Noise Adaptive Soft-Switching Median Filter

How-Lung Eng, *Student Member, IEEE,* and Kai-Kuang Ma, *Senior Member, IEEE*

***Abstract*—Existing state-of-the-art switching-based median filters are commonly found to be nonadaptive to noise density variations and prone to misclassifying pixel characteristics at high noise density interference. This reveals the critical need of having a sophisticated switching scheme and an adaptive weighted median filter. In this paper, we propose a novel switching-based median filter with incorporation of fuzzy-set concept, called the *noise adaptive soft-switching median* (NASM) filter, to achieve much improved filtering performance in terms of effectiveness in removing impulse noise while preserving signal details and robustness in combating noise density variations. The proposed NASM filter consists of two stages. A soft-switching noise-detection scheme is developed to classify each pixel to be *uncorrupted pixel, isolated impulse noise, nonisolated impulse noise* or image object's *edge pixel*. "No filtering" (or *identity filter*), *standard median* (SM) filter or our developed *fuzzy weighted median* (FWM) filter will then be employed according to the respective characteristic type identified. Experimental results show that our NASM filter impressively outperforms other techniques by achieving fairly close performance to that of *ideal*-switching median filter across a wide range of noise densities, ranging from 10% to 70%.**

***Index Terms*—Adaptive median filter, fuzzy weighted median filter, image enhancement, impulse noise detector, median filter, nonlinear filter, switching-based median filter.**

I. Introduction

THE acquisition or transmission of digital images through sensors or communication channels is often interfered by impulse noise. It is imperative, and even indispensable, to remove these corrupted pixels to facilitate subsequent image processing operations, such as edge detection, image segmentation and object recognition, to name a few.

Impulse noise randomly and sparsely corrupts pixels to two intensity levels—relative high or relative low, when compared to its neighboring pixels. *standard median* (SM) filter (e.g., [1], [2]) was initially introduced to eliminate impulse noise and achieves reasonably well performance. SM filter exploits the rank-order information (i.e., order statistics) [3], [4] of the input data to effectively remove impulse noise by substituting the considered pixel with the middle-position element (i.e., median) of the re-ordered input data. Since its inception, SM filter has been intensively studied and extended to promising approaches such as *weighted median* (WM) [5] and *center weighted median* (CWM) [6] filters. The WM filter, proposed by Brownrigg in 1984, used a set of weighting parameters to control the filtering performance in order to preserve more signal details than what SM filtering can accomplish. CWM filter, proposed by Ko and Lee in 1991, is a special case of the WM filter, where only the center pixel of the filtering window has a weighting factor.

Intuitively, and ideally, the filtering should be applied to corrupted pixels only while leaving those uncorrupted ones intact. Applying median filter unconditionly across the entire image as practiced in the conventional schemes would inevitably alter the intensities and remove signal details of those uncorrupted pixels. Therefore, a noise-detection process to discriminate the uncorrupted pixels from the corrupted ones prior to applying nonlinear filtering is highly desirable. Sun and Neuvo [7] and Florencio and Schafer [8] have proposed their *switching-based* median filtering methodologies by applying "no filtering" to preserve true pixels and SM filter to remove impulse noise. However, we observed the following fundamental concerns inherited in these schemes.

First, the algorithms make use of a fixed noise-detection threshold obtained at a pre-assumed noise density level and hence lack of adaptivity to noise density variation. The mismatch between the designed algorithms and the actual noise density, which is often unknown *in priori*, will cause noticeable and even substantial degradation on filtering performance. Second, when the noise density increases, more misclassifications of pixel characteristic are going to occur and subsequently result in more degraded filtering performance. Therefore, an intelligent noise-detection process will be highly desirable and instrumental in correctly detecting various types of pixel characteristic. In addition, an adaptive filtering scheme is essential to effectively remove the corrupted pixels while preserving image details when misclassification of pixel characteristic happens. These indicate that both noise detection and corresponding filtering operation are crucial to achieve good median filtering performance, especially at high noise density interference.

In this paper, a novel *noise adaptive soft-switching median* (NASM) filter is proposed to address the above-mentioned concerns with architecture as shown in Fig. 1. It contains a switching mechanism steered by a soft-switching noise-detection scheme to identify each pixel's characteristic, followed by invoking proper filtering operation as outlined in Fig. 2. In the noise-detection scheme, global (i.e., based on the entire picture) or local (i.e., based on a small window) pixel statistics are utilized in the first and the remaining two decision-making levels, respectively. Most of the true pixel are successfully identified as "*uncorrupted pixels*" in the first decision-making level. Other remaining unidentified pixels will be further discriminated in the remaining two decision levels as "*isolated impulse noise*," "*nonisolated impulse noise*" or "*edge pixel*." The concept of fuzzy logic is exploited in the latter stage to achieve "*soft*" switching. In the filtering scheme, action of "no filtering" is applied to those identified uncorrupted pixels.

Manuscript received August 17, 1999; revised August 15, 2000. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Scott T. Acton.

The authors are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: ekkma@ntu.edu.sg).

Publisher Item Identifier S 1057-7149(01)00818-1.

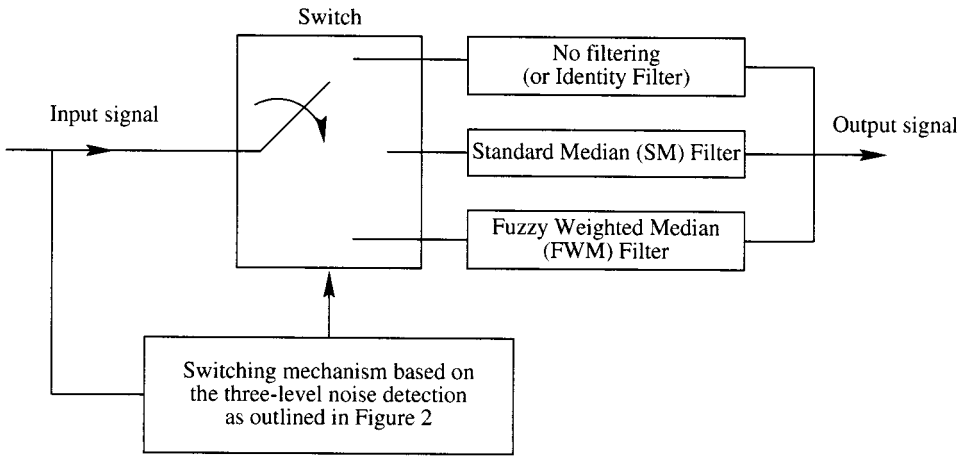

Fig. 1. *Noise adaptive soft-switching median* (NASM) filter.

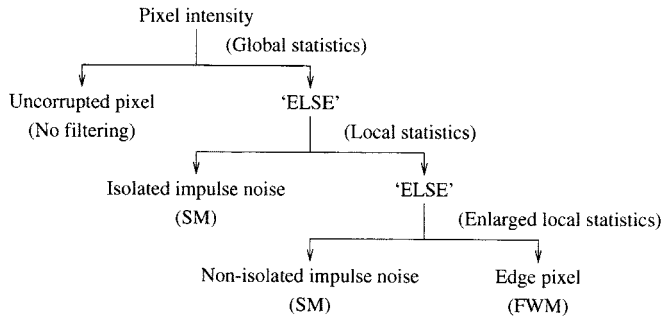

Fig. 2. Hierarchical soft-switching noise detection for identifying each pixel's characteristic.

