Mem. Faculty. B. O. S. T. Kinki University No. 36 : 9 ～ 18 (2015)

Which is better Median Filter or Linear Filter as a Smoother?

Sho Kikkawa, Takeshi Kohama, Hisashi Yoshida

Abstract

It is believed that median filtering outperforms traditional linear filtering in preserving signal edges while reducing noise. However, this does not always mean that the median filter is superior to the linear filter as a smoother. By comparing the bias-variance characteristics of median filtering and linear filtering, we demonstrate that a linear filter (moving average filter) is a better smoother than the median filter for inputs that have Gaussian or uniform distributions and relatively small signal-to-noise ratios, whereas the median filter outperforms the linear filter for inputs with Laplace distributions.

Key words: median filter, linear filter, smoother, edge, Laplace distribution, Gaussian distribution, uniform distribution, bias-variance characteristic.

1. Introduction

Linear filtering cannot preserve edges when smoothing noise in the presence of edges. By contrast, the median filtering approach introduced by Tukey (1) is believed to be capable of eliminating noise without smearing edges(2, 3, 4, 5), and it has been extensively used in the fields of image processing and signal processing. However, Arias-Castro & Donoho have cast doubt on the abilities of median filtering (6). They reported that for an input signal consisting of a unit step signal and random Gaussian noise, the bias of the output of the median filter is as large as that of a linear filter (moving average filter) for an input signal-to-noise ratio (SNR) of order 1, whereas the bias becomes considerably less and finally vanishes as the SNR $\to \infty$. However, they didn't mention whether the median filter can be outperformed by the linear filter as a smoother or not. On the other hand, Justusson showed that for Gaussian noise, the output mean square error (MSE) at the points in a region of filter-width around edges for the 3-points linear filter is somewhat smaller than the MSE for the 3-points median filter for SNR < 2 and for SNR > 3, the MSE of the median is considerably smaller than that of the linear filter (7). Although these results are important, they are not enough, only by themselves, to properly use the filter, linear or median, according to a situation.

In this work, we evaluate the bias-variance characteristics of median filtering and linear filtering and demonstrate that the linear filter is a better smoother than the median filter for inputs that have Gaussian or uniform distributions and relatively small SNRs, although median filtering outperforms linear filtering for inputs with Laplace distributions.

2. Output bias and variance

Denote the window width of a filter by $M = 2w + 1$, where w is a non-negative integer, and a discrete input signal by $x(t)$. Assume that $x(t)$ for $-(N + w) \leq t \leq -1$ is an i.i.d. random sequence that has a probability density function (pdf) of $g_1(x)$ with mean λ_1 and variance σ_1^2 and that $x(t)$ for $0 \leq t \leq N + w - 1$ is another i.i.d. random sequence that has a pdf of $g_2(x)$ with mean λ_2 and variance σ_2^2. That is, the pdf $g(x, t)$ of the input $x(t)$ is given by

$$g(x,t) = \begin{cases} g_1(x) & (-(N+w) \leq t \leq -1), \\ g_2(x) & (0 \leq t \leq N + w - 1). \end{cases} \tag{1}$$

Received 1 May 2015

This work was partially supported by JSPS KAKENHI Grant numbers 25420422, 26420401, 20560406 and 24560534.

Department of Computational Systems Biology, Faculty of Biology Oriented Science and Technology, Kinki University, Wakayama 649-6493, Japan

The input $x(t)$ is written as

$$x(t) = s(t) + n(t), \tag{2}$$

where $s(t) = \lambda_1$ $(-(N+w) \le t \le -1), = \lambda_2$ $(0 \le t \le N+w-1)$ and $n(t) = x(t) - s(t)$ have a pdf with a zero mean and a variance of σ_1^2 for $-(N+w) \le t \le -1$ and σ_2^2 for $0 \le t \le N+w-1$.

