M11.12

3-DIMENSIONAL MEDIAN FILTERS FOR IMAGE SEQUENCE PROCESSING

M. Bilge Alp†, *Yrjö Neuvo*‡

†Dept. of Elec. Eng., Bilkent Univ., TR-06533 Bilkent, Ankara, TURKEY.
‡Signal Proc. Lab., Tampere Univ. of Tech., P. O. Box 527, SF-33101, Tampere, FINLAND.

Abstract

Restoration of noisy images has been an active area of research for many years. Most of the research and development carried in this area have been on the restoration of 2-D images. Processing of image sequences brings a third dimension to the problem. In this paper, we develop two 3-D median-based filtering algorithms that preserve the motion in the image sequence while attenuating noise effectively. Some observations are made on the root signals in binary domain based on the positive Boolean functions corresponding to the filters. From the Boolean expressions the output distribution functions are derived. The performances of both filters under various noise types are examined theoretically and experimentally. The structures are simulated on a video sequencer (DVSR 100) on real image sequences. Comparisons are made with other 2- and 3-D algorithms from the literature based on mean square error, mean absolute error, and subjective criteria.

1. INTRODUCTION

Many applications in image processing require the processing of 3-D signals, namely image sequences. TV applications, target tracking, robot navigation, dynamic monitoring of industrial processes, study of cell motion by microcinematography, highway traffic monitoring, and video transmission are only a few examples where the signal to be processed is 3-D, the third dimension being time. It has been shown in many cases that 1-D algorithms do not produce optimum results in image processing. In other words, while processing images, their 2-D nature has to be taken into account. Likewise, in processing image sequences, 1- or 2-D algorithms do not yield optimum results. Although similar in some respects, the extension of 2-D algorithms to 3-D signal processing is not straightforward. The motion content of the image sequence requires the time dimension to be approached in a different manner.

Temporal filters have been developed to make use of the information in the time dimension in many image processing problems [1]. However, temporal filters usually blur the moving parts of the image sequence, resulting in poor image quality. It is known that 2-D spatial processing gives better results in moving parts, whereas temporal processing gives better results in still parts of the image sequence. This observation leads to the development of adaptive algorithms that require motion estimation or motion compensation to obtain acceptable image quality [2]. However, motion estimation and motion compensation are critical processes which increase the complexity and the cost of processing. Therefore, it is highly desirable to have 3-D filters which would be insensitive to motion in image sequences.

In image processing, median filters preserve edges and high frequency details in the image, resulting in improved image quality [3]. Here, we present two 3-D filtering algorithms, that are insensitive to the motion content of the image sequence, based on the multilevel median operation introduced in [4]. The filtering structures are defined in Section 2. Section 3 gives some observations on the root signals in binary domain based on the Boolean expressions corresponding to the filters. The observations made in binary domain can be extended to multi-valued signals using the threshold decomposition property [5]. The output distributions of the filters are presented in Section 4 for a given input noise distribution. The filters are applied to real image sequences and the results are presented in Section 5. Finally, Section 6 gives the conclusions.

2. FILTERING STRUCTURES

In this section, two new median-based 3-D filtering structures will be introduced. There are very few examples in the literature of this kind [6]. In developing these filtering structures, the detail preserving property of the median operation has been made use of.

2.1 3-D Planar Filter (P3D)

The first 3-D algorithm is based on the multilevel median structure introduced in [4]. The structure is shown in Figure 2.1. It consists of four standard median filters. Each of the 5-point median operations in the first level operate on a different plane of the 3-D image sequence, i.e., on the x-y, x-t, and y-t planes. This is the reason why the filter is called the 3-D planar filter. Using the notation given in Figure 2.2 P3D can be defined as follows. First

CH2977-7/91/0000-2917 $1.00 © 1991 IEEE

level filters are

$$
\begin{aligned}
m_{xy} &= MED[D_1, E_1, F_1, B_1, H_1] , \\
m_{xt} &= MED[D_1, E_1, F_1, E_0, E_2] , \\
m_{yt} &= MED[B_1, E_1, H_1, E_0, E_2] ,
\end{aligned}
\tag{2.1}
$$

and the final output is

$$
y = MED[m_{xy}, m_{xt}, myt] .
$$

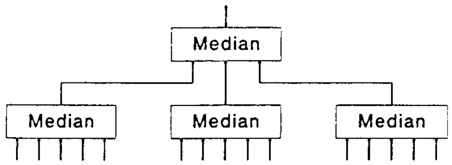

Figure 2.1 The multilevel structure for the 3-D planar filter (P3D).