SM or a proposed *fuzzy weighted median* (FWM) filtering would be subsequently carried out to remove impulse noise or preserve image object's edge details, depending on the pixel's characteristic identified. Instead of exploiting "no filtering" to the edge pixels, the proposed FWM is developed to effectively compensate possible degraded performance due to misclassification of *nonisolated impulse noise* as *edge pixel*. FWM is essentially an *adaptive* WM in which larger weights are assigned to more correlated pixels, based on the local statistics of pixel intensity. By doing so, FWM incorporates the pixel-intensity correlation to enhance its filtering capability in attenuating impulse noise while preserving image details.

This paper is outlined as follows. Section II describes our soft-switching noise-detection scheme used in identifying four different types of pixel characteristic. Section III discusses various types of median filters employed in response to the pixel characteristic type determined in the decision stage. These two sections establish the fundamental principles and structure of the proposed NASM filter. Sections IV and V present a summary of implementation procedures and test results, respectively. The conclusion is drawn in Section VI.

II. Soft-Switching Noise Detection

For each image pixel, a hierarchical soft-switching noise-detection process is performed to identify it as one of the four characteristic types: 1) *uncorrupted pixel*, 2) *isolated impulse noise*, 3) *nonisolated impulse noise*, and 4) *edge pixel*, as indicated at the decision tree nodes in Fig. 2, respectively. After identifying each pixel's characteristic, the corresponding filtering action will be recorded according to

$$g_{i,j} = \begin{cases} 1, & \text{"No filtering"} \\ 2, & \text{SM filtering} \\ 3, & \text{FWM filtering} \end{cases} \quad (1)$$

where indices i and j denote the coordinate of each pixel's position in the image. The *filtering action map* is thus formed for the entire image and to be referred in the filtering stage later. Initially, we set $g_{i,j} = 0$ for all i and j (i.e., the entire image).

A. Detection of Uncorrupted Pixel

The first-level noise detection involves the identification of "*uncorrupted*" pixel by utilizing the *global* statistics based on the pixel intensities of the entire image. Impulse noise corrupts the image pixel by altering its intensity to either relatively high or relatively low value. By analyzing the gray-level difference between the noisy image and an estimation of the original image pixel-wise, it is expected that uncorrupted pixels should yield much smaller differences as compared to that of corrupted ones. This intuition lays the core foundation on accurately identifying uncorrupted pixels as demonstrated in Section IV-A.

An estimation of the original image could be obtained by passing the noisy image through a SM filter with an adaptively determined window size of $W_{D1} \times W_{D1}$. SM filter is exploited to ensure the filtered image (i.e., estimator) is free from impulse noise. At low noise density level, small window size is desirable as it is capable of removing impulse noise without causing noticeable blurring effect. However, the use of smaller window might not be able to remove noise blotches at high noise density and hence cause more difficulties for the remaining noise-detection processes. On the contrary, large window size is more effective in removing impulse noise at high noise density situation but result in much serious blurring side effect. Based on the above-mentioned considerations and the fact that noise density level is usually unknown *in priori*, it is essential to make a good estimation of noise density p in order to determine the proper window size of $W_{D1} \times W_{D1}$. For that, the steps of the first-level noise detection will be iterated twice to estimate actual noise density p in the first iteration such that appropriate

TABLE I
SUGGESTED WINDOW SIZE FOR THE NOISE DENSITY LEVEL p ESTIMATED

Noise density	Suggested $W_{D1} \times W_{D1}$
$0\% < p \leq 15\%$	3×3
$15\% < p \leq 30\%$	5×5
$30\% < p \leq 45\%$	7×7
$45\% < p \leq 60\%$	9×9
$60\% < p \leq 70\%$	11×11

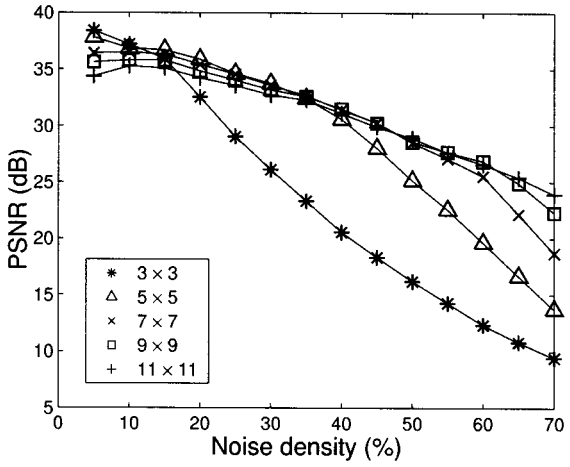

Fig. 3. Filtering performance of "Lena" under different $W_{D1} \times W_{D1}$ decision window sizes and noise densities.

window size $W_{D1} \times W_{D1}$ could be adaptively determined according to Table I. The same processing steps are then repeated in the second iteration using the window $W_{D1} \times W_{D1}$ determined in the first iteration.

To start with the first iteration, a fixed window size of 7×7 is applied and followed by calculating the number of pixels with $g_{i,j} = 0$ to estimate the noise density p. After analyzing four commonly used test images (i.e., "Lena," "Bridge," "Peppers," and "Sailboat") using various $W_{D1} \times W_{D1}$ decision windows for noise densities ranging from 10% to 70%, Table I is established based on the filtering performance achieved. A set of typical filtering performance curves resulted from exploiting different window sizes on 512×512 "Lena" test image is shown in Fig. 3. Analysis reveals that the use of smaller window size of $W_{D1} \times W_{D1}$ achieves superior filtering performance at lower range of noise densities; whereas, larger window size is more appropriate to be used at higher range of noise densities.

To improve the detection performance, the estimated original image is decomposed into nonoverlapping homogeneous rectangular blocks based on conventional quadtree decomposition technique. The concept of Weber–Fechner law [9] was adopted as the criterion for block splitting with the following modifications. According to the Weber–Fechner law, given two visual stimuli from two adjacent regions with intensities I and $I + \Delta I$ respectively, if the ratio of $\Delta I/I$ is less than the *just noticeable difference* threshold, these two regions would be visually indiscernible. In our case, four equally divided quadrants of each considered block represent four adjacent visual stimuli and the iterative decomposing process involves:

1) Starting with the entire image itself, the average pixel intensities of all the four equally divided rectangular blocks are computed and collectively denoted by the set $\mathbf{I} = \{I_1, I_2, I_3, I_4\}$. If any one of the absolute differences among the elements of set $\mathbf{I}$, i.e., $|I_1 - I_2|, |I_1 - I_3|, \cdots, |I_3 - I_4|$, is greater than the empirically determined threshold $T = 255/32$, the considered image block would be claimed as an inhomogeneous block, and further splitting is necessary.
2) Repeat step 1) on each divided block independently and recursively until all the sub-blocks are either homogeneous or reaching to the minimum size of 8×8.

Note that threshold T does not need to be highly accurate as long as it performs "reasonably" well in decomposing the whole image into homogeneous rectangular blocks. The fixed threshold suggested above provides a fairly robust decomposition for the four test images we experimented, and one example on decomposing "Lena" image is presented in Fig. 4(a).

