

Generalized Correlativity of Median Filtering Operator on Signals*

Wanzhou Ye, Zhao Liao

Department of Mathematics, Shanghai University, Shanghai, China
Email: wzhy@shu.edu.cn

Received April 21, 2012; revised May 28, 2012; accepted June 17, 2012

ABSTRACT

The generalized correlativity of input signal and output signal of a stack filtering operator is defined and used for numerously measuring these filtering operators' behavior in removing noise in signals. We show that under the criterion of the generalized correlativity, of stack filtering operators the median filtering operator is optimal, which implies that this filtering operator possesses better filtering behavior than the others.

Keywords: Median Filter; Stack Filter; Generalized Correlativity

1. Introduction

The median filter was introduced in 1974 by Tukey [1], who used the moving median as a smoothing technique in time series analysis. Since then, there have been many studies on deterministic and statistical properties of median filtering operator, such as properties of roots of median filters ([2-6]), structures of recurrent sequences of median filters ([7-9]), convergence properties of a signal ([2,10-12]), and the output distribution of median-type filters under known distributions [13]. L-filters, a generalization of median filters, were introduced [14], their output is given by a linear combination of the order statistics of the input sequence. Assuming a constant signal in white noise, the coefficients in the linear combination are chosen to minimize the output MSE for several noise distributions, *i.e.*, the optimal design of L-filters, and a new methodology for the design of L-filters is presented in which the L-filter coefficients are obtained by approximating the covariance matrix of the ordered samples through the use of Taylor expansion [15]. Most practical applications deal with unknown signals, which means that the noise distributions of signals are difficult to characterize. There are cases such as in detection, and an example was given [11], which is an original two-dimensional image obtained by Dr. J. Johnson of Redstone Arsenal with an experimental laser imaging system at the far infrared wavelength of 1.2 mm. It is very difficult to characterize the noise source in the image. In these situations, the statistical methods above do not work. Therefore, a significant filtering problem is:

Given a signal with some unknown noise and a filter, how do we numerically measure its filtering behavior for the signal? This problem is general and complicated. So, in this paper, we will study this problem only for stack filters.

The organization of this paper is as follows. In Section 2 we review stack filters and their threshold decomposition property, and define the generalized correlativity between of a discrete binary input signal and its output signal of stack filters. In Section 3 we prove the optimality of median filters based on the generalized correlativity in terms of stack filters. In Section 4 we give some properties of the median filtering operator to show when and how to use the filtering operator better.

2. Stack Filtering Operators and Generalized Correlativity on Signals

In this section, first we briefly review the definition of stack filters. Stack filters are a class of sliding window, nonlinear digital filters that possess the stacking property and the threshold decomposition architecture [11]. The threshold decomposition of an M -valued signal $x(n)$ is the set of $M - 1$ binary signals, called threshold signals, $x^1(n), x^2(n), \dots, x^{M-1}(n)$, which are defined by

$$x^m(n) = T^m(x(n)) = \begin{cases} 1, & \text{if } x(n) \geq m; \\ 0, & \text{otherwise} \end{cases}$$

Note that the sum of the threshold signals $x^m(n)$ is $x(n)$, *i.e.*,

$$\sum_{m=1}^{M-1} x^m(n) = x(n).$$

*Supported by Shanghai Leading Academic Discipline Project (J50101) NSF of Shanghai (10ZR1412100), and NNSF of China (61071186).

Also, the $x^m(n)$'s are ordered, *i.e.*,

$$x^1(n) \geq x^2(n) \geq \dots \geq x^{M-1}(n).$$

This property is called the stacking property of sequences. A very important property of median filters operating within a sliding window was observed by Fitch *et al.* [16]: Applying a median filter to an M -valued signal is equivalent to decomposing the signal to $M - 1$ binary threshold signals, filtering each binary signal separately with the median filter, and then adding the binary output signals together. This is called the threshold decomposition architecture of median filters.

In the above architecture, the median operation on binary input signals reduces to a Boolean function that possesses the stacking property.

Definition 1. A $2k + 1$ -input Boolean function

$g : \{0, 1\}^{2k+1} \rightarrow \{0, 1\}$ is said to possess the stacking property if whenever two input vectors x and y stack, *i.e.*, $x_i \geq y_i$ for each $i \in \{1, 2, \dots, 2k+1\}$, then also their outputs stack $g(x) \geq g(y)$.

It has been shown that a Boolean function has the stacking property if and only if it can be expressed as a Boolean expression that contains no complements of input variables [11]. Such functions are called positive Boolean functions (PBF).