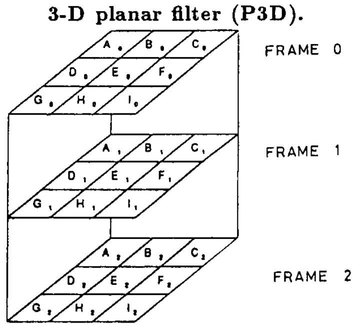

Figure 2.2 $3 \times 3 \times 3$ **cubic mask representing 3 successive frames.**

2.2 3-D Multilevel Filter (ML3D)

The second filter developed is based on the preservation of different features in the first level of the multilevel structure. The first level consists of two 7-point median filters each preserving different features of the input image. The multilevel structure is shown in Figure 2.3. Using the notation given in Figure 2.2, ML3D can be defined as follows. The first level 7-point filters are

$$
\begin{aligned}
m_{+} &= MED[D_1, E_1, F_1, B_1, H_1, E_0, E_2] , \\
m_{\times} &= MED[A_1, C_1, E_1, G_1, I_1, E_0, E_2] ,
\end{aligned}
\tag{2.2}
$$

and the final output is

$$
y = MED[m_{+}, m_{\times}, E_1] .
$$

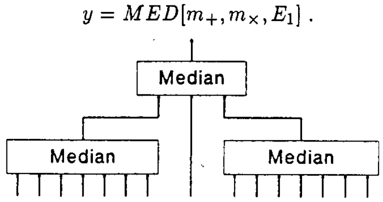

Figure 2.3 The multilevel structure for the 3-D multilevel filter (ML3D).

Usually the recursive versions of median-based filters have higher noise attenuations. It is possible to define the recursive versions of these filters, denoted by P3DR, and ML3DR as usual.

3. BINARY DOMAIN ANALYSIS

Threshold decomposition property of the median operation makes it possible to analyze the behaviour of these filters in binary domain. Since multilevel median filters belong to the class of stack filters, there is a positive Boolean function (PBF) corresponding to each of the filters defined in Section 2. Based on these PBF's, the following observations can be made on the root signals of the filters. In binary domain we can distinguish three cases according to the motion content of the image sequence.

Case 1 : $E_0 = E_1 = E_2$. This case represents static image sequences. In this case the outputs of the filters can be expressed as follows.

$$
E_0 = E_1 = E_2 \Rightarrow \begin{cases} f_{P3D}(\cdot) = E_1 ; \\ f_{ML3D}(\cdot) = A_1 B_1 C_1 D_1 F_1 G_1 H_1 I_1 \\ \qquad + E_1(A_1 + B_1 + C_1 + D_1 \\ \qquad + F_1 + G_1 + H_1 + I_1). \end{cases} \tag{3.1}
$$

The above expressions show that P3D preserves all high frequency details in a stationary sequence, i.e., all stationary image sequences are root signals of P3D. The 3-D multilevel filter, ML3D, still eliminates an impulsive point even if it repeats in successive frames. As will be seen in the simulations, ML3D has the highest noise attenuation, which is expected.

Case 2 : $E_0 = E_1 \neq E_2$ or $E_0 \neq E_1 = E_2$. In the second case, only two successive pixels out of three frames are equal. This may be considered as slow motion in a binary image sequence. In this case, the output of P3D reduces to the 5-point standard median filter $MED[B_1, D_1, E_1, F_1, H_1]$. Since the filter reduces to a 2-D algorithm, it is expected to preserve slow motion in the image sequence. On the other hand, the 3-D multilevel filter, ML3D, preserves the input pixel, E_1, only if at least two other pixels in one of the substructures corresponding to $+$ or $\times -$ shaped features are equal to the input pixel. This implies that the filter preserves all lines of arbitrary width under slow motion.

Case 3 : $E_0 \neq E_1 \neq E_2 (E_0 = E_2)$. In the case of fast motion, all successive pixels in three frames are different. In the binary domain this corresponds to oscillation in the time dimension. In this case, the following observation can be made on the output of P3D.

$$
f_{P3D}(\cdot) = E_1 \iff B_1 = E_1 = H_1 \text{ or } D_1 = E_1 = F_1 . \tag{3.2}
$$

This implies that, under fast motion, the filter preserves vertical and horizontal lines of arbitrary width, and diagonal lines that are at least two pixels wide. Under fast motion, the 3-D multilevel filter, ML3D, preserves the input pixel only if at least 3 other pixels in one of the substructures corresponding to + or $\times-$ shaped features are equal to the current pixel, E_1. This implies that the filter preserves all features at least two pixels wide under fast motion. This reduction in resolution is not critical since the eye does not require high spatial resolution under fast motion.