For each homogeneous block obtained from the quadtree decomposition, the corresponding pixel-wise difference Δ_i between the noisy image and the estimated original image are computed for each block independently and viewed as the local estimation errors. A typical histogram distribution of Δ_i from the block highlighted with white-boundary bounding box in Fig. 4(a) is presented in Fig. 4(b). Distribution around the center is mainly contributed by those *uncorrupted* pixels as they tend to yield much smaller Δ_i values individually. Distributions appeared at both tails are contributed by those *corrupted* pixels and/or *edge* pixels.

Two optimal partition parameters, p_l and p_u, are derived as the two boundary positions of the center range such that all the pixels with Δ_i falling on this range are considered as being "uncorrupted," and their corresponding $g_{i,j}$ will be set to "1" according to (2). As the histogram of Δ_i is not symmetrical with respect to the origin, we consider positive and negative Δ_i separately to obtain p_u and p_l, respectively, as follows.

Denote $x_0, \cdots, x_m$ as the bin indices of the error histogram of Δ_i, and $0 \leq x_0 \leq x_1 \leq \cdots \leq x_m$. Each n_i (for $i = 0, \cdots, m$) indicates the number of elements falling on the bin i. Parameter p_u can be obtained by minimizing the following expression:

$$J_u = \sum_{i=0}^{m/2-1} n_i \left(x_i - \frac{x_0 + p_u}{2} \right)^2 + \sum_{i=m/2}^{m} n_i \left(x_i - \frac{x_m + p_u}{2} \right)^2. \quad (2)$$

Differentiating function J_u with respect to parameter p_u and setting it to zero, we obtain

$$p_u = \frac{2}{\sum_{i=0}^{m} n_i} \left[\sum_{i=0}^{m/2-1} n_i \left(x_i - \frac{x_0}{2} \right) + \sum_{i=m/2}^{m} n_i \left(x_i - \frac{x_m}{2} \right) \right]. \quad (3)$$

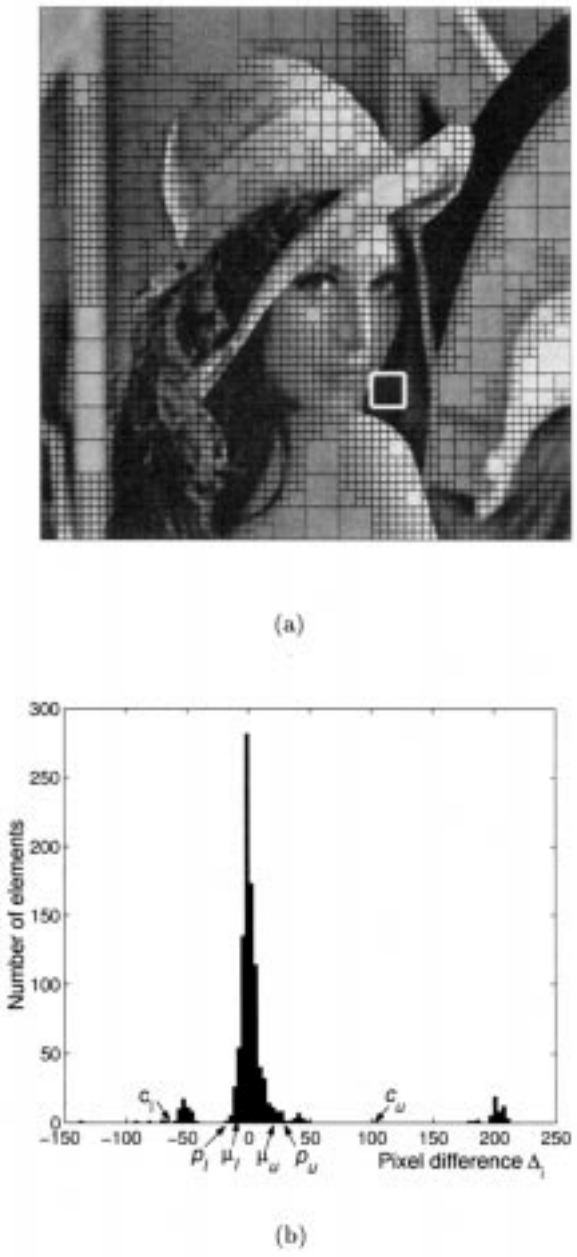

Fig. 4. (a) Example of quadtree decomposition on "Lena" image with threshold $T = 255/32$. (b) A histogram of pixel difference—between noisy image (with noise density $p = 10\%$) and estimated original image from a homogeneous block highlighted by a white bounding box as shown in (a).

Similar analysis is repeated for the negative part of the distribution. Let bin indices $x_{-m} \le x_{-m+1} \le \cdots \le x_{-1} < 0$, and n_i represents the number of elements in bin i. Parameter p_l can be obtained by minimizing the following function J_l

$$J_l = \sum_{i=-m}^{-m/2-1} n_i \left(x_i - \frac{x_{-m} + p_l}{2} \right)^2 + \sum_{i=-m/2}^{-1} n_i \left(x_i - \frac{x_{-1} + p_l}{2} \right)^2. \quad (4)$$

With the same mathematical approach, we obtain

$$p_l = \frac{2}{\sum_{i=-m}^{-1} n_i} \left[\sum_{i=-m}^{-m/2-1} n_i \left(x_i - \frac{x_{-m}}{2} \right) + \sum_{i=-m/2}^{-1} n_i \left(x_i - \frac{x_{-1}}{2} \right) \right]. \quad (5)$$

Since the second-order derivative yields

$$\frac{\partial^2 J_u}{\partial p_u^2} = \sum_{i=0}^{m} \frac{n_i}{2} > 0 \quad \text{and}$$
$$\frac{\partial^2 J_l}{\partial p_l^2} = \sum_{i=-m}^{-1} \frac{n_i}{2} > 0 \quad (6)$$

this ensures that the computed p_u and p_l correspond to the minimum values of functions J_u and J_l, respectively.

Fig. 4(b) shows that our approach in determining parameters p_u and p_l avoids the biasing toward to the origin (i.e., $\Delta_0 = 0$) as compared to the approach by taking the mean of $\Delta_i \ge 0$ and $\Delta_i < 0$, individually (denoted by μ_u and μ_l, respectively). Furthermore, this approach also avoids the biasing toward to both extreme ends of the distribution as compared to parameter c_u and c_l, respectively, where $c_u = (x_m + x_0)/2$ and $c_l = (x_{-m} + x_0/2)$.

For those pixels falling on the range of $[p_l \ p_u]$ are considered as uncorrupted pixels and the corresponding $g_{i,j}$ is set to "1," indicating that "no filtering" action will be taken. The remaining pixels with $g_{i,j} = 0$ require to be further processed in the second-level noise detection as described in the following.

B. Detection of Isolated Impulse Noise

The second-level noise detection involves the identification of "*isolated*" impulse noise by utilizing local statistics of pixel intensities extracted from $W_{D2} \times W_{D2}$ decision window, where W_{D2} is an odd integer and satisfies $3 \le W_{D2} \le W_{D1}$. Fig. 5 demonstrates a simplified one-dimensional representation for two cases where the considered pixel is 1) an *isolated impulse noise* or 2) part of a *correlated pixel block* in an N-pixel sliding window. The former case demonstrates that an isolated impulse noise possesses intensity which is relatively higher or lower than that of its neighboring samples; whereas, the latter may be a small noise blotch or an edge pixel of an image object. For this level of noise detection, we incorporate fuzzy-set concept as follows.

