Optimization of Threshold Boolean Filters

Ki Dong Lee and Yong Hoon Lee

Department of Electrical Engineering
Korea Advanced Institute of Science and Technology
373–1, Kusong–dong, Yusong–gu, Taejon 305–701, Korea
phone: +82–42–869–3437, fax: +82–42–869–3410
e-mail: yohlee@ee.kaist.ac.kr

Abstract

Recently, a new class of nonrecursive digital filters, which is called the *threshold Boolean filter* (TBF), has been introduced as an extension of stack filters. While a procedure for finding a TBF that yields a smaller mean absolute error (MAE) than an optimal stack filter is known, no efficient algorithm for finding the best TBF in the MAE sense is available at this time. As an alternative, the optimal TBF under the mean square error (MSE) criterion is considered in this paper. It is shown that the optimization problem can be formulated into a known problem: the minimization of a *pseudo-Boolean quadratic* function, which is an *NP-hard* problem. In this paper, a numerical method embedded in a *branch and bound* algorithm is adopted to solve the minimization problem. Some computer experiments with real signals are performed, and the results of optimal threshold Boolean filtering are compared with those of linear FIR Wiener and optimal stack filters.

1. Introduction

The threshold Boolean filter (TBF)[1] is an extension of stack filters[2]. While the class of stack filters is defined based on both the threshold decomposition and the stacking property, the TBF is based only on the threshold decomposition and does not necessarily possess the stacking property. It is shown in [3] that one can find very easily a TBF which has a smaller MAE than an optimal stack filter[4] minimizing the MAE. At this time, however, there exists no efficient algorithm for finding an optimal TBF under the MAE criterion. As an alternative, the problem of designing an optimal TBF under the mean square error (MSE) criterion is investigated in this paper.

In Section 2, the definition and representations of the TBF are briefly reviewed. The optimization problem minimizing the MSE is formally defined in Section 3. It is shown that the problem of finding an optimal TBF under the MSE criterion can be formulated as a minimization of a *pseudo-Boolean quadratic* function, which is known as an *NP-hard* problem[5]. An algorithm for solving the optimization problem, which is originally introduced in [6], is described in Section 4. Some computer experiments are performed with real signals, and the results are presented in Section 5. Finally, conclusions are given in Section 6.

2. Threshold Boolean Filters

The TBF is defined as

$$TBF_f(\vec{X}) = \sum_{t=1}^{M} f(I_t(\vec{X})), \qquad (1)$$

where $f(\cdot)$ is a Boolean function, $\vec{X} = (X_1, \cdots, X_N)$ is an input vector observed within a sliding window of size N, each input sample X_i is an $(M+1)$–valued integer, i.e., $X_i \in \{0, 1, \cdots, M\}$, and $I_t(\vec{X}) \equiv (I_t(X_1), \cdots, I_t(X_N))$ is the indicator function defined as $I_t(X_i) = 1$ if $X_i \geq t$ and 0, otherwise. When $f(\cdot)$ is a positive Boolean function, the TBF becomes a stack filter. An alternative TBF representation which is useful for the optimal design of TBF's is obtained from the following propositions[1].

Proposition 1. *Suppose that a Boolean function* $f : \{0,1\}^N \mapsto \{0,1\}$ *has a single true vector* $\vec{v} =$

$(v_1, \cdots, v_N) \in \{0,1\}^N$, i.e., $f(\vec{x}) = 1$ iff $\vec{x} = \vec{v}$ for any binary vector $\vec{x} \in \{0,1\}^N$. Then,

$$TBF_f(\vec{X}) = max[0, min\{X_j \mid j \in B_1(\vec{v})\} - max\{X_j \mid j \in B_0(\vec{v})\}], \quad (2)$$

where min and max of the empty set are defined to be M and 0, respectively, and $B_0(\vec{v})(B_1(\vec{v}))$ is the set of all indices of v_i, $i = 1, 2, \cdots, N$ which are zero(one). ∎

Example 1. Suppose $N = 3$ and $f(x_1, x_2, x_3) = x_1\bar{x}_2x_3$. Then the only true vector of $f(\vec{x})$ is $(1,0,1)$, $B_0(\vec{v}) = \{2\}$ and $B_1(\vec{v}) = \{1,3\}$. Thus, $TBF_f(\vec{X}) = max\{0, min\{X_1, X_3\} - X_2\}$. ∎