4. STATISTICAL ANALYSIS

By using the Boolean expressions corresponding to the filters, it is possible to express the output probability distribution functions in terms of the input distributions. These can be expressed as follows.

$$\begin{aligned} F_{P3D}(j) =& F(j)^3[3 + 20F(j) - 57F(j)^2 \\ & + 49F(j)^3 - 14F(j)^4] , \\ F_{ML3D}(j) =& F(j)^4\big[40 - 106F(j) + 84F(j)^2 \\ & + 60F(j)^3 - 195F(j)^4 + 190F(j)^5 \\ & - 88F(j)^6 + 16F(j)^7\big] , \end{aligned} \tag{4.1}$$

where $F(j)$ is the input distribution function. The output probability density functions (pdf) can be obtained by taking the derivative of these expressions. Thus, the noise attenuation of the filters can be obtained for the homogeneous parts of the image where the problem is to estimate a constant signal in additive white noise. For comparison purposes, the output distributions of two other 3-D median-based filters that are introduced by Arce *et al.* are also derived. These are the unidirectional (UNI3D) and bidirectional (BI3D) multistage filters that are defined in [6].

$$\begin{aligned} F_{UNI3D}(j) =& F(j)\big[1 - (1 - F(j))^{10}\big] \\ & + (1 - F(j))\, F(j)^{10} , \\ F_{BI3D}(j) =& F(j)^3\big[21 - 80F(j) + 166F(j)^2 \\ & - 224F(j)^3 + 202F(j)^4 + 120F(j)^5 \\ & + 45F(j)^6 - 11F(j)^7 + 2F(j)^8\big] . \end{aligned} \tag{4.2}$$

Although the analytical expressions are rather complicated, it is possible to make the statistical analysis by numerical methods. The probability density functions of the filters are plotted in Figures 4.1 and 4.2 for Gaussian and biexponential noise distributions.

5. SIMULATIONS

The filters that are defined in Section 2 are simulated on a VTE DVSR 100 [7] image sequencer. Noise attenuations of these filters and their recursive versions are calculated for both Gaussian and biexponential independent, identically distributed (i.i.d.) additive white noise using

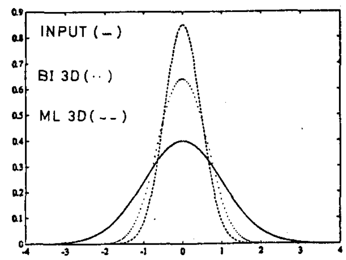

Figure 4.1 The pdf of the output of BI3D, and ML3D for zero mean, unit variance Gaussian noise.

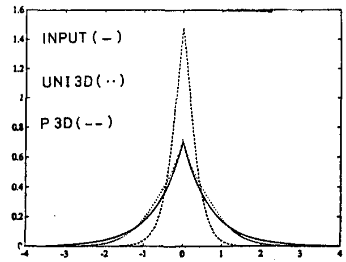

Figure 4.2 The pdf of the output of P3D, and UNI3D for zero mean, unit variance biexponential noise.

a 4 frame (256×128), zero mean, unit variance noise sequence. The same calculations are carried out for Arce's filters, the 2-D 5-point median filter, which is given as $y = MED[B_1, D_1, E_1, F_1, H_1]$, and the linear averaging filter in a 3×3 square window. The results are given in Table 5.1.

Table 5.1 Output variance of various filters when the input is zero mean, unit variance i.i.d. noise with Gaussian and biexponential distributions.

FILTER	GAUSS.	BIEXP.
P3D	0.238	0.137
P3DR	0.117	0.061
ML3D	0.222	0.124
ML3DR	0.119	0.059
UNI3D	0.735	0.579
UNI3DR	0.735	0.579
BI3D	0.363	0.231
BI3DR	0.298	0.181
MEDIAN5	0.293	0.178
MEDIAN5R	0.152	0.083
LAVE	0.113	0.113
LAVER	0.101	0.100

The same filters are also applied to still and moving image sequences with additive Gaussian, and impulsive noise. For impulsive noise, the probability of an impulse is 0.1 with equal probability for positive and negative impulses. For additive Gaussian noise the variance is 30 and the mean is zero. The still image sequence is a 4 frame sequence created using the image "BRIDGE". The motion sequence is a 19 frame sequence called "COSTGIRLS". The mean square error (MSE) and the mean absolute error (MAE) between the original sequence and the filter outputs are given for the "BRIDGE" sequence with impulsive and Gaussian noise distributions in Table 5.2.

