

By using this approximation (27), equation (26) can be written as follows:

$$\frac{\partial}{\partial w_j} J(w) = \frac{-1}{2} \int_{-\infty}^{\infty} e^{q}(n)\, dq \bullet \int_{-\infty}^{\infty} \sec h^2 (W_a^T s^q) \operatrm{sgn}(w_j) s_j^q \, dq \tag{28}$$

Where s_j^q is the j-th component of s^q i.e. ($s_j^q = \operatorname{sgn}(S_j - q)$, $S_j \in R$).

Now, we have to determine the value of each integral in equation (28). First, we can evaluate the first integral as follows: - Let

$$I_n = \int_{-\infty}^{\infty} e^{q}(n)\, dq \tag{29}$$

Since the threshold decomposition of the error term $e^{q}(n)$ takes non-zero values only if (q) is between the desired output D (n) and the actual filter's output Y (n) [3], assuming that D (n) is one of the signed samples (say $S_{(m)}$) and the actual output Y (n) (say $S_{(j)}$), so it can be shown that :

$$e^{q}(n) \neq 0 \ \ if \quad Min\,(S_{(m)}, S_{(j)}) \leq q \leq Max\,(S_{(m)}, S_{(j)})$$

i.e. $e^{q}(n) = 0 \ for\ q \in \{(-\infty, Min\,(S_{(m)}, S_{(j)})) \cup (Max\,(S_{(m)}, S_{(j)}), \infty)\}$

Then $$I_n = \int_{-\infty}^{\infty} e^{q}(n)\, dq = \int_{Min\,(S_{(m)}, S_{(j)})}^{Max\,(S_{(m)}, S_{(j)})} e^{q}(n)\, dq \tag{30}$$

From the properties of integrals, equation (30) can be written in the following form :

$$I_n = \sum_{i=Min(m,j)}^{max(m,j)-1} \int_{S_{(i)}}^{S_{(i+1)}} e^{S^+_{(i)}} dq$$

$$I_n = \sum_{i=Min\,(m,j)}^{Max\,(m,j)-1} (S_{(i+1)} - S_{(i)})\, e^{S^+_{(i)}} \tag{31}$$

similarly, the second integral $J_n = \int_{-\infty}^{\infty} \sec h^2 (W_a^T s^q) \operatorname{sgn}(w_j) s_j^q \, dq$ (32)

can be evaluated in the same manner as follows: -

$$J_n = \int_{-\infty}^{\infty} \operatorname{sech}^2 (W_a^T s^q) \operatorname{sgn}(w_j) s_j^q \, dq$$

$$\begin{aligned} J_n = & \int_{-\infty}^{S_{(1)}} \sec h^2 (W_a^T s^{S^-_{(1)}}) \operatorname{sgn}(w_j) \; s_j^{S^-_{(1)}} dq \\ & + \sum_{k=1}^{N-1} \int_{S_{(k)}}^{S_{(k+1)}} \sec h^2 (W_a^T s^{S^+_{(k)}}) \operatorname{sgn}(w_j) \; s_j^{S^+_{(k)}} dq \\ & + \int_{S_{(N)}}^{\infty} \sec h^2 (W_a^T s^{S^+_{(N)}}) \operatorname{sgn}(w_j) \; s_j^{S^+_{(N)}} dq \end{aligned} \tag{33}$$

Since we have $s^q = [s_1^q, s_2^q, \ldots, s_N^q] = [\operatorname{sgn}(S_1 - q), \operatorname{sgn}(S_2 - q), \ldots, \operatorname{sgn}(S_N - q)]$

Then, it can be shown that $s^{S_{(1)}^{-}} = [1, 1, ..., 1]^T$, Where $S_{(1)}$ is the smallest signed sample.

Also, it is clear that $s^{S_{(N)}^{+}} = [-1, -1, ..., -1]^T$, where $S_{(N)}$ is the largest signed sample.

Thus, we have :

$$(W_a^T s^{S_{(1)}^{-}}) = \begin{bmatrix} |w_1| \\ |w_2| \\ \cdot \\ \cdot \\ |w_N| \end{bmatrix} \begin{bmatrix} s_1^{S_{(1)}^{-}} \\ s_2^{S_{(1)}^{-}} \\ \cdot \\ \cdot \\ s_N^{S_{(1)}^{-}} \end{bmatrix} = \begin{bmatrix} |w_1| \\ |w_2| \\ \cdot \\ \cdot \\ |w_N| \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \cdot \\ \cdot \\ 1 \end{bmatrix} = \sum_{l=1}^{N} |w_l| \tag{34}$$

Thus, $\operatorname{sech}^2 (W_a^T s^{S_{(1)}^{-}}) = \operatorname{sech}^2 (\sum_{l=1}^{N} |w_l|)$ and the j-th $(s_j^{S_{(1)}^{-}})$ component of $s^{S_{(1)}^{-}}$ is 1.

In the same manner: - $(W_a^T s^{S_{(N)}^{+}}) = -(\sum_{l=1}^{N} |w_l|)$ and the j-th $(s_j^{S_{(N)}^{+}})$ component of $s^{S_{(N)}^{+}}$ is -1.

And thus, $\operatorname{sech}^2 (W_a^T s^{S_{(N)}^{-}}) = \operatorname{sech}^2 (-(\sum_{l=1}^{N} |w_l|))$.