An arbitrary Boolean function $f(\cdot)$ is specified by its truth table which lists all possible binary input sequences and the corresponding outputs. Let $\{\vec{x}_i,\ i = 1, \cdots, 2^N\} \equiv \chi$ denote all the possible binary vectors in $\{0,1\}^N$, where each $\vec{x}_i$ is the radix–2 representation of $i - 1$. The output of a Boolean function $f(\cdot)$ for $\vec{x}_i \in \chi$ is denoted as a_i: that is, $f(\vec{x}_i) = a_i$, $i = 1, \cdots, 2^N$ where $a_i \in \{0,1\}$. For each $\vec{x}_k \in \chi$, we specify a Boolean function $f_k : \{0,1\}^N \mapsto \{0,1\}$ which has only one true vector $\vec{x}_k$. Note that $f_k(\vec{x})$ is uniquely determined. For example, if $N = 3$ and $\vec{x}_k = (1,0,1)$, then $f_k(\vec{x}) = x_1\bar{x}_2x_3$. Now we are ready to describe the following multi-level representation of the TBF.

Proposition 2. *A multi-level representation of the TBF corresponding to a general Boolean function $f(\cdot)$ is*

$$TBF_f(\vec{X}) = \sum_{k=1}^{2^N} a_k W_k(\vec{X}) = \vec{a} \cdot \vec{W}(\vec{X}) \quad (3)$$

where $\vec{a} = (a_1, \cdots, a_{2^N})$, and $W_k(\vec{X}) = TBF_{f_k}(\vec{X})$ which is the multi-level TBF expression of $f_k(\vec{x})$. ∎

Since $f_k(\vec{x})$ is unique for each $\vec{x}_k$, $k = 1, \cdots, 2^N$, $W_k(\vec{X})$ is also determined uniquely for each k. The example presented below illustrates $W_k(\vec{X})$.

Example 2. Consider the case $N = 3$. Then the 8 $(= 2^3)$ entries of $\vec{W}(\vec{X})$ are

$W_1(\vec{X}) = M - max\{X_1, X_2, X_3\}$,
$W_2(\vec{X}) = max\{0, X_3 - max\{X_1, X_2\}\}$,
$W_3(\vec{X}) = max\{0, X_2 - max\{X_1, X_3\}\}$,
$W_4(\vec{X}) = max\{0, min\{X_2, X_3\} - X_1\}$,
$W_5(\vec{X}) = max\{0, X_1 - max\{X_2, X_3\}\}$,
$W_6(\vec{X}) = max\{0, min\{X_1, X_3\} - X_2\}$,
$W_7(\vec{X}) = max\{0, min\{X_1, X_2\} - X_3\}$, and
$W_8(\vec{X}) = min\{X_1, X_2, X_3\}$. ∎

3. Optimization Problem

Suppose that $\mathbf{R}(n) = \mathbf{S}(n) + \mathbf{N}(n)$ where $\mathbf{S}(n)$ and $\mathbf{R}(n)$ are a desired signal and an observed signal contaminated by the additive noise $\mathbf{N}(n)$, respectively. It is assumed that both $\mathbf{S}(n)$ and $\mathbf{R}(n)$ are $(M+1)$-valued, i.e., they take values from $\{0, 1, \cdots, M\}$. The observation within a window located at n is denoted by $\vec{X}(n) = (\mathbf{R}(n-L), \cdots, \mathbf{R}(n), \cdots, \mathbf{R}(n+L))$ where the window size $N = 2L+1$. Now we want to find the TBF which minimizes the MSE $E[|\mathbf{S} - TBF_f(\vec{X})|^2]$, where the time index n has been dropped for the notational simplicity. By using (3), we get the following optimization problem:

Optimization #1.