Definition 2. The stack filter S_g based on the PBF $g : \{0, 1\}^{2k+1} \rightarrow \{0, 1\}$ is defined as follows:

$$S_g(x) = \sum_{m=1}^{M-1} g(x^m),$$

where

$$x = [x_1, x_2, \dots, x_{2k+1}]$$

and

$$x^m = T^m(x) = [T^m(x_1), T^m(x_2), \dots, T^m(x_{2k+1})].$$

Stack filters include median filters. Since stack filters possess the threshold decomposition property and the stacking property, a complete characterization of the effect of the stack filter on binary signals is sufficient to characterize the behavior of the M -valued filter. Since any $\{a(n)\}$ with $a(n) \in \{0, 1\}$ can be transformed into $\{b(n)\}$ with $b(n) \in \{-1, 1\}$ by the operation: $b(n) = 2a(n) - 1, n \in Z$. In what follows, attention will therefore be focused mainly on stack filters with binary input taking values 1 and -1 , and k is a fixed positive integer, Z is the set of integers.

Thus, as Definition 1, we can define the PBF $g : \{-1, 1\}^{2k+1} \rightarrow \{-1, 1\}$, and for each $\{x(n)\} \in \{-1, 1\}^Z$, let, for each $n \in Z$

$$S_g(\{x(n)\})(n) = g(x(n-k), x(n-k+1), \dots, x(n+k))$$

We call S_g a stack filtering operator with width $2k + 1$. Let Ω_k denote the set of stack filtering operators

with width $2k + 1$.

From the definition, it is clear that for each $n \in Z$, $S_g(\{x(n)\})(n)$ depends on $x(n-k), x(n-k+1), \dots, x(n+k)$. Therefore, we will investigate the relation between $\{x(n-k)\}, \{x(n-k+1)\}, \dots, \{x(n+k)\}$ and $S_g(\{x(n)\})$. Usually, if two sequences $\{a(n)\}$ and $\{b(n)\}$ belong to l^2 , that is,

$$\sum_{n \in Z} (a(n))^2 < \infty \text{ and } \sum_{n \in Z} (b(n))^2 < \infty,$$

then we use

$$\sum_{n \in Z} a(n)b(n)$$

to study the correlativity of $\{a(n)\}$ and $\{b(n)\}$, but all $\{x(n-k)\}, \{x(n-k+1)\}, \dots, \{x(n+k)\}$ and $S_g(\{x(n)\})$ do not belong to l^2 . Therefore, we first introduce such a window sequence with width $2k + 1$: $\lambda(0), \lambda(1), \dots, \lambda(2k + 1)$ satisfying

- $1 = \lambda(0) > \lambda(1) > \dots > \lambda(2k) \geq \frac{k}{k+1}$.

Further, we define a real sequence $\{\lambda(n)\}_{n \in Z}$ satisfying

- $\lambda(i) = \frac{k}{k+1} \lambda(i-2k)$ for each $i > 2k$.
- $\lambda(n) = \lambda(-n), n \in Z$;

For each $\{x(n)\} \in \{-1, 1\}^Z$, we have

$$\begin{aligned} \sum_{n \in Z} (\lambda(n)x(n))^2 &= \sum_{n \in Z} \lambda^2(n) \\ &= 1 + 2 \sum_{n=1}^{\infty} \left[(\lambda^2(1) + \lambda^2(2) + \dots + \lambda^2(2k)) \cdot \left(\frac{k}{k+1}\right)^{2(n-1)} \right]. \end{aligned}$$

Since $0 < \frac{k}{k+1} < 1$, $\sum_{n \in Z} (\lambda(n)x(n))^2 < \infty$. Therefore,

$\{\lambda(n)x(n)\} \in l^2$. This is why we introduce such a real sequence.

Again we give the following definition.

Definition 3. Suppose S_g is a window filtering operator with width $2k + 1$. For each $\{x(n)\} \in \{-1, 1\}^Z$, let

$$\begin{aligned} C(\{x(n)\}, S_g(\{x(n)\})) \\ = \sum_{i=-k}^k \sum_{n \in Z} \lambda(n+i)x(n+i)\lambda(n)S_g(\{x(n)\})(n). \end{aligned}$$

We call it the generalized correlativity of $\{x(n)\}$ and $S_g(\{x(n)\})$.

In particular, let $m : \{-1, 1\}^{2k+1} \rightarrow \{-1, 1\}$ satisfy that for each $(x(1), x(2), \dots, x(2k + 1)) \in \{-1, 1\}^{2k+1}$,

$$m(x(1), x(2), \dots, x(2k + 1)) = \begin{cases} 1, & \text{if } \sum_{j=1}^{2k+1} x(j) > 0; \\ -1, & \text{otherwise} \end{cases}$$

Then S_g is the median filtering operator and $S_g \in \Omega_k$.