Table 5.2 MSE and MAE between the original "BRIDGE" sequence and the outputs of the filters for various noise distributions. For impulsive noise, the probability of an impulse is 0.1 and for Gaussian noise, the variance is 30.

FILTER	IMP. MAE	IMP. MSE	GAUSS. MAE	GAUSS. MSE
P3D	0.989	37.641	12.544	251.794
P3DR	0.954	22.322	10.619	186.492
ML3D	1.242	34.314	12.368	244.233
ML3DR	1.457	29.047	10.872	194.157
UNI3D	6.347	833.632	21.045	666.403
UNI3DR	6.347	833.632	21.045	666.403
BI3D	1.301	106.394	14.909	347.280
BI3DR	1.011	58.936	14.028	308.323
MEDIAN5	5.720	154.661	14.881	359.088
MEDIAN5R	5.421	119.197	12.992	283.006
LAVE	15.184	422.844	12.440	270.612
LAVER	14.589	400.602	12.224	259.358

For subjective evaluation part of the original sequence and the filter outputs are shown in Figure 5.1.

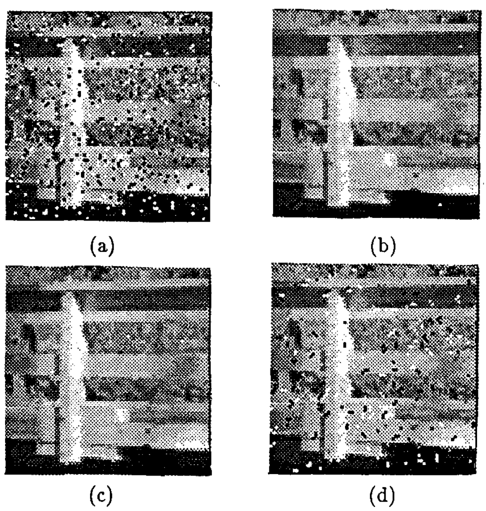

Figure 5.1 Part of the noisy "BRIDGE" sequence and the filter outputs for impulsive noise with probability 0.1. (a) Original image, (b) P3D, (c) ML3D, (d) UNI3D output.

6. CONCLUSIONS

As the simulations show, the two proposed 3-D algorithms (P3D and ML3D) have higher noise attenuation than Arce's unidirectional and bidirectional filters. Although the 2-D linear average filter seems to have the best noise attenuation, it is not preferable since it also filters the high frequency details in the image causing blurring. This is why it does not give good results when applied to real image sequences as can be seen from the results presented in Table 5.2. The simulations made on moving image sequences show that the 3-D filters presented here do not disturb the motion content of the image. The only disadvantage of 3-D filters compared to their 2-D counterparts is that they require more memory, two frames in our case. However, with current VLSI technology, the algorithms can already be implemented at video rates [8].

REFERENCES

[1] S. S. H. Naqvi, N. C. Gallagher, and E. J. Coyle, "An application of median filters to digital television," ICASSP 86, pp. 2451-2454, Tokyo, 1986.

[2] T. S. Huang, and Y. P. Hsu, "Image sequence enhancement", in *Image Sequence Analysis*, T. S. Huang, Ed., Springer-Verlag, New York, 1981.

[3] J. Salo, Y. Neuvo, and V. Hämeenaho, "Improving TV picture quality with linear-median type operations," IEEE Trans. Consumer Electronics, vol. 34, no. 3, pp. 373-379, Aug. 1988.

[4] A. Nieminen, P. Heinonen, and Y. Neuvo, "A new class of detail preserving filters for image processing," in Proc. 1989 IEEE Int. Conf. Circuits and Systems, Portland, OR, pp. 983-986, May 1989.

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

[6] G. R. Arce, and E. Malaret, "Motion-preserving ranked-order filters for image sequence processing," Proc. 1989 IEEE Int. Conf. Circuits and Systems, pp. 983-986, Portland, OR, May 1989.

[7] VTE Digitalvideo GMBH DVSR 100 Documentation.

[8] J. Juhola, P. Haavisto, O. Vainio, T. Raita-aho, and Y. Neuvo, "On VLSI implementation of median based field rate up-conversions," in Proc. 1990 IEEE Int. Symp. on Circuits and Systems, New Orleans, Louisiana, pp. 3042-3045, May 1990.