$$\underset{\vec{a}}{minimize}\ J(\vec{a}) = E[|\mathbf{S} - \vec{a} \cdot \vec{W}(\vec{X})|^2]$$

$$\text{subject to } a_k \in \{0,1\},\ k = 1, \cdots, 2^N.$$

The object function $J(\vec{a})$ in Optimization #1 is given by

$$J(\vec{a}) = E[\mathbf{S}^2] - 2\vec{a} \cdot E[\mathbf{S}\vec{W}(\vec{X})] + \vec{a}E[\vec{W}(\vec{X})\vec{W}^t(\vec{X})]\vec{a}^t \equiv K - 2\vec{a} \cdot \vec{\phi} + \vec{a}\Psi\vec{a}^t, \quad (4)$$

where $K \equiv E[\mathbf{S}^2]$, $\vec{\phi} \equiv E[\mathbf{S}\vec{W}(\vec{X})]$, $\Psi \equiv E[\vec{W}(\vec{X})\vec{W}^t(\vec{X})]$, and the superscript t denotes the transposition of a row vector. Note that, $\vec{\phi}$ is a $2^N \times 1$ row vector and Ψ is a $2^N \times 2^N$ matrix. The entries of $\vec{\phi}$ and Ψ are all nonnegative because $\mathbf{S}$ is a nonnegative integer and $W_k(\vec{X}) \geq 0$ for all k and $\vec{X}$. Furthermore, $\Psi \equiv [\psi_{ij}]$ is symmetric, i.e., $\psi_{ij} = \psi_{ji}$. If $\vec{x}_i$ and $\vec{x}_j$ are incomparable, $\psi_{ij} = \psi_{ji} = 0$ since $W_i(\vec{X})$ and $W_j(\vec{X})$ cannot have nonzero values simultaneously.

Example 3. Consider the case where $N = 3$. Since the binary sequences $\vec{x}_i \in \chi$ are given by

$\vec{x}_1 = (0,0,0)$, $\vec{x}_2 = (0,0,1)$, $\vec{x}_3 = (0,1,0)$,
$\vec{x}_4 = (0,1,1)$, $\vec{x}_5 = (1,0,0)$, $\vec{x}_6 = (1,0,1)$,

$\vec{x}_7 = (1,1,0)$, and $\vec{x}_8 = (1,1,1)$, the matrix Ψ has the form

$$\begin{pmatrix} \psi_{11} & \psi_{12} & \psi_{13} & \psi_{14} & \psi_{15} & \psi_{16} & \psi_{17} & \psi_{18} \\ \psi_{21} & \psi_{22} & 0 & \psi_{24} & 0 & \psi_{26} & 0 & \psi_{28} \\ \psi_{31} & 0 & \psi_{33} & \psi_{34} & 0 & 0 & \psi_{37} & \psi_{38} \\ \psi_{41} & \psi_{42} & \psi_{43} & \psi_{44} & 0 & 0 & 0 & \psi_{48} \\ \psi_{51} & 0 & 0 & 0 & \psi_{55} & \psi_{56} & \psi_{57} & \psi_{58} \\ \psi_{61} & \psi_{62} & 0 & 0 & \psi_{65} & \psi_{66} & 0 & \psi_{68} \\ \psi_{71} & 0 & \psi_{73} & 0 & \psi_{75} & 0 & \psi_{77} & \psi_{78} \\ \psi_{81} & \psi_{82} & \psi_{83} & \psi_{84} & \psi_{85} & \psi_{86} & \psi_{87} & \psi_{88} \end{pmatrix}$$

where ψ_{ij} are all positive. ▮

Assume that $\vec{\phi}$ and Ψ are known. The decision vector $\vec{a}$ minimizing $J(\vec{a})$ in (4) also minimizes the following pseudo-Boolean quadratic function[8][9] $g(\vec{a}) = \frac{1}{2}\vec{a}\Psi\vec{a}^t - \vec{a}\cdot\vec{\phi}$ and *vice versa*. Therefore, Optimization #1 can be restated equivalently as follows:

Optimization #2.

$$\underset{\vec{a}}{minimize}\ g(\vec{a}) = \frac{1}{2}\vec{a}\Psi\vec{a}^t - \vec{a}\cdot\vec{\phi} \qquad (5.a)$$

$$\text{subject to } \vec{a} \in \{0,1\}^N. \qquad (5.b)$$

Without the constraint in (5.b), the solution to this optimization would be $\vec{a} = \vec{\phi}\Psi^{-1}$. Due to the constraint, however, Optimization #2 becomes an NP-hard problem[5]. Thus, there exists no polynomial algorithm to solve the optimization except for some special cases[7][8], and a numerical search or heuristic technique should be applied. In [6],[9] and [10], algorithms for solving Optimization #2 are proposed. Among them we use one of the algorithms in [6] to optimize the TBF. The allgorithm, which is relatively simpler to implement, will be reviewed in the following section.