It is clear that the generalized correlativity is a measure of the relation between $\{x(n)\}$ and $S_g(\{x(n)\})$. From a view of signal processing, it is shown that if we apply a stack filtering operator S_g to a signal $\{x(n)\}$ with noise, then the generalized deviation $\{x(n)\}$ and $S_g(\{x(n)\})$ may be a measure of the S_g 's capability of removing noise in the signal.

3. Optimality of Median Filtering Operator

In the optimal design of filters, it is usual to select the optimal filter using *MSE* criterion or *MAE* criterion, for example, the optimal design of L-filters is based on *MSE* criterion. Both *MSE* criterion and *MAE* criterion mean that output signal is closest to input signal, which implies their high correlativity. Therefore, we have the following definition.

Definition 4. Suppose $S_f, S_g \in \Omega_k$. If, for each $\{x(n)\} \in \{-1, 1\}^Z$,

$$C(\{x(n)\}, S_f \{x(n)\}) \geq C(\{x(n)\}, S_g \{x(n)\}),$$

then we say that S_f is better than S_g in terms of filtering behavior, denoted by " $S_f \geq S_g$ ".

Based on the definition, we have the following result.

Theorem 1. For any $S_g \in \Omega_k$, we have $S_m \geq S_g$.

Proof. Suppose that $\{x(n)\} \in \{-1, 1\}^Z$ and $S_g \in \Omega_k$. Since $x(n), S_g(\{x(n)\})(n) \in \{-1, 1\}$, we have

$$\begin{aligned} & C(\{x(n)\}, S_m(\{x(n)\})) - C(\{x(n)\}, S_g(\{x(n)\})) \\ &= \sum_{i=-k}^k \sum_{n \in Z} \lambda(n+i) \lambda(n) x(n+i) S_m(\{x(n)\})(n) \\ &\quad - \sum_{i=-k}^k \sum_{n \in Z} \lambda(n+i) \lambda(n) x(n+i) S_g(\{x(n)\})(n) \\ &= \sum_{n \in Z} \sum_{i=-k}^k \lambda(n+i) \lambda(n) x(n+i) \\ &\quad \cdot [S_m(\{x(n)\})(n) - S_g(\{x(n)\})(n)] \\ &= \sum_{n \in Z} \lambda(n) [S_m(\{x(n)\})(n) - S_g(\{x(n)\})(n)] \\ &\quad \cdot \left[\sum_{i=-k}^k \lambda(n+i) x(n+i) \right]. \end{aligned}$$

Now we prove that for each $n \in Z$,

$$\sum_{i=-k}^k \lambda(n+i) x(n+i) \cdot \sum_{i=-k}^k x(n+i) > 0 \quad (1)$$

In fact, since $x(n+i) \in \{-1, 1\}$ for $i = -k, -k+1, \dots, k$, $\sum_{i=-k}^k |x(n+i)| = 2k+1$. Therefore,

$$\sum_{i=-k}^k x(n+i) \neq 0.$$

Case 1. $\sum_{i=-k}^k x(n+i) > 0$.

In this case, we have

$$\sum_{x(n+i)=1} x(n+i) \geq k+1 \quad \text{and} \quad \sum_{x(n+i)=-1} (-x(n+i)) \leq k.$$

Thus

$$\begin{aligned} & \sum_{i=-k}^k \lambda(n+i) x(n+i) \\ &= \sum_{x(n+i)=1} \lambda(n+i) x(n+i) - \sum_{x(n+i)=-1} \lambda(n+i) (-x(n+i)) \\ &> \min \lambda[n-k, n+k] \sum_{x(n+i)=1} x(n+i) \\ &\quad - \max \lambda[n-k, n+k] \sum_{x(n+i)=-1} (-x(n+i)) \\ &\geq (k+1) \min \lambda[n-k, n+k] - k \max \lambda[n-k, n+k], \end{aligned}$$

where

$$\begin{aligned} \min \lambda[i, j] &= \min \{\lambda(i), \lambda(i+1), \dots, \lambda(j)\}, \\ \max \lambda[i, j] &= \max \{\lambda(i), \lambda(i+1), \dots, \lambda(j)\}. \end{aligned}$$