2. 1. Linear filter

Here, we consider a moving average filter. The running filter window slides over the input signal from $t = -N$ to $t = N-1$, yielding the output signal $y_{\mathrm{L}}(t) = (1/M)\sum_{k=-w}^{w} x(t+k)$. The output $y_{\mathrm{L}}(t)$ is expressed as

$$y_{\mathrm{L}}(t) = s_{\mathrm{L}}(t) + n_{\mathrm{L}}(t), \tag{3}$$

where $s_{\mathrm{L}}(t) = (1/M)\sum_{k=-w}^{w} s(t+k)$ and $n_{\mathrm{L}}(t) = (1/M)\sum_{k=-w}^{w} n(t+k)$. By defining the mean variance per unit time $\overline{\sigma_{y\mathrm{L}}^2}$ and the mean square bias per unit time $\overline{b_{y\mathrm{L}}^2}$ as $\overline{\sigma_{y\mathrm{L}}^2} = \sum_{t=-N}^{N-1} \mathrm{E}[\{y_{\mathrm{L}}(t) - \mathrm{E}[y_{\mathrm{L}}(t)]\}^2]/2N$ and $\overline{b_{y\mathrm{L}}^2} = \sum_{t=-N}^{N-1} \mathrm{E}[\{s(t) - \mathrm{E}[y_{\mathrm{L}}(t)]\}^2]/2N$, respectively, we obtain the relationship between $\overline{\sigma_{y\mathrm{L}}^2}$ and $\overline{b_{y\mathrm{L}}^2}$ as

$$\overline{\sigma_{y\mathrm{L}}^2} = \frac{-12N\overline{b_{y\mathrm{L}}^2} + \sqrt{(12N\overline{b_{y\mathrm{L}}^2})^2 + \Delta\lambda^4}}{\Delta\lambda^2}\overline{\sigma_0}^2, \tag{4}$$

where $\Delta\lambda = \lambda_2 - \lambda_1$ and $\overline{\sigma_0^2} = (\sigma_1^2 + \sigma_2^2)/2$. The relationship (4) is independent of the distribution type.

2. 2. Median filter

The output of the median filter is given by

$$y_{\mathrm{M}}(t) = \mathrm{med}\{x(t-w), \ldots, x(t), \ldots, x(t+w))\}. \tag{5}$$

The pdf of $y_{\mathrm{M}}(t)$ is given by

$$f_{\mathrm{M}}(x,t) = \begin{cases} f_{\mathrm{M}}^{(1)}(x,t')_{t'=-(w+1)} & (-N \le t \le -(w+1)), \\ f_{\mathrm{M}}^{(1)}(x,t) + f_{\mathrm{M}}^{(2)}(x,t) & (-(w+1) \le t \le w), \\ f_{\mathrm{M}}^{(2)}(x,t')_{t'=w} & (w \le t \le N-1]), \end{cases} \tag{6}$$

where $f_{\mathrm{M}}^{(1)}(x,t)$ and $f_{\mathrm{M}}^{(2)}(x,t)$ are given by

$$f_{\mathrm{M}}^{(1)}(x,t) = \begin{cases} (w-t)g_1(x)\sum_{k=-t-1}^{w} \binom{w-t-1}{k}\binom{w+t+1}{w-k} & \\ \quad \cdot G_1^k(x)(1-G_1(x))^{w-t-1-k}G_2^{w-k}(x)(1-G_2(x))^{k+t+1} & (-(w+1) \le t \le -1), \\ (w-t)g_1(x)\sum_{k=0}^{w-t-1} \binom{w-t-1}{k}\binom{w+t+1}{w-k} & \\ \quad \cdot G_1^k(x)(1-G_1(x))^{w-t-1-k}G_2^{w-k}(x)(1-G_2(x))^{k+t+1} & (0 \le t \le w), \end{cases} \tag{7}$$

and

$$f_{\mathrm{M}}^{(2)}(x,t) = \begin{cases} (w+t+1)g_2(x)\sum_{k=-t}^{w} \binom{w-t}{k}\binom{w+t}{w-k} & \\ \quad \cdot G_1^k(x)(1-G_1(x))^{w-t-k}G_2^{w-k}(x)(1-G_2(x))^{k+t} & (-(w+1) \le t \le -1), \\ (w+t+1)g_2(x)\sum_{k=0}^{w-t} \binom{w-t}{k}\binom{w+t}{w-k} & \\ \quad \cdot G_1^k(x)(1-G_1(x))^{w-t-k}G_2^{w-k}(x)(1-G_2(x))^{k+t} & (0 \le t \le w), \end{cases} \tag{8}$$