4. A Branch and Bound Algorithm

Carter[6] used a method for transforming an indefinite pseudo-Boolean quadratic problem into an equivalent positive definite problem, and presented three variations of the transformation which are embedded in a branch and bound algorithm. Algorithm COLALL, which was experimentally proved to be the best among the three in [6], is adopted in this paper and reviewed below.

Since $a_i^2 = a_i$ for every zero-one solution, the transformation

$$\tilde{\psi}_{ii} = \psi_{ii} - \rho \ \text{ and } \ \tilde{\phi}_i = \phi_i - \frac{1}{2}\rho \qquad (6)$$

for a constant ρ preserves the function values $g(\vec{a})$ at any $\vec{a} \in \{0,1\}^N$. Assume that Ψ is positive definite, and let $\vec{a}^*$ be the real-valued unconstrained solution to (5.a). If the i-th diagonal element of Ψ is decreased by ρ, assuming ρ is sufficiently small enough such that the new matrix $\tilde{\Psi}$ is still positive definite, then the unconstrained minimum will be changed by

$$\Delta g = \frac{\rho}{2}\left[\frac{(\frac{1}{2} - a_i^*)^2}{\rho h_{ii} - 1} + \frac{1}{4}\right] \qquad (7)$$

where h_{ii} represents the i-th diagonal element of $H \equiv \Psi^{-1}$. The effect of decreasing ψ_{ii} by ρ on Optimization #2 can be summarized as follows:

Algorithm CHDIAG(i,ρ)

Step 1. Set $\psi_{ii} = \psi_{ii} - \rho$ and $\phi_i = \phi_i - \frac{1}{2}\rho$.
Step 2. $\tau = \rho/(\rho h_{ii} - 1)$.
Step 3. $H = H - \tau\vec{h}_i\vec{h}_i^t$.
Step 4. $\vec{a}^* = \vec{a}^* + \tau\vec{h}_i(\frac{1}{2} - a_i^*)$. ▮

In Step 3, $\vec{h}_i$ represents the i-th column vector of H.

For any $\vec{a}^I \in \{0,1\}^N$, the convex set $F = \{\vec{a} \mid g(\vec{a}) \le g(\vec{a}^I)\}$ is bounded by $l_j \le a_j \le u_j$ where $l_j = a_j^* - \sqrt{2h_{jj}(g(\vec{a}^I) - g(\vec{a}^*))}$ and $u_j = a_j^* + \sqrt{2h_{jj}(g(\vec{a}^I) - g(\vec{a}^*))}$. Thus, the square of the half-width of the spread between the upper and lower bounds on a_j is given by $s_j^2 = 2h_{jj}(g(\vec{a}^I) - g(\vec{a}^*))$. If ψ_{ii} is decreased by ρ, the new spread $\tilde{s}_j^2 = 2\tilde{h}_{jj}(g(\vec{a}^I) - g(\vec{a}^*) - \Delta g)$ can be expressed as a function of ρ by using the relationship $\tilde{h}_{jj} = h_{jj} - \tau h_{ij}^2$ in CHDIAG, the reverse relationship $g(\vec{a}^I) - g(\vec{a}^*) = s_i^2/2h_{ii}$ from the definition of s_i^2, and Δg in (7). Let $\vec{a}^I$ be the current best feasible zero-one solution. Then for each i, we choose ρ, the amount of the change in the i-th diagonal element of Ψ, minimizing $\sum_j \tilde{s}_j^2$. The algorithm is summarized as follows:

Algorithm COLALL

Repeat for $i = 1$ to 2^N:
Step 1. Compute $s_i^2 = 2h_{ii}(g(\vec{a}^I) - g(\vec{a}^*))$.
Step 2. Calculate
$\alpha_1 = \frac{1}{4}(\sum_j h_{ij}^2 - h_{ii}\sum_j h_{jj})h_{ii}^2$,
$\alpha_2 = \frac{3}{4}h_{ii}(h_{ii}\sum_j h_{jj} - \sum_j h_{ij}^2)$,
$\alpha_3 = -h_{ii}\sum_j h_{jj}[\frac{3}{4} - (\frac{1}{2} - a_i^*)^2] + \sum_j h_{ij}^2[s_i^2 - 2(\frac{1}{2} - a_i^*)^2 + \frac{1}{2}]$,
$\alpha_4 = [h_{ii}\sum_j h_{ij}a_i^*(1 - a_i^*) - \sum_j h_{ij}^2 s_i^2]/h_{ii}$.
Step 3. Let $u = \alpha_2^2 - 3\alpha_1\alpha_3$.