By the definition of sequence $\{\lambda(n)\}$, for any $n, i \in Z$ with $n-k \leq i \leq n+k$, we have

$$\lambda(n-k) \geq \lambda(i) \geq \lambda(n+k)$$

and

$$\lambda(n+k) = \frac{k}{k+1} \lambda(n-k) \quad \text{for } n \geq k;$$

$$\lambda(n-k) \leq \lambda(i) \leq \lambda(n+k)$$

and

$$\lambda(n-k) = \frac{k}{k+1} \lambda(n+k) \quad \text{for } n \leq -k;$$

$$\frac{k}{k+1} \leq \lambda(i) \leq 1 \quad \text{for } |n| < k.$$

So we have

$$\begin{aligned} & (k+1) \min \lambda[n-k, n+k] \\ & - k \max \lambda[n-k, n+k] \geq 0. \end{aligned} \quad (2)$$

Thus

$$\sum_{i=-k}^k \lambda(n+i) x(n+i) > 0.$$

Therefore, (1) holds.

Case 2. $\sum_{i=-k}^{n+k} x(n+i) < 0$.

In this case, we have

$$\sum_{x(n+i)=-1} (-x(n+i)) \geq k+1 \quad \text{and} \quad \sum_{x(n+i)=1} (x(n+i)) \leq k.$$

Thus

$$\begin{aligned} & \sum_{i=-k}^k \lambda(n+i)x(n+i) \\ &= \sum_{x(n+i)=1} \lambda(n+i)x(n+i) - \sum_{x(n+i)=-1} \lambda(n+i)(-x(n+i)) \\ &< \max \lambda[n-k, n+k] \sum_{x(n+i)=1} x(n+i) \\ &\quad - \min \lambda[n-k, n+k] \sum_{x(n+i)=-1} (-x(n+i)) \\ &\geq k \max \lambda[n-k, n+k] - (k+1) \min \lambda[n-k, n+k]. \end{aligned}$$

By (2), we have

$$k \max \lambda[n-k, n+k] - (k+1) \min \lambda[n-k, n+k] \leq 0$$

Thus

$$\sum_{i=-k}^k \lambda(n+i)x(n+i) < 0.$$

Therefore, (1) holds.

For any $n \in Z$, if $\sum_{i=-k}^k \lambda(i)x(i) > 0$, then

$\sum_{i=-k}^{+k} x(i) > 0$. By the definition of S_m , $S_m(\{x(n)\})(n) = 1$. Again $S_g(\{x(n)\})(n) \in \{-1, 1\}$, so $S_m(\{x(n)\})(n) - S_g(\{x(n)\})(n) \geq 0$. Therefore,

$$\begin{aligned} & v(n) [S_m(\{x(n)\})(n) - S_g(\{x(n)\})(n)] \\ & \cdot \left[\sum_{i=-k}^k \lambda(n+i)x(n+i) \right] \geq 0. \end{aligned}$$

If $\sum_{i=-k}^k v(i)x(i) < 0$, then $\sum_{i=-k}^{+k} x(i) < 0$. By the definition of S_m , $S_m(\{x(n)\})(n) = -1$. Again

$S_g(\{x(n)\})(n) \in \{-1, 1\}$, so $S_m(\{x(n)\})(n) - S_g(\{x(n)\})(n) \leq 0$. Therefore,

$$\begin{aligned} & v(n) [S_m(\{x(n)\})(n) - S_g(\{x(n)\})(n)] \\ & \cdot \left[\sum_{i=-k}^k \lambda(n+i)x(n+i) \right] \geq 0. \end{aligned}$$

So

$$\begin{aligned} & \sum_{n \in Z} v(n) [S_m(\{x(n)\})(n) - S_g(\{x(n)\})(n)] \\ & \cdot \left[\sum_{i=-k}^k \lambda(n+i)x(n+i) \right] \geq 0. \end{aligned}$$

Therefore, for any $S_g \in \Omega_k$,

$$C(\{x(n)\}, S_g(\{x(n)\})) \leq C(\{x(n)\}, S_m(\{x(n)\})),$$

that is,

$$C(\{x(n)\}, S_m(\{x(n)\})) = \max_{S_g \in \Omega_k} \{C(\{x(n)\}, S_g(\{x(n)\}))\}$$

This completes the proof of Theorem 1.

This result shows that, in the sense of the generalized correlativity, the median filtering operator is optimal of the stack filtering operators. Also it is globally optimal, compared with the optimal design of L-filters.

4. Conclusion

In this paper we defined the generalized correlativity of input signal and output signal of a stack filtering operator, which can be used for numerously measuring these filtering operators's behavior in removing noise in signals. In the sense of the criterion of the generalized correlativity, the median filtering operator is optimal of stack filtering operators, which implies that this filtering operator possesses better filtering behavior than the others.