Given a pixel as the center pixel, the membership values of its neighboring pixels s are defined as

$$\mu_s = \left(\sum_i \frac{d_s}{d_i} \right)^{-1} \quad (7)$$

for $-(N-1)/2 \le i, s \le (N-1)/2$, $i \ne 0$ and $s \ne 0$. Parameters d_i and d_s are the intensity differences between the pixels i and s with respect to the center pixel in the N-pixel sliding

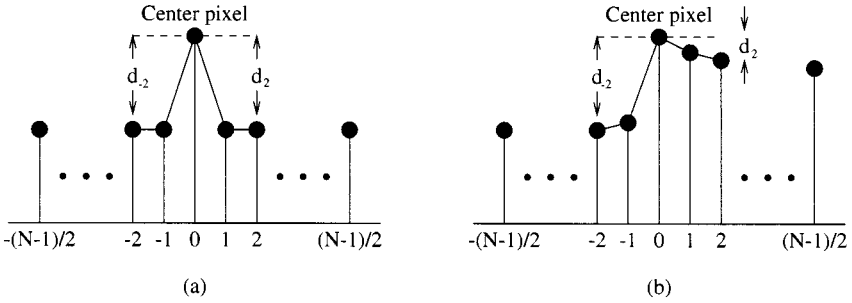

Fig. 5. One-dimensional illustration of (a) isolated impulse noise and (b) small block of correlated pixels.

window, respectively. The computing of membership values essentially transforms the "*pixel-intensity* map" into the "*membership-value* map." Denote μ_{left} and μ_{right} as the means of the membership values of the considered pixels from each side. Equivalently, μ_{left} and μ_{right} indicate the degree of "confidence" of the center pixel belonging to their respective sides.

Note that, if the center pixel's value is of equal intensity differences from both sides, this will yield $\mu_{left} = \mu_{right} = 0.50$. Hence, based on the same binarization principle of the *absolute moment block truncation coding* (AMBTC) [10], [11], the confidence threshold of the membership value is set to be $0.75 (= (0.50 + 1.0)/2)$. That is, the center pixel in Fig. 5 will be assigned to the left side if $\mu_{left} \geq 0.75$ (thus, $\mu_{right} \leq 0.25$ and $\mu_{left}/\mu_{right} \geq 3$). On the contrary, the center pixel will be more associate to the right region if $\mu_{right} \geq 0.75$ (thus, $\mu_{left} \leq 0.25$ and $\mu_{left}/\mu_{right} \leq 1/3$). Therefore, the decision rule for detecting an *isolated impulse noise* in one-dimensional case can be concluded as follows:

1) One-dimensional case:

Condition 1 If $1/3 < \mu_{left}/\mu_{right} < 3$, an isolated impulse noise is detected, and $g_{i,j}$ is set to "2" according to (2). SM filtering will be carried out to remove this impulse noise later.

Condition 2 If $\mu_{left}/\mu_{right} \leq 1/3$ or $\mu_{left}/\mu_{right} \geq 3$, the pixel will be considered as part of a small correlated pixel block. In this case, the considered pixel is either a *nonisolated impulse noise* or an *edge pixel* according to Fig. 2.

The aforementioned 1-D approach can be straightforwardly extended into two-dimensional (2-D) case with analysis window $W_{D2} \times W_{D2}$. Only those *uncorrupted* pixels (i.e., $g_{s,t} = 1$) within the 2-D sliding window are considered on computing their membership values associated to the center pixel. The membership value of uncorrupted pixel at coordinate (s,t) within $W_{D2} \times W_{D2}$ is defined as

$$\mu_{s,t} = \left(\sum_i \sum_j \frac{d_{s,t}}{d_{i,j}} \right)^{-1} \tag{8}$$

for $-(W_{D2}-1)/2 \leq s,t \leq (W_{D2}-1)/2$ and $(s,t) \neq (0,0)$. Coordinate (i,j) corresponds to all $g_{i,j} = 1$ within the window. Parameters $d_{i,j}$ and $d_{s,t}$ are the intensity differences between pixels $(i,j)$and (s,t) with respect to the center pixel, respectively. Starting with $W_{D2} = 3$, the decision window iteratively extends outwards by one pixel in all the four window sides provided that the number of uncorrupted pixels are less than $W_{D2} \times W_{D2}/2$, or until $W_{D2} = W_{D1}$.

By adopting the same binarization method used in AMBTC, the mean of $\mu_{s,t}$ is used to divide the membership map $\mu_{s,t}$ into two groups—higher-value group representing "closely correlated pixels" and lower-value group indicating "noncorrelated pixels." The means of each group's membership values are computed individually and denoted as μ_{high} and μ_{low}, respectively. Therefore, the decision rule for detecting an *isolated impulse noise* in the 2-D case can be concluded as follows.

2) Two-dimensional case:

Condition 1 If $1/3 < \mu_{low}/\mu_{high} < 3$, the center pixel is claimed to be comparatively far away from both groups. Therefore, it would be recognized as an isolated impulse noise, and the corresponding $g_{i,j}$ is set to "2."

Condition 2 If $\mu_{low}/\mu_{high} \leq 1/3$ or $\mu_{low}/\mu_{high} \geq 3$, further discrimination will be required as described in Section II-C.

C. Discrimination between Nonisolated Impulse Noise and Edge Pixel

The third-level noise detection distinguishes the considered pixel as being a *nonisolated impulse noise* or an *edge pixel*. Nonisolated impulse noise refers to the considered pixel that belongs to a noise blotch; whereas, edge pixel is simply a true pixel that falls on the edge of an image object. Note that these two categories are most difficult to be discerned from each other, since both are high frequency signals in essence. Fig. 6 illustrates an example and reveals the limitation of window $W_{D2} \times W_{D2}$ exploited in the previous detection level on "capturing" sufficient local statistics to distinguish these two categories. Although the pixel values covered within the $W_{D2} \times W_{D2}$ windows in Fig. 6(a) and (b) are identical, the considered pixel (being circled) is an impulse noise in Fig. 6(a) but an edge pixel in Fig. 6(b) after examining more surrounding pixels.

Intuitively, more reliable pixel statistics could be obtained by properly extending $W_{D2} \times W_{D2}$ window only to those directions that include more correlated pixels. If the pixel being considered is an edge pixel, such extension will include more correlated pixels from its surrounding, and subsequently increase the percentage of closely correlated pixels (i.e., enhancing local statistics reliability). On the other hand, if the considered pixel is a nonisolated impulse noise, the inclusion of more impulse noise will not be possible since only uncorrupted pixels (i.e.,

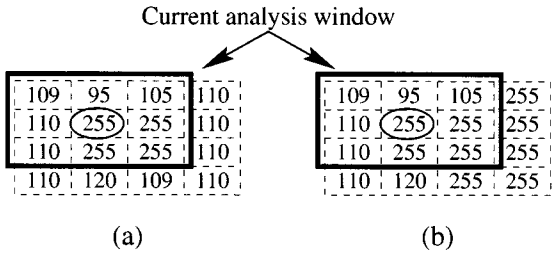

Fig. 6. Pixel under consideration (being circled) is an impulse noise in (a), but an edge pixel in (b), after observing more surrounding pixels.

pixel with $g_{i,j} = 1$) in the enlarged window are considered. The proposed algorithm individually checks each window boundary of the $W_{D2} \times W_{D2}$ decision window to examine whether it contains at least one "closely correlated pixel," which corresponds to those pixels exploited in computing parameter μ_{high} as described in Section II-B. If so, the corresponding window boundary will be subsequently extended outwards by one pixel in that side. The same analysis steps described in Section II-B are then repeated for the enlarged window.