Step 4. If $u \leq 0$, let $\rho_0 = 0$.
Otherwise, $\rho_0 = 1/h_{ii} + 2\sqrt{u}/3\alpha_1$.

Step 5. Use *Newton's method* to find a zero for the function $\alpha_1\rho^3 + \alpha_2\rho^2 + \alpha_3\rho + \alpha_4 = 0$ beginning at ρ_0.

Step 6. Use algorithm CHDIAG(i, ρ) to change the i-th diagnal of Ψ by ρ. ∎

For more details, see [6].

5. Computer Simulations

As an example for designing an optimal TBF under the MSE criterion by using COLALL algorithm, some experimental results with real signals are presented below. We shall compare the results for the optimal TBF minimizing the MSE with those of linear FIR Wiener filters and of the optimal stack filters minimizing the MAE.

The original signal with 256 gray levels has been taken from the 128th line of 256 × 256 *Lena* image. Three different noisy versions are synthesized by adding Gaussian, impulsive, and both Gaussian and impulsive noises, respectively. The generated Gaussian noise has zero-mean and variance of 15^2, $G(0, 15^2)$. The impulsive noise which consists of positive and negative impulses with values 200 and −200, respectively, occurs with probability $P_e = 0.2$. The compound noise is composed of Gaussian noise with $G(0, 10^2)$ and impulsive noise with $P_e = 0.1$. For each case, the necessary statistics required for the optimal linear FIR, stack, and TBF filterings are estimated from the original and noisy signals. The MSE and MAE between the original and the filtered signals are evaluated. The results are summarized in Table 1. The FIR Wiener filter resulted in the smallest MSE and MAE under the impulse-free environment, but it performed poorer than the TBF's and the stack filters when impluses exist. It is interesting to note that the MAE's (MSE's) of the stack filters are always less (greater) than those of the TBF's. Now we compare the subjective quality of the filtered signals. Fig. 1 illustrates the results for the suppression of impulsive noise (the second case in Table 1). Among the three filtered signals, the result for the TBF looks the best: the FIR Wiener filter severely blurred the signal and the stack filter failed to suppress all impulses.

6. Conclusions

A design of an optimal TBF under the MSE criterion has been investigated. It was shown that the optimization problem can be formulated into a known problem, which is the minimization of a pseudo-Boolean quadratic fnction. To solve the problem, the algorithm COLALL proposed by Carter[6] was used. Computer experiments have been performed to verify the procedure for optimizing the TBF. The performance of the designed TBF was compared with those of the FIR Wiener and of the optimal stack filters. The results indicate that the TBF outperforms the others in suppressing impulses.

If the stacking constraint is additionally imposed on Optimization #2, an optimal stack filter under the MSE criterion can be obtained. Finding an efficient algorithm for this case and a simpler algorithm for optimizing the TBF requires further research.

References

[1] K. D. Lee and Y. H. Lee, "Threshold Boolean filters," *Proc. Canadian Conf. Electr. Comput. Eng.*, Toronto, Sep. 1992, *also submitted to IEEE Trans. Signal Processing.*

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

[3] J. Song and Y. H. Lee, "Linear combination of weighted order statistic filters: an extension of stack filters," *Proc. 1992 Conf. Inform. Sci. Syst.*, Princeton, Mar. 1992.

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

[5] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness,* San Francisco: Freeman, 1979.

[6] M. W. Carter, "The indefinite zero-one quadratic problem," *Discrete Appl. Math.*, vol. 7, pp. 23–44, 1984.

[7] J. C. Picard and H. D. Ratliff, "Minimum cuts and related problems," *Networks,* 5, pp. 357–370, 1974.

[8] F. Barahona, "A solvable case of quadratic 0–1 programming," *Discrete Appl. Math.*, vol. 13, pp. 23–26, 1986.

[9] F. Barahona, M. Junger, and G. Reinelt, "Experiments in quadratic 0–1 programming," *Math. Programming*, vol. 44, pp. 127–137, 1989.

[10] H. Lee, *Maximizing a Submodular Function by Integer Programming: A Polyhedral Approach*, Ph. D. Dissertation, Georgia Institute of Technology, 1991.