where $G_i(x)$ $(i = 1, 2)$ are the cumulative distribution functions (cdfs) of $g_i(f)$. Eqs.(7) and (8) coincide with the results by Justusson [7]. While he also gives the proof [7], we will show more detail proof in Appendix A.
The average $\mathrm{E}[y_{\mathrm{M}}(t)]$ and the variance $\sigma^2_{y\mathrm{M}}(t)$ of the output of the median filter are obtained by

$$\mathrm{E}[y_{\mathrm{M}}(t)] = \int_{-\infty}^{\infty} x f_{\mathrm{M}}(x,t)dx, \tag{9}$$

$$\sigma^2_{y\mathrm{M}}(t) = \int_{-\infty}^{\infty} (x - E[y_{\mathrm{M}}(t)])^2 f_{\mathrm{M}}(x,t)dx. \tag{10}$$

2. 3. Bias-variance analysis

One of the simplest and most practically useful models of $g(x,t)$ is that in which, for a zero-mean symmetrical pdf $g_0(x)$ with variance σ_0^2, the following assumption holds

$$\begin{cases} g_1(x) = & g_0(x+\lambda), \\ g_2(x) = & g_0(x-\lambda). \end{cases} \tag{11}$$

Under this assumption, we obtained the mean $\mathrm{E}[y_{\mathrm{M}}(t)]$ and the variance $\sigma^2_{y\mathrm{M}}(t)$ by numerically calculating (9) and (10) for $\lambda_1 = -\lambda = -1/2, \lambda_2 = \lambda = 1/2$, $\sigma_1^2 = \sigma_2^2 = \sigma_0^2$, $-N \le t \le N-1$ $(N = 971)$ and $w = 2^n (n = 0, 1, \ldots, 7)$. The integrals were approximated by

$$\mathrm{E}[y_{\mathrm{M}}(t)] \approx \sum_{k=-K}^{K-1} x_k f_{\mathrm{M}}(x_k, t)\Delta x, \tag{12}$$

$$\sigma^2_{y\mathrm{M}}(t) \approx \sum_{k=-K}^{K-1} (x_k - E[y_{\mathrm{M}}(t)])^2 f_{\mathrm{M}}(x_k, t)\Delta x \tag{13}$$

where $x_k = k\Delta x, \Delta x = 10^{-5}\sigma_0, K = [(1/2 + c\sigma_0)/\Delta x]$ and c was selected corresponding to the pdf $g_0(x)$ as $c = 5$ for a uniform distribution, $c = 10$ for a Gaussian distribution and $c = 30$ for a Laplace distribution.
There is a trade-off between the variance and bias of the filters. Similar to the case of the linear filter, let us define the mean square bias and the mean variance per unit time of the median filter output $y_{\mathrm{M}}(t)$ by

$$\overline{b^2_{y\mathrm{M}}}(w) = \frac{1}{2N} \sum_{t=-N}^{N-1} \{s(t) - \mathrm{E}[y_{\mathrm{M}}(t)]\}^2, \tag{14}$$

$$\overline{\sigma^2_{y\mathrm{M}}}(w) = \frac{1}{2N} \sum_{t=-N}^{N-1} \{y_{\mathrm{M}}(t) - \mathrm{E}[y_{\mathrm{M}}(t)]\}^2 \tag{15}$$