Denote N_c as the number of "closely correlated pixel" identified within the enlarged $W_{D2} \times W_{D2}$ window. A threshold S_{in} is conservatively defined to be half of the total number of uncorrupted pixels within the enlarged window. If $N_c > S_{in}$, the considered pixel will be identified to be an *edge pixel*, and the corresponding $g_{s,t}$ is set to "3." Otherwise, it is recognized to be a *nonisolated impulse noise*, and $g_{s,t}$ is set to "2." Note that the window extension conducted in the second-level noise detection is for detecting isolated impulse noise. Thus, each window extension should be applied to all the four window boundaries. On the contrary, "closely correlated pixels" are of interest in the third-level noise detection. Therefore, only those window boundaries that contain potentially correlated uncorrupted pixel(s) are being extended.

Experiments have been carried out according to the aforementioned procedures by iteratively extending the $W_{D2} \times W_{D2}$ window and up to the maximum size of $W_{D2} = 64$ pixels for test images "Lena," "Bridge," "Peppers" and "Sailboat." However, the obtained filtering performance is comparable to that obtained by extending the $W_{D2} \times W_{D2}$ window only once. Thus, we only extend $W_{D2} \times W_{D2}$ window once in this paper.

Note that, in the third-level noise detection, if the considered pixel is actually uncorrupted and happens to be very close to a particular neighboring pixel in intensity, the resulting *membership value* of that particular pixel will be much larger than that of other uncorrupted pixels (within the window), owing to the exponentiality of the membership function and $\Sigma_s \, \Sigma_t \mu_{s,t} = 1$. Hence, this will exclude other uncorrupted pixels to be considered as "closely correlated pixels" and subsequently misinterpret the considered pixel to be a small noise blotch. To avoid such numerical peculiarity, each $d_{s,t}$ will be clipped at $d_{\min}$ if it were found that $d_{s,t} < d_{\min}$. Such clipping operation has effect of assigning *equal* membership value to majority of those uncorrupted pixels and successfully classify them to be "closely correlated pixel." In this paper, parameter $d_{\min}$ of a homogeneous block (obtained after quadtree decomposition as mentioned in Section II-A) is defined as half of the dynamic range between the maximum and minimum uncorrupted pixel intensities, denoted as $P_{\max}$ and $P_{\min}$ respectively, of that block; i.e., $d_{\min} = (P_{\max} - P_{\min})/2$. If $N_c > S_{in}$, the considered pixel will then be classified as an *edge pixel*.

III. Filtering Scheme

Besides the action of "no filtering" applied to those uncorrupted pixels identified (i.e., $g_{i,j} = 1$), SM and the proposed *fuzzy weighted median* (FWM) filters are exploited for the detected *impulse noise* and *edge pixels*, indicated by $g_{i,j} = 2$ and $g_{i,j} = 3$ in the filtering action map, respectively.

Since both *nonisolated impulse noise* and *edge pixels* are high-frequency signals in essence, they are most difficult to be discriminated from each other; thus, leading to higher probability of misclassification. Action of "no filtering" applied to the misclassified *edge pixels* will cause the unremoval of noise pixels. Reversely, applying SM filter to the misclassified noise pixels will, theoretically speaking, lead to a certain degree of smearing on the image object's edges. A *fuzzy weighted median* (FWM) filter is derived and shown in (15) to compensate the former case, as the human visual system is fairly sensitive to the presence of impulse noise. The proposed FWM filter adaptively assigns different weighting factors to all the uncorrupted pixels within the $W_F \times W_F$ filtering window. Greater weights are assigned to those closely correlated pixels and smaller weights to those less correlated ones. In this way, the pixel-intensity correlation is incorporated to enhance FWM's filtering capability in preserving edge pixels while removing impulse noise.

By exploiting SM filter with the window size of $W_F \times W_F$ to a noise pixel, the output pixel Y_{ij} is

$$Y_{ij} = \text{median} \, \{X_{i-s,\, j-t} | (s,t) \in W\} \tag{9}$$

where $W = \{(s,t) | -(W_F - 1/2) \leq s,t \leq (W_F - 1)/2\}$. Note that only *uncorrupted* pixels within the window W (i.e., $g_{i-s,j-t} = 1$) are considered for the ranking process. The filtering window $W_F \times W_F$ is obtained in the same way as that of the decision window $W_{D2} \times W_{D2}$ in the second-level noise detection; thus, $3 \leq W_F \leq W_{D1}$.

In the proposed FWM filter, the fuzzy membership values $\mu_{s,t}$ computed earlier using (8) are re-used to determine all the weights of uncorrupted pixels within the $W_F \times W_F$ window, except for the center pixel. Larger weights are assigned to more correlated pixels, and less weights to those otherwise. The weight of the center pixel $\omega_{0,0}$ is determined by minimizing the output data variance σ^2_{wm} as defined in [12, Eq. (3)] so that the noise attenuation will be maximized. The relevant foundation of [12] and the derivation of FWM filter are described in the following.

For i.i.d. inputs with cumulative distribution function $F(t)$ and density $f(t)$, the output of the WM window is asymptotically normal, and its output variance is

$$\sigma^2_{wm} = \frac{\sum_{i=1}^{N} \omega_i^2}{4f(t_o)} \tag{10}$$

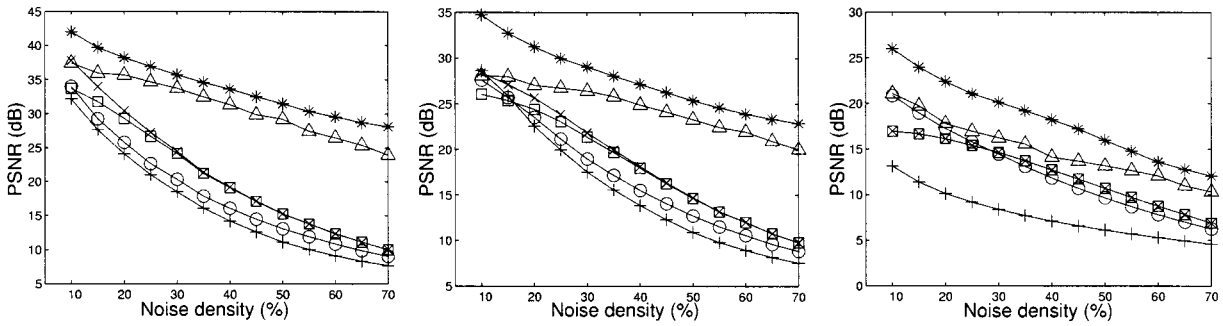

Fig. 7. Performance comparison using various median filtering techniques under various noise densities. (Legends: "□" for median filter; "∘" for center weighted median filter; "+" for Florencio and Schafer's switching scheme; "×" for Sun and Neuvo's switching scheme-I; "△" for noise adaptive soft-switching median filter; "*" for Ideal-switching filter). (a) Lena, (b) Bridge, and (c) Text.

where $t_o = F^{-1}(\frac{1}{2})$; i.e., t_o is obtained such that $F(t_o) = \frac{1}{2}$. Parameter N is the total number of uncorrupted pixels within the $W_F \times W_F$ window, and ω_i is the weight of the ith element of the WM filter. Minimizing (10) is equivalent to minimizing its numerator.