	Gaussian		impulsive		Gaussian + impulsive	
	MAE	MSE	MAE	MSE	MAE	MSE
FIR Wiener	8.53	113.39	25.48	1225.54	17.13	733.41
TBF	9.54	147.26	11.77	826.89	14.11	664.85
stack	9.39	153.58	11.75	960.74	11.94	708.50

Table 1. The MAE's and MSE's between the original and filtered signals.

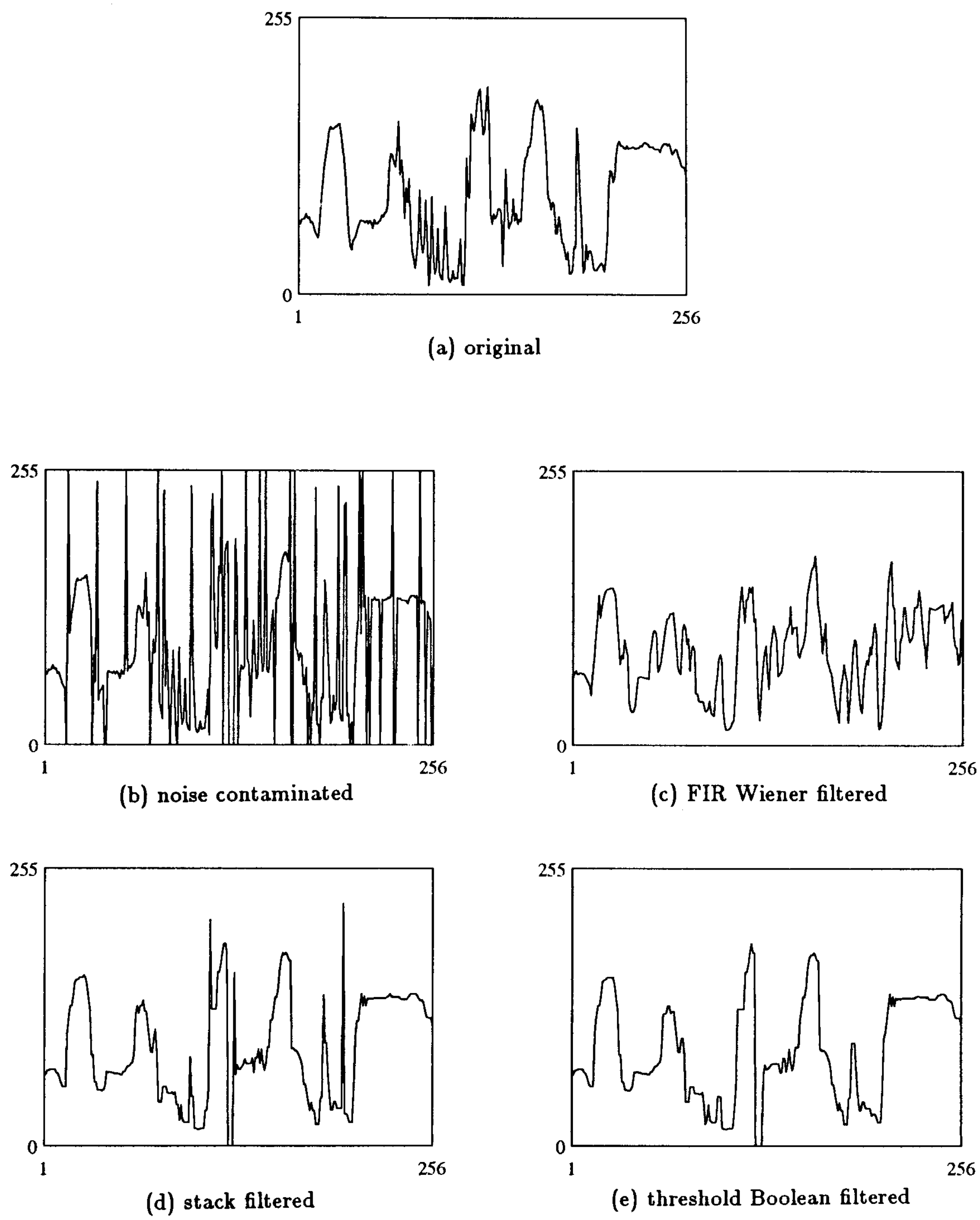

Fig. 1. The results of filtering for impulsive noise with $P_e = 0.2$.