Fig.1 presents a comparison of the bias-variance characteristics ($\overline{b^2_y} - \overline{\sigma^2_y}$ characteristics) of median filtering with those of linear filtering for three different types of distributions, $g_0(x)$s, i.e., uniform, Gaussian, and Laplace distributions. The calculations of the bias-variance characteristics were performed for different σ_0^2s for each distribution. The abscissa shows the mean square bias $\overline{b^2_{y*}}$ and the ordinate shows the mean variance $\overline{\sigma^2_{y*}}$ standardized by the input variance σ_0^2, where the symbol $*$ refers to L and M corresponding to the linear filter and the median filter, respectively. From the top panel to bottom, the results for a uniform distribution, Gaussian distribution and Laplace distribution are presented. This figure shows that although median filtering is superior to linear filtering for Laplace inputs with any input variance σ_0^2, the median filter is inferior to the linear filter for uniform inputs with variances greater than 0.3^2 and for Gaussian inputs with variances greater than approximately 0.4^2. In other words, the linear filter is superior to the median filter as a smoother for Gaussian and uniform distributions when the input SNR is not very large.
Additionally, we performed simulations to validate the results of the above theoretical study. For each distribution $g_0(x)$, input random sequences $\{x_i(t)\}_{i=1,\ldots,R}$ of length $2N'$ $(-N' \le t \le N'-1)$ were generated using Matlab, where