REFERENCES

- [1] J. W. Tukey, "Nonlinear (Nonsuperposable) Methods for Smoothing Data," *Proceedings of Congress Record EASCON*, Washington DC, 7-9 October 1974, p. 673.
- [2] H. X. Chen, R. K. Yang and M. Gabbouj, "On Root Structures and Convergence Properties of Weighted Median Filters," *Circuits and System Signal Processing*, Vol. 14, No. 6, 1995, pp. 735-747. [doi:10.1007/BF01204682](https://doi.org/10.1007/BF01204682)
- [3] U. Eckhardt, "Root Images of Median Filters," *Journal of Mathematical Imaging and Vision*, Vol. 19, No. 1, 2003, pp. 63-70. [doi:10.1023/A:1024489020930](https://doi.org/10.1023/A:1024489020930)
- [4] D. Eberly, H. Longbotham and J. Aragon, "Complete Classification of Roots to One-Dimensional Median and Rank-Order Filters," *IEEE Transactions on Signal Processing*, Vol. 39, No. 1, 1991, pp. 197-199. [doi:10.1109/78.80781](https://doi.org/10.1109/78.80781)
- [5] S. G. Tyan, "Median Filters: Deterinistic Properties," In: *Two-Dimension Digital Signal Processing II: Transforms and Median Filters*, Springer, Berlin, 1981. [doi:10.1007/BFb0057598](https://doi.org/10.1007/BFb0057598)
- [6] P.-T. Yu and W.-L. Wang, "Root Properties of Median Filters under There Appending Strategies," *IEEE Transactions on Signal Processing*, Vol. 41, No. 2, 1993, pp. 965-970. [doi:10.1109/78.193236](https://doi.org/10.1109/78.193236)
- [7] J. Brandt, "Cycles of Medians," *Utilitas Mathematica*, Vol. 54, 1998, pp. 111-126.
- [8] W. Z. Ye, L. Wang and L. G. Xu, "Properties of Locally Convergent Sequences with Respect to Median Filter," *Discrete Mathematics*, Vol. 309, No. 9, 2009, pp. 2775-2781. [doi:10.1016/j.disc.2008.07.002](https://doi.org/10.1016/j.disc.2008.07.002)
- [9] W. Z. Ye, M. T. Zhang and Y. L. Ma, "Structure of Recurrent Sequences of Median Filters," *Discrete Mathematics*, Vol. 310, No. 6-7, 2010, pp. 1253-1258. [doi:10.1016/j.disc.2009.12.005](https://doi.org/10.1016/j.disc.2009.12.005)
- [10] Z.-J. Gan and M Mao, "Two Convergence Theorems on Deterministic Properties of Median Filters," *IEEE Transactions on Signal Processing*, Vol. 39, No. 7, 1991, pp. 1689-1690. [doi:10.1109/78.134410](https://doi.org/10.1109/78.134410)
- [11] P. D. Wendt, E. J. Coyle and N. C. Gallagher Jr., "Stack Filter," *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 34, No. 4, 1986, pp. 898-911. [doi:10.1109/TASSP.1986.1164871](https://doi.org/10.1109/TASSP.1986.1164871)
- [12] W. Z. Ye and X. W. Zhou, "Criteria of Convergence of Median Filters," *IEEE Transactions on Signal Processing*, Vol. 49, No. 2, 2001, pp. 360-363. [doi:10.1109/78.902118](https://doi.org/10.1109/78.902118)

- [13] T. A. Nodes and N. C. Gallagher, "The Output Distribution of Median-Type Filters," *IEEE Transactions on Communication*, Vol. 32, No. 5, 1984, pp. 532-541. [doi:10.1109/TCOM.1984.1096099](https://doi.org/10.1109/TCOM.1984.1096099)
- [14] A. C. Bovik, T. S. Huang and D. C. Munson, "A generalization of Median Filtering Using Linear Combinations of Order Statistics," *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 31, No. 6, 1983, pp. 1342-1349. [doi:10.1109/TASSP.1983.1164247](https://doi.org/10.1109/TASSP.1983.1164247)
- [15] R. Oten and R. J. P. de Figueiredo, "An Efficient Method for L-Filter Design," *IEEE Transactions on Signal Processing*, Vol. 51, No. 1, 2003, pp. 193-203. [doi:10.1109/TSP.2002.806573](https://doi.org/10.1109/TSP.2002.806573)
- [16] J. P. Fitch, E. J. Coyle and N. C. Gallagher Jr., "Median Filtering by Threshold Decomposition," *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 32, No. 6, 1984, pp. 1183-1188.