The weighting factors of the considered pixels within the $W_F \times W_F$ filtering window is defined as follows. For those $g_{s,t} = 1$ within the window $W = \{(s,t)| -(W_F-1)/2 \le s,t \le (W_F-1)/2\}$

$$\omega_{s,t} = \begin{cases} \dfrac{\mu_{s,t}}{X}, & \text{for } (s,t) \neq (0,0), \\ \dfrac{\mu_c}{X}, & \text{if } s = t = 0 \end{cases} \tag{11}$$

where $X = \Sigma\ \mu_{s,t} + \mu_c$, and μ_c/X is the weighting factor assigned to the center pixel. Parameter μ_c can be optimally derived as follows.

Define K as the summation of the square of all the weighting parameters, i.e.,

$$K = \sum_s \sum_t \left(\frac{\mu_{s,t}}{X}\right)^2 + \left(\frac{\mu_c}{X}\right)^2, \quad \text{for } (s,t) \neq (0,0) \text{ and } g_{s,t} = 1. \tag{12}$$

Taking the partial derivative of K with respect to μ_c and setting the result to zero, i.e.,

$$\frac{\partial K}{\partial \mu_c} = \sum_s \sum_t \frac{-2\mu_{s,t}^2}{X^3} + \frac{2\mu_c}{X^2} - \frac{2\mu_c^2}{X^3} = 0, \quad \text{for } (s,t) \neq (0,0) \text{ and } g_{s,t} = 1. \tag{13}$$

By further substituting $X = \Sigma_s\ \Sigma_t\ \mu_{s,t} + \mu_c$ into (13), we arrive at the expression of μ_c as

$$\mu_c = \frac{\sum_s \sum_t \mu_{s,t}^2}{\sum_s \sum_t \mu_{s,t}}, \quad \text{for } (s,t) \neq (0,0) \text{ and } g_{s,t} = 1. \tag{14}$$

Therefore, the filtered value of pixel $X_{i,j}$ is

$$Y_{i,j} = \text{median } \{\omega_{i-s,j-t} \diamondsuit X_{i-s,j-t} | (s,t) \in W\} \tag{15}$$

where symbol $\diamondsuit$ denotes the duplication operation.

IV. Experimental Results

A. Soft-Switching Noise-Detection Performance

Multiple commonly used gray-scale test images were experimented. Among them, "Lena," "Bridge," and "Text" are chosen and presented in Fig. 8. In our experiments, *salt-and-pepper* noise with uniform distribution were injected as practiced in [7] and [8]. That is, each image pixel has equal probability of being corrupted to either "white" (with value 255) or "black" (with value 0). Simulations were carried out for a wide range of noise density levels—$10\% \le p \le 70\%$ with an increment step of 5%.

To appreciate the performance contributed from each decision level in Fig. 2, parameters *correct detection* ξ and *misclassification* ζ for both corrupted and uncorrupted cases are defined as follows:

$$\text{Correct detection } \xi = \frac{\text{number of corrupted (uncorrupted) pixels detected}}{\text{total corrupted (uncorrupted) pixels in the image}} \tag{16}$$

and (17), as shown at the bottom of the page.

These parameters are used to measure the percentages of corrupted and uncorrupted pixels being correctly and incorrectly classified at each decision node, respectively. (See Table II for the results based on "Lena.") The percentages of ξ and ζ at each noise-detection level are possible to be calculated, as the exact position of injected impulse noise and the pixel characteristic identified for each pixel are known in the simulation. From Table II, it shows that the percentage of correct detection of *uncorrupted* pixels reaches almost 100%. This indicates that the first-level noise detection plays the dominant role in preserving image details. Majority of

$$\text{Misclassification } \zeta = \frac{\text{number of corrupted (uncorrupted) pixels misclassified}}{\text{total corrupted (uncorrupted) pixels in the image}}. \tag{17}$$