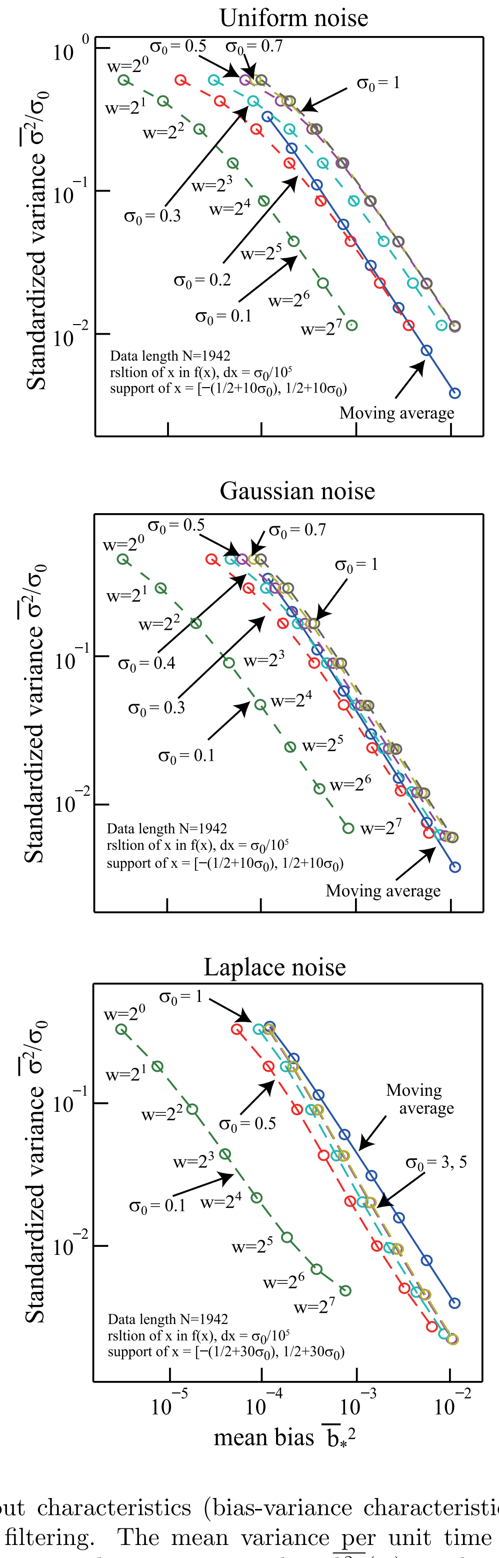

Figure 1: Comparison of the output characteristics (bias-variance characteristics, $\overline{b_y^2} - \overline{\sigma_y^2}$ characteristics) of median filtering with those of linear filtering. The mean variance per unit time normalized by the input variance $\overline{\sigma_{y*}^2}(w)/\sigma_0^2$ is plotted on the y-axis versus the mean square bias $\overline{b_{y*}^2}(w)$ on the x-axis for different input variances σ_0^2. The circle marks ($\circ$) on each dashed and solid line are the coordinate points that respectively correspond to $w = 2^0, 2^1, \ldots, 2^7$ from the largest value of $\overline{\sigma_{y*}^2}(w)/\sigma_0^2$ (the smallest $\overline{b_{y*}^2}(w)$) to the smallest of $\overline{\sigma_{y*}^2}(w)/\sigma_0^2$ (the largest $\overline{b_{y*}^2}(w)$). The dashed lines and the solid line with marks in each panel refer to the median filtering $\overline{b_y^2} - \overline{\sigma_y^2}$ characteristics and linear filtering $\overline{b_y^2} - \overline{\sigma_y^2}$ characteristics, respectively. $2N = 1942$, $\Delta x = \sigma_0/10^5$, $K = [(1/2 + c\sigma_0)/\Delta x], c = 5(\text{uniform distribution}), = 10(\text{Gaussian distribution}), = 30(\text{Laplace distribution})$.

$R = 10^4$ and N' is sufficiently larger than $N = 971$. The output data $\{y_{*i}(t)\}_{i=1,\dots,R}$ were shortened from $2N'$ to $2N = 1942$ by removing some data points from both ends to avoid end effects. The mean $\mathrm{E}[y_*(t)]$ and the variance $\sigma^2_{y_*}(t)$ of output $y_*(t)$ were estimated by

$$\mathrm{E}[y_*(t)] \approx \frac{1}{R}\sum_{i=1}^{R} y_{*i}(t), \tag{16}$$

$$\sigma^2_{y_*}(t) \approx \frac{1}{R}\sum_{i=1}^{R}(y_{*i}(t) - E[y_*(t)])^2. \tag{17}$$

The simulation confirmed the validity of the numerical study; that is, the results obtained using (12) and (13) were virtually identical to the results obtained using (16) and (17).

3. Conclusions

In order to use the right filter in real situations, we compared the performance of the median filter to that of the linear filter as a smoother for input signals that contain an abrupt change buried in random noise. We found that linear filtering is superior to median filtering for inputs with short-tailed distributions (uniform distributions and Gaussian distributions) when the SNR of the input is not very high, whereas median filtering is superior to linear filtering for inputs with long-tailed distributions (Laplace distributions) or for inputs with very high SNRs when the distribution of the input is not long-tailed (Gaussian distribution or uniform distribution). This result indicates that careful attention should be paid to the type of distribution and SNR of the input and is useful when judging in practice which of the linear fiter or the median filter should be used.

A Derivations of (6), (7), and (8)

The pdf $f_{\mathrm{M}}(x,t)$ of the median filter output depends on only $g_1(x)$ ($g_2(x)$) for $-N \le t \le -(w+1)$ $(w \le t \le N-1)$ and on both $g_1(x)$ and $g_2(x)$ for $-w \le t \le w-1$.

A-1 $-N \le t \le -(w+1)$ $(w \le t \le N-1)$

For $-N \le t \le -(w+1)$ $(w \le t \le N-1)$, $x(t-w), \dots, x(t), \dots, x(t+w)$ are i.i.d. random variables that have the common pdf $g_1(x)$ $(g_2(x))$. Then, the pdf $f_{\mathrm{M}}(x,t)$ is given [8, 4] by

$$f_{\mathrm{M}}(x,t) = \begin{cases} \frac{M!}{w!w!} g_1(x) G_1^w(x)(1-G_1(x))^w & (-N \le t \le -(w+1)), \\ \frac{M!}{w!w!} g_2(x) G_2^w(x)(1-G_2(x))^w & (w \le t \le N-1). \end{cases} \tag{18}$$

A-2 $-(w+1) \le t \le w$

Divide the sample set $S = \{x(t-w), \dots, x(t),\ \dots, x(t+w))\}$ $(-(w+1) \le t \le w)$ into two subsets: $S_1 = \{x(t-w), \dots, x(-1)\}$ of $w-t$ samples and $S_2 = \{x(0), \dots, x(t+w)\}$ of $w+t+1$ samples. For a sample $x(t) \in S$, the probability of the event $\{x < x(t) \le x + \Delta x\}$ is given by

$$Pr[x < x(t) \le x + \Delta x] = \begin{cases} (w-t) g_1(x)\Delta x + O(\Delta x^2) & (x(t) \in S_1), \\ (w+t+1) g_2(x)\Delta x + O(\Delta x^2) & (x(t) \in S_2). \end{cases} \tag{19}$$