Then, from the last notes the first part of the integral (say J_{nf}) in (33) can be evaluated as improper integral as follows :

$$\begin{aligned} J_{nf} &= \int_{-\infty}^{S_{(1)}} \operatorname{sech}^2 (W_a^T s^{S_{(1)}^{-}}) \operatorname{sgn}(w_j)\, s_j^{S_{(1)}^{-}}\, dq \\ &= \operatorname{sech}^2 (\sum_{l=1}^{N} |w_l|).\operatorname{sgn}(w_j).\lim_{Q\to\infty} (S_{(1)} + Q) \end{aligned} \tag{35}$$

Similarly, the last part of (say J_{nl}) in (33)is given by :

$$\begin{aligned} J_{nl} &= \int_{S_{(N)}}^{\infty} \operatorname{sech}^2 (W_a^T s^{S_{(N)}^{+}}) \operatorname{sgn}(w_j)\, s_j^{S_{(N)}^{+}}\, dq \\ &= -\operatorname{sech}^2 (-\sum_{l=1}^{N} |w_l|).\operatorname{sgn}(w_j).\lim_{Q\to\infty} (Q - S_{(N)}) \end{aligned} \tag{36}$$

Since $\operatorname{sech}(x) = \operatorname{sech}(-x)$, the sum of the first term J_{nf} and the last term J_{nl} of the integral J_n can be reduced to :

$$\begin{aligned} J_{nf} + J_{nl} &= \operatorname{sech}^2 (\sum_{l=1}^{N} |w_l|).\operatorname{sgn}(w_j).\lim_{Q\to\infty}(S_{(1)} + Q) - \operatorname{sech}^2 (-\sum_{l=1}^{N} |w_l|).\operatorname{sgn}(w_j).\lim_{Q\to\infty}(Q - S_{(N)}) \\ &= \operatorname{sech}^2 (\sum_{l=1}^{N} |w_l|).\operatorname{sgn}(w_j).\lim_{Q\to\infty}(S_{(1)} + Q - Q + S_{(N)}) \\ &= (S_{(1)} + S_{(N)}).\operatorname{sgn}(w_j).\operatorname{sech}^2 (\sum_{l=1}^{N} |w_l|) \end{aligned} \tag{37}$$

Then, the integral $J_n = \int_{-\infty}^{\infty} \operatorname{sech}^2 (W_a^T s^q) \operatorname{sgn}(w_j)\, s_j^q\, dq$ is equal to

$$J_n = J_{nf} + \sum_{k=1}^{N-1}(S_{(k+1)} - S_{(K)}).\ \mathrm{sgn}(w_j).s_j^{S^+_{(k)}}.\mathrm{sech}^2(W_a^T s^{S^+_{(k)}}) + J_{nl}$$

$$J_n = (S_{(1)} + S_{(N)}).\mathrm{sgn}(w_j).\mathrm{sech}^2\left(\sum_{l=1}^{N}|w_l|\right) + \sum_{k=1}^{N-1}(S_{(k+1)} - S_{(K)}).\ \mathrm{sgn}(w_j).s_j^{S^+_{(k)}}.\mathrm{sech}^2(W_a^T s^{S^+_{(k)}}) \quad (38)$$

Thus,

$$w_j(n+1) = w_j(n) + u\left\{\begin{array}{l} \left(\sum_{i=Min(m,j)}^{Max(m,j)-1}(S_{(i+1)} - S_{(i)})\ e^{S^+_{(i)}}\right). \\ \left\{\begin{array}{l} (S_{(1)} + S_{(N)}).Sgn(w_j).\sec h^2\left(\sum_{l=1}^{N}|w_l|\right) + \\ \sum_{k=1}^{N-1}(S_{(k+1)} - S_{(K)}).\ Sgn(w_j).s_j^{S^+(k)}.\sec h^2(W_a^T s^{S^+_{(k)}}) \end{array}\right\} \end{array}\right\} \quad (39)$$

Since the MSE criterion was used in the derivation, the recursive in equation (39) is referred to as the least mean square (LMS) weighted median adaptive algorithm (AWM).

References

[1] L. Yin,R. Yang, M. Gabbouj, and Y. Neuvo "Weighted median filters: A Tutorial," IEEE Trans. On circuits and systems, Vol. 43, Mar. 1996.

[2] B.Widrow and S.D.Stearns, adaptive signal processing Engel wood Cliffs, NJ: Prentice-Hall, 1985

[3] G.R.Arce and J. L. Paredes,"Recursive Weighted median filters admitting negative weights and their optimization" IEEE Trans. Signal Processing, Vol .48, Mar. 2000

[4] D.R.K.Brownrigg, " The weighted median filter" Commun. Assoc. Comput. Mach., Vol.27, Aug. 1984.

[5] 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, Feb. 1991.

[6] L. Yin and Y.Neuvo , " Fast adaptation and performance characteristic of FIR- Wos hybrid filters." IEEE Trans. Signal Processing , Vol .42, July 1994.

[7] G.R.Arce, " A general weighted median filter structure admitting negative real valued weights." IEEE Trans. Signal Processing, Sp 46, Dec.1998.

[8] S. J. Ko and Y.H.Lee, " Center weighted median filters and their applications to image enhancement." IEEE Trans. On Circuits and Systems. 38, Sept.1991.

[9] ALBOVIK, Hand book of image and video Processing, Acadimic press2000.

[10] O. Yli-Harja, J. Astola, and Y. Neuvo, "Analysis of the properties of filter and weighted median filters using threshold logic and stack filter representation,"IEEE Trans. Acoust., Speech, Signal Processing, Vol. 39 ,Feb. 1991.