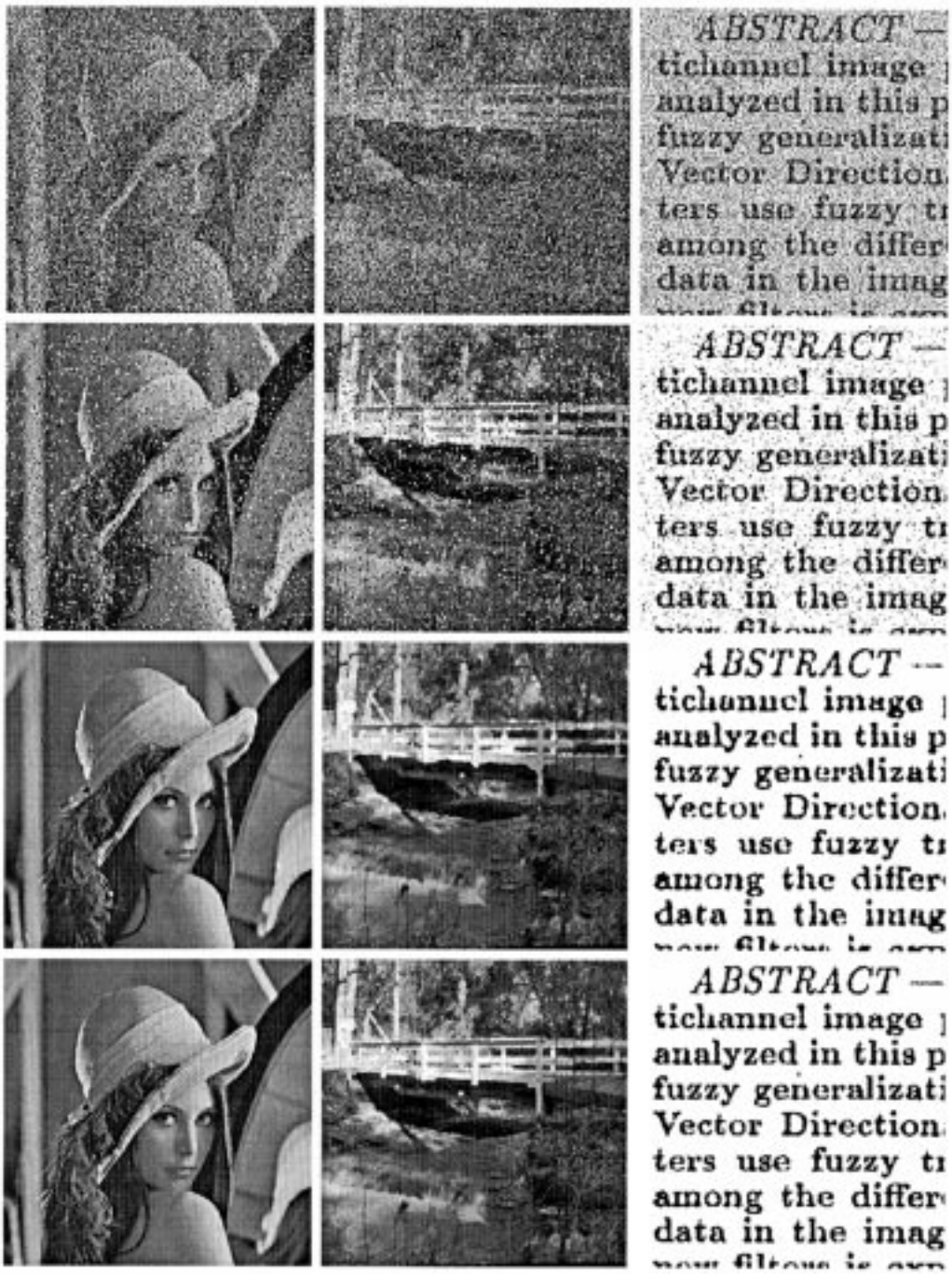

Fig. 8. Corrupted images "*Lena*", "*Bridge*," and "*Text*" with injected impulse noise density $p = 50\%$ each are shown in the first row. The corresponding filtered images resulted from exploiting SM, Sun and Neuvo's switching scheme-I, our proposed NASM and ideal-switching filters are shown in the second, third and fourth row, respectively.

the remaining unidentified uncorrupted pixels are successfully re-detected as *edge pixels* in the last decision level. For "Lena" image at the $p = 30\%$ in Table II for example, our first-level noise-detection scheme achieves 99.474% of correct detection. Among the remaining 0.526% unidentified uncorrupted pixels, 0.319% of the total true pixels have been successfully classified as edge pixels in the third-level noise detection. Hence, only 0.207% ($= 0.137\% + 0.070\%$) of uncorrupted pixels are misclassified as isolated or nonisolated impulse noise.

The classification of *impulse noise* also achieves superb performance with over 97% being correctly detected as *isolated* impulse noise for $p \leq 50\%$. This implies that the second-level noise-detection process provides the core mechanism in removing impulse noise, even when the noise density is high. For "Lena" image at $p = 30\%$ for example, our proposed noise detection scheme has correctly classified 99.444% isolated and 0.136% nonisolated impulse noise. Only 0.420% ($= 0.195\% + 0.225\%$) of impulse noise have been misclassified as uncorrupted pixels.

TABLE II
AN EXAMPLE OF *CORRECT DETECTION* ξ AND *MISCLASSIFICATION* ζ YIELDED BASED ON "LENA" IMAGE

Noise density	Uncorrupted pixel		Isolated impulse noise		Non-isolated impulse noise		Edge pixel	
	ξ (%)	ζ (%)	ξ (%)	ζ (%)	ξ (%)	ζ (%)	ξ (%)	ζ (%)
10%	97.678	0.076	99.585	1.004	0.175	0.085	1.233	0.164
30%	99.474	0.195	99.444	0.137	0.136	0.070	0.319	0.225
50%	99.492	0.575	96.832	0.069	1.705	0.077	0.362	0.888
70%	99.350	1.822	80.407	0.052	13.539	0.091	0.507	4.232

At high noise density level, impulse noise tends to form noise blotches rather than isolated ones. The third-level noise-detection re-detects the presence of noise blotches and classified them as *nonisolated* impulse noise. This is shown by the increase of parameter ξ in the category of nonisolated impulse noise when the noise density level increases. For example, for "Lena" image at high noise density $p = 70\%$, 93.946% ($= 80.407\% + 13.539\%$) of impulse noise have been correctly identified. Only 1.822% and 4.232% of impulse noise are misclassified as *uncorrupted* pixels and *edge* pixels, respectively. In this example, there are 17.771% ($= 100\% - 1.822\% - 80.407\%$) of impulse noise remain unidentified after the second-level noise detection. The third-level noise detection has successfully re-detected 76.186% ($= (13.539 \times 100/17.771)\%$) of remaining unidentified impulse noise. In conclusion, our fuzzy-set approach on conducting soft switching in the second-level and third-level noise detections is fairly effective in detecting impulse noise.

It has been further observed that the percentage of misclassifying uncorrupted pixels as nonisolated impulse noise increases for the images with high-activity content such as "Bridge." This trend is expected as high-activity images normally contain more intensity variations. Therefore, it is difficult to differentiate between true pixels and impulse noise; thus, leading to more misclassifications.

On the other way, it is also observed that more impulse noise tend to be misclassified as uncorrupted pixels in high-activity images as well. This is due to the facts that more impulse noise might have close pixel intensity to that of uncorrupted pixels in the same local area—leading to the misclassification as "uncorrupted." Also, high frequency areas with large intensity variations tend to camouflage small noise blotches—leading to the misclassification as "edge pixels."

B. Overall Filtering Performance

The performance of the filtering result is quantified by PSNR

$$\text{PSNR} = 10 \log_{10}\left(\frac{255^2}{\text{MSE}}\right) \text{ dB} \tag{18}$$

and

$$\text{MSE} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(Y_{i,j} - S_{i,j})^2 \tag{19}$$

where M and N are the total number of pixels in the horizontal and vertical dimensions of the image; $S_{i,j}$ and $Y_{i,j}$ are the original and filtered image pixels, respectively.

The PSNR performance of the proposed NASM filter was compared to that of the 3×3 SM filter, 3×3 CWM filter (with center weight $\omega_c = 3$) [6], Florencio and Schafer's switching scheme [8], Sun and Neuvo's switching scheme-I [7] and the ideal-switching filtering. For Sun and Neuvo's switching scheme-I, decision threshold of $T = 30$ obtained at noise density p = 10% [7] is used throughout the test. The *ideal*-switching filter is obtained by performing $W_F \times W_F$ SM filtering on those impulse noise only. This is possible to achieve on simulation, since the exact positions of injected impulse noise were recorded in the filter action map. The performance of the ideal-switching filter we suggested here is fairly useful to be served as the theoretical upper bound, in terms of PSNR, on gauging other switching-based median filters.

TABLE III
RUNTIME (IN SECONDS) CONSUMED AT VARIOUS NOISE DENSITIES p USING THE PROPOSED NASM FILTER AND OTHER MEDIAN FILTERS BASED ON "LENA" IMAGE

Filters	p=10%	p=30%	p=50%	p=70%
SM (3×3)	1.26	1.37	1.27	1.04
SM (7×7)	4.45	4.34	3.62	2.85
SM (11×11)	8.51	7.80	6.92	5.66
CWM (3×3)	1.82	1.70	1.65	1.43
CWM (7×7)	6.76	6.15	5.38	4.23
CWM (11×11)	12.80	11.97	10.49	8.40
Florencio and Schafer's switching scheme	1.05	1.04	0.94	0.88
Sun and Neuvo's switching scheme	1.59	1.42	1.43	1.27
NASM	8.41	9.45	12.96	17.92

The extrapolated PSNR curves resulted from using various median filters at different noise densities, ranging from 10% to 70%, are shown in Fig. 7. The proposed NASM filter significantly outperforms other median filtering schemes considered here and is much close to the ideal-switching filter in PSNR measurement. The consistency of the performance curves indicates that the proposed NASM is fairly robust against wide variation of impulse noise densities. A subjective visual comparison of the noise reduction and image detail-preserving using three test images "Lena," "Bridge," and "Text" are presented in Fig. 8. The proposed NASM filter achieves almost unnoticeable difference on subjective visual comparison as compared to that of the ideal-switching filter.