The w order statistics of S for which the orders are smaller than $w+1$ must be less than or equal to $y_{\mathrm{M}}(t)$ because $y_{\mathrm{M}}(t)$ is the $(w+1)$th-order statistic. In the following, assume that the k order statistics of the w order statistics of S belong to S_1 and that the other $w-k$ belong to S_2.

1. Assume $y_{\mathrm{M}}(t) \in S_1$. Because k order statistics in S_1 that are less than or equal to $y_{\mathrm{M}}(t)$ can occur in $\binom{w-t-1}{k}$

different ways, the probability of the event that these k samples are less than or equal to x is given by

$$\binom{w-t-1}{k} G_1^k(x)(1-G_1(x))^{w-t-1-k}. \tag{20}$$

Furthermore, because the $w-k$ order statistics in S_2 can occur in $\binom{w+t+1}{w-k}$ different ways, the probability of the event occurring in which they are less than or equal to x is given by

$$\binom{w+t+1}{w-k} G_2^{w-k}(x)(1-G_2(x))^{k+t+1}. \tag{21}$$

The inequalities $0 \le k \le w$, $0 \le k \le w-t-1$ and $0 \le w-k \le w+t+1$ (i.e., $-t-1 \le k \le w$) hold. Because $w \le w-t-1$ and $0 \le -t-1$ for $-(w+1) \le t \le -1$ and $w-t-1 < w$ and $-t-1 < 0$ for $0 \le t \le w$, the next inequalities hold,

$$\begin{cases} -t-1 \le k \le w & (-(w+1) \le t \le -1), \\ 0 \le k \le w-t-1 & (0 \le t \le w). \end{cases} \tag{22}$$

Overall, from Eqs. (19), (20), (21) and (22), the probability of the two events $\{y_{\mathrm{M}}(t) \in S_1\}$ and $\{x < y_{\mathrm{M}}(t) \le x+\Delta x\}$ occurring simultaneously is expressed as

$$\begin{aligned} Pr^{(1)}(x,t|\Delta x) &= Pr[y_{\mathrm{M}}(t) \in S_1 \cap x < y_{\mathrm{M}}(t) \le x+\Delta x] \\ &= \begin{cases} (w-t)g_1(x)(\Delta x + O(\Delta x^2)\sum_{k=-t-1}^{w}\binom{w-t-1}{k}\binom{w+t+1}{w-k} \\ \qquad \cdot G_1^k(x)(1-G_1(x))^{w-t-1-k}G_2^{w-k}(x)(1-G_2(x))^{k+t+1} & (-(w+1)\le t \le -1), \\ (w-t)g_1(x)(\Delta x + O(\Delta x^2)\sum_{k=0}^{w-t-1}\binom{w-t-1}{k}\binom{w+t+1}{w-k} \\ \qquad \cdot G_1^k(x)(1-G_1(x))^{w-t-1-k}G_2^{w-k}(x)(1-G_2(x))^{k+t+1} & (0 \le t \le w). \end{cases} \end{aligned}$$

Then, the pdf of $y_{\mathrm{M}}(t)$ corresponding to $Pr^{(1)}[x,t|\Delta x]$ is obtained as

$$\begin{aligned} f_{\mathrm{M}}^{(1)}(x,t) &= \lim_{\Delta x \to 0} \frac{Pr^{(1)}[x,t|\Delta x]}{\Delta x} \\ &= \begin{cases} (w-t)g_1(x)\sum_{k=-t-1}^{w}\binom{w-t-1}{k}\binom{w+t+1}{w-k} \\ \quad \cdot G_1^k(x)(1-G_1(x))^{w-t-1-k}G_2^{w-k}(x)(1-G_2(x))^{k+t+1} & (-(w+1)\le t \le -1), \\ (w-t)g_1(x)\sum_{k=0}^{w-t-1}\binom{w-t-1}{k}\binom{w+t+1}{w-k} \\ \quad \cdot G_1^k(x)(1-G_1(x))^{w-t-1-k}G_2^{w-k}(x)(1-G_2(x))^{k+t+1} & (0 \le t \le w). \end{cases} \end{aligned} \tag{7}$$

2. Assume $y_{\mathrm{M}}(t) \in S_2$. The k order statistics in S_1 less than or equal to $y_{\mathrm{M}}(t)$ can occur in $\binom{w-t}{k}$ different ways, and furthermore, the $w-k$ order statistics in S_2 can occur in $\binom{w+t}{w-k}$ different ways. Then, the probabilities of the event that these k samples in S_1 are less than or equal to x and the event that the $w-k$ ones in S_2 are less than or equal to x are respectively given by

$$\binom{w-t}{k} G_1^k(x)(1-G_1(x))^{w-t-k} \tag{23}$$

and

$$\binom{w+t}{w-k} G_2^{w-k}(x)(1-G_2(x))^{k+t}. \tag{24}$$

The inequalities $0 \le k \le w$, $0 \le k \le w-t$ and $0 \le w-k \le w+t$ (i.e., $-t \le k \le w$) hold. Because $w < w-t$

and $0 < -t$ for $-(w+1) \le t \le -1$ and $0 \le w - t \le w$ and $-t \le 0$ for $0 \le t \le w$, the next inequalities hold,

$$\begin{cases} -t \le k \le w & (-(w+1) \le t \le -1), \\ 0 \le k \le w - t & (0 \le t \le w). \end{cases} \tag{25}$$

Overall, from Eqs. (19), (23), (24) and (25), the probability of the two events $\{y_{\mathrm{M}}(t) \in S_2\}$ and $\{x < y_{\mathrm{M}}(t) \le x + \Delta x\}$ occurring simultaneously is expressed as

$$\begin{aligned} Pr^{(2)}(x,t|\Delta x) &= Pr[y_{\mathrm{M}}(t) \in S_2 \cap x < y_{\mathrm{M}}(t) \le x + \Delta x] \\ &= \begin{cases} (w+t+1) g_2(x) (\Delta x + O(\Delta x^2) \sum_{k=-t}^{w} \binom{w-t}{k} \binom{w+t}{w-k} & \\ \quad \cdot G_1^k(x)(1-G_1(x))^{w-t-k} G_2^{w-k}(x)(1-G_2(x))^{k+t} & (-(w+1) \le t \le -1), \\ (w+t+1) g_2(x) (\Delta x + O(\Delta x^2) \sum_{k=0}^{w-t} \binom{w-t}{k} \binom{w+t}{w-k} & \\ \quad \cdot G_1^k(x)(1-G_1(x))^{w-t-k} G_2^{w-k}(x)(1-G_2(x))^{k+t} & (0 \le t \le w). \end{cases} \end{aligned}$$

Then, the pdf of $y_{\mathrm{M}}(t)$ corresponding to $Pr^{(2)}[x,t|\Delta x]$ is obtained as

$$\begin{aligned} f_{\mathrm{M}}^{(2)}(x,t) &= \lim_{\Delta x \to 0} \frac{Pr^{(2)}[x,t|\Delta x]}{\Delta x}. \\ &= \begin{cases} (w+t+1) g_2(x) \sum_{k=-t}^{w} \binom{w-t}{k} \binom{w+t}{w-k} & \\ \quad \cdot G_1^k(x)(1-G_1(x))^{w-t-k} G_2^{w-k}(x)(1-G_2(x))^{k+t} & (-(w+1) \le t \le -1), \\ (w+t+1) g_2(x) \sum_{k=0}^{w-t} \binom{w-t}{k} \binom{w+t}{w-k} & \\ \quad \cdot G_1^k(x)(1-G_1(x))^{w-t-k} G_2^{w-k}(x)(1-G_2(x))^{k+t} & (0 \le t \le w). \end{cases} \end{aligned} \tag{8}$$