Among all the test images, "Text" image is the most difficult one to filter since character symbols create substantial amount of sharp edges. As expected, when the font size gets much smaller, the difficulty will be significantly increased and leading to much degraded performance.

Simulation testing also reveals that the size of the decision windows (i.e., $W_{D1} \times W_{D1}$ and $W_{D2} \times W_{D2}$) and filtering window (i.e., $W_F \times W_F$) of NASM is independent of the size of the test image.

C. Runtime Analysis

The runtime analysis of the proposed NASM filter and other concerned filters were conducted for "Lena" image using Pentium III 450 MHz Personal Computer and documented in Table III. Results reveal that NASM's total processing time is longer than others' in general. When noise density

increases, more noise blotches tend to occur. With the use of larger $W_{D1} \times W_{D1}$ and $W_{D2} \times W_{D2}$ decision windows and $W_F \times W_F$ filtering window, more processing time is needed for processing a larger amount of input data. On the other hand, majority of the pixels are detected as *uncorrupted* when noise density is low. This means that only the first level detection is involved, and no filtering is required. Consequently, the overall runtime needed is much shorter than that at high noise density.

In NASM filter, SM filtering and quadtree decomposition performed in the first-level noise detection are the most computationally intensive processing steps. For example, these two processes consume around 90% of the total runtime for $p \leq 30\%$ (e.g., 8.08 s out of the overall runtime of 8.41 s at $p = 10\%$). The filtering process occupies only about 10% of the total runtime, ranging from 2.02% at $p = 10\%$ to 18.42% at $p = 70\%$.

V. Conclusion

Our proposed NASM filter has simultaneously addressed the following issues commonly encountered in certain state-of-the-art *switching-based* median filters 1) nonadaptive to the changes of noise density 2) lack of sufficient sophistication in noise detection and adaptivity in median filtering, especially at high noise density.

The performance of our NASM filter has been extensively compared with that of SM, CWM, Florencio and Schafer's switching, and Sun and Neuvo's switching (scheme-I). Experimental results reveal that the proposed NASM filter significantly outperforms other techniques by having (much) higher PSNR with consistent and stable performance across a wide range of noise densities, varying from 10% to 70%. The *ideal*-switching filtering performance is also introduced in this paper to serve as the upper bound of the PSNR performance measurement. Note that the PSNR performance of the NASM filter is fairly close to that of the ideal-switching filter, and their subjective visual comparison is also hardly discernible from each other.

The proposed NASM filter is generic to be used in 1-D and multidimensional signals. Besides two-dimensional image processing concerned in this paper, we have also applied our NASM filter for smoothing out irregular macroblock motion vectors extracted directly from MPEG-encoded bitstreams for the application of video indexing and retrieval [13]. Compared with other median filters, superior performance on automatically identifying multiple video objects based on motion vectors has demonstrated that our proposed NASM strikes a good balance between preserving the details of motion-vector field while removing those irregular motion vectors which are considered as noise.

Acknowledgment

The authors would like to express their gratitude to the anonymous reviewers for their comments to improve the quality of this paper (particularly on runtime analysis and experimenting text images).

References

[1] I. Pitas and A. Venetsanopou, *Nonlinear Digital Filters: Principles and Application*. Norwell, MA: Kluwer, 1990.
[2] J. Astola and P. Kuosmanen, *Fundamentals of Nonlinear Digital Filtering*. Boca Raton, FL: CRC, 1997.
[3] H. A. David, *Order Statistics*. New York: Wiley, 1980.
[4] E. L. Lehmann, *Theory of Point Estimation*. New York: Wiley, 1983.
[5] D. Brownrigg, "The weighted median filter," *Commun. Assoc. Computer*, pp. 807–818, Mar. 1984.
[6] S.-J. Ko and S.-J. Lee, "Center weighted median filters and their applications to image enchancement," *IEEE Trans. Circuits Syst.*, vol. 15, pp. 984–993, Sept.i 1991.
[7] T. Sun and Y. Neuvo, "Detail-preserving median based filters in image processing," *Pattern Recognit. Lett.*, vol. 15, pp. 341–347, 1994.
[8] D. Florencio and R. Schafer, "Decision-based median filter using local signal statistics," in *Proc. SPIE Int. Symp. Visual Communications Image Processing*, Chicago, Sept. 1994.
[9] J.-S. Lim, *Two-Dimensional Signal and Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1990.
[10] K.-K. Ma, "Put absolute moment block truncation in perspective," *IEEE Trans. Commun.*, pp. 284–286, Mar. 1997.
[11] K.-K. Ma and S. Zhu, "Fundamental error analysis and geomeetric interpretation for block trunction coding techniques," *Signal Process.*, vol. 15, no. 10, pp. 859–867, 2000.
[12] R. Yang, M. Gabbouj, and Y. Neuvo, "An efficient design method for optimal weighted median filtering," in *IEEE Int. Symp. Circuits Systems (ISCAS'94)*, Chicago, IL, Sept. 1994.
[13] H.-L. Eng and K.-K. Ma, "Motion trajectory extraction based on macroblock motion vectors for video indexing," in *Proc. Int. Conf. Image Processing*, Kobe, Japan, Oct. 23–25, 1999.

How-Lung Eng (S'97) received the B.Eng. degree in electrical engineering from Nanyang Technological University, Singapore, in 1998, where he is currently pursuing the Ph.D. degree.

His research interests include content-based image/video indexing and retrieval, pattern recognition, clustering and digital image/video coding and processing.

Kai-Kuang Ma (S'80–M'84–SM'95) received the B.E. degree from Chung Yuan Christian University, Chung-Li, Taiwan, R.O.C, in electronic engineering, the M.S. degree from Duke University, Durham, NC, and the Ph.D. degree from North Carolina State University, Raleigh, both in electrical engineering.

From 1984 to 1992, he was with IBM Corporation, Kingston, NY, and IBM Research, Triangle Park, NC, and engaged on various advanced DSP and VLSI product developments. From 1992 to 1995, he was with the Institute of Microelectronics (IME), National University of Singapore, working on MPEG video research. He joined the School of Electrical and Electronic Engineering, Nanyang Technological University (NTU), Singapore, in 1995, and currently, he is an Associate Professor. His research interests mainly focus on digital image/video coding, standards, and procesing, content-based image/video indexing and retrieval, pattern recognition, and multimedia networking and services. He has had numerous publications in these areas.

Dr. Ma has been serving as the Singapore MPEG Chairman and Head of Delegation, as well as the Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS FOR VIDEO AND SIGNAL PROCESSING since 1997. He is currently the Chairman of Singapore Chapter of IEEE Signal Processing Society. He has been acting as program committee member and session chair of multiple IEEE international conferences, and reviewing papers for various IEEE Transactions, conferences, and other international journals.