From 1) and 2), we obtain the pdf of $y_{\mathrm{M}}(t)$ as

$$f_{\mathrm{M}}(x,t) = f_{\mathrm{M}}^{(1)}(x,t) + f_{\mathrm{M}}^{(2)}(x,t) \qquad (-(w+1) \le t \le w). \tag{26}$$

Furthermore, from (7), (8) and (18), we can easily prove the relation

$$f_{\mathrm{M}}(x,t) = \begin{cases} f_{\mathrm{M}}^{(1)}(x,t')_{t'=-(w+1)} & (-N \le t \le -(w+1)), \\ f_{\mathrm{M}}^{(2)}(x,t')_{t'=w} & (w \le t \le N-1). \end{cases} \tag{27}$$

Therefore, (6) is derived.

References

(1) Tukey, J. W. (1974) Nonlinear (nonsuperimposable) methods for smoothing data. Cong. Rec., EASCON,74, p.673,

(2) Bottema, M. J. (1991) Deterministic properties of analog median filters. IEEE Trans. on Information Theory .37(6), 1629-1640.

(3) Al-amri, S. S., Kalyankar, N., Khamitkar S. D. (2010) A comparative study of removal noise from remote sensing image. International Journal of Computer Science 7(1), 32-36.

(4) Barner, K. E., Arce, G. R. (1998) Order-Statistic Filtering and Smoothing of Time-Series: Part II, Northholland.

(5) Caselles, V., Spario, G., Chung, D. H. (2000) Vector median filters, inf-sup operations, and coupled PDE's: Theoretical connections. Journal of Mathematical Imaging and Vision 8, 109-119.

(6) Arias-Castro, E., Donoho, D. L. (2009) Does median filtering truly preserve edges better than linear filtering? The Annals of Statistics 37(3), 1172-1206

(7) Justusson, B. I. (1981) Median Filtering; Statistical Properties, Springer.

(8) Arnold, B. C., Balakrishnan, N., Nagaraja, H. N. (1992) A first course in order statistics. John Wiley & Sons.

和文抄録

メディアンフィルタと線形フィルタ、平滑化フィルタとしてどちらがよいか？

吉川　昭、小濱　剛、吉田　久

メディアンフィルタは、それが提案されて以来、決定論的信号に対するエッジ保存能力が注目され、エッジ保存が重視される状況下での平滑化フィルタとして、線形フィルタに勝るものとして頻繁に用いられるようになってきている。しかし、ガウス雑音の仮定のもとでは、メディアンフィルタのエッジ保存能力が線形フィルタを凌駕するのは高 SN 比入力に対してのみであり、SN 比が低下するに従いほとんど両者のエッジ保存能に差はなくなるとの報告がある。また、同じくガウス雑音に対して、高 SN 比でない限り、出力の平均二乗誤差で見る性能評価では、むしろ線形フィルタの方が優れているとの報告もある。これらは両フィルタを比較するための重要な情報ではあるが、データ処理の現場で実際にどちらを使うべきかの判断材料とするには十分でない。そこで本報告では、各種雑音に埋もれたエッジ信号入力に対する、メディアンフィルタと線形フィルタ（移動平均フィルタ）の出力のバイアスー分散特性を比較し、長い尾を持たない雑音（ガウス分布雑音と一様分布雑音）でかつ高 SN 比でない場合はむしろ線形フィルタ性能の方が良く、長い尾を持つラプラス分布雑音の場合は SN 比に関わらずメディアンフィルタが良いことを示す。これは一般に流布しているメディアンフィルタの線形フィルタに対する優位性「神話」への警告であり、また、実際にどちらのフィルタを使うべきかの判断材料の提供でもある。

キーワード：メディアンフィルタ、線形フィルタ、平滑化フィルタ、エッジ、ガウス分布、一様分布、ラプラス分布、バイアスー分散特性

近畿大学生物理工学部 システム生命科学科　〒649-6493 和歌山県紀の川市西三谷 930