Journal of Intelligent and Robotic Systems 7: 257–285, 1993.

Noise Removal from Color Images

J. ZHENG
UPS Research and Development, 51–53 Kenosia Avenue, Danbury, CT 06810, U.S.A.

K.P. VALAVANIS
Intelligent Robotic Systems Laboratory, The Center for Advanced Computer Studies, The University of Southwestern Louisiana, Lafayette, LA 70504-4330, U.S.A.

and

J.M. GAUCH
College of Computer Science, Northeastern University, Boston, MA 02115, U.S.A.

(Received: 6 August 1990; in final form: 7 May 1992)

Abstract. The noise effects in color images are studied from the human perception and machine perception point of view. Three justifiable observations are made to illustrate problems related to individual color signal processing. To minimize the noise effects, two solutions are studied: One is a 'rental scheme' and the other is a vector signal processing technique. The 'rental scheme' adopts filters originally developed for grey scale images to color images. A set of heuristic criteria is defined to reconstruct an output with minimum artifacts. The vector signal processing technique utilizes a median vector filter based on the well developed median filter for grey scale images. Since the output of the filter does not have the same physical meaning as the median defined in one-dimensional space, the search of a vector median is considered as a minimum problem. The output is guaranteed to be one of the inputs. Both approaches are shown to be very effective in removing speckle noise. Results from real and synthetic images are obtained and compared.

Key words. Color image analysis, chromaticity-preserving median filter, color perception, monochromatic-like processing, vector signal processing, mean type filtering.

1. Introduction

In computer natural scene analysis, color images are being widely used due to the superior characteristics over black-and-white images [6, 8, 9, 11]. Consider, for example, several adjacent objects with different color but identical luminosity. These objects may be interpreted as a merged object in a grey level image due to the constant intensities, but can be reflected distinctively in a color image. The added color information clarifies every possible ambiguity. Besides carrying pictorial information of reflected energy, color images contain object spectral information and provide valuable details for quick visual search, inspection and accurate object classification. Moreover, the surface color is relatively invariant to (intensity) illumination changes due to the property of color constancy. Therefore, color information processing is relatively consistent when the environmental illumination conditions change, and provides more reliable and accurate results for machine perception and natural scene analysis.

Any image acquired by optical, electo-optical or electronic means is likely to be degraded by the imperfection of the sensing mechanisms. Potential degradations may

occur in the form of a sensor noise, photographic grain noise, blur (camera out-of-focus, relative object-camera motion), random atmospheric turbulence, etc. The presence of noise in images represents an irrecoverable loss of information. Several filters have been designed to remove noise from images.

All filters proposed to process grey scale images may be classified either as linear or nonlinear filters. Linear type filters include inverse filters, Wiener filters, geometrical mean filters. Nonlinear type filters include the various types of median filters and nonlinear mean filters. Modifications have also been proposed [22–25]. Median-type filters perform very well and sharpen the image edges during the first few iterations. The operating principle is to replace the center pixel in a defined window with the median value of the pixels within the window. However, repeated iteration blurs the edges and removes thin details. Due to their superior characteristics in removing impulse noise, there have been several attempts to develop various median filters. For example, the KNN-median filter proposed by Davis *et al.* [22] is similar to the KNN-mean filter, except that the median of the K selected pixels is substituted as the new value of the center pixel, instead of the mean. To some extent, this has the same advantages over the KNN-mean filters, as the simple median filter has over the mean filter; it pulls the substituted value to one side or the other of the mean, sharpens the edges better and smears the details less. The proposed sigma filter [24], motivated by the sigma probability of the normal distribution, smooths the image noise by averaging only those pixels in the neighborhood which have their values within a fixed sigma range from the center pixel. Consequently, edges are preserved and thin details are retained.

Impulse noise, sometimes called 'outliers' because of the outlying observations, may be created during the image acquisition or transmission. The data recorded in this case are completely outlying due to errors in the acquisition device (e.g. in A/D converter), film grain, recording media, and transmission channel. They appear either as very high positive values or very negative values, called 'salt-and-pepper' noise. The mathematical expression of the noise in the spectrum domain is usually modeled as a long-tailed distribution. It has been shown that linear filters have poor performance in the presence of such noise, and quite often the neighborhood of the noise-contaminated pixel is affected after the images are processed; nonlinear filters are quite successful. Although the exact characteristics of the human visual system are not well understood, experimental results indicate that the first processing levels of the human visual system possess nonlinear characteristics. Various nonlinear filters have been therefore devised. In addition, various criteria, such as the maximum entropy criterion, lead to nonlinear solutions. One of the most popular families of nonlinear filters for noise removal are order statistic filters. Of them, the median filters which represent one of the simplest type of such filters were first suggested by Tukey in 1971 [26]. It has been proven that the median type filters are the most effective in removing impulse noise in grey images [23].

Color images, just like the formation of grey images, are gone through a series of noise-contamination processes. However, the noise removal problems are generally

not considered in most of the research related to natural (color) scene analysis and computer color vision [11, 14–16]. The work by Faugeras [27] represents one of the few papers on color image enhancement. A simple model of human color vision is introduced. It is shown that the perceptual space defined by the model, offers a meaningful way of thinking about important perceptual parameters such as brightness, hue and saturation. A concise formalism is also given to describe those parameters quantitatively. Both, color image enhancement and transmission coding are tried and proved to be quite successful based on the idea of structuring the perceptual space as a vector space, where vector addition corresponds to tristimulus values multiplication.

Color images as raw data in most vision systems consist of red (R), green (G) and blue (B) components. Thus, one may argue that noise could be removed by processing these three single spectrum intensity images individually. This technique, called *monochromatic-like processing* [21] is being widely used regardless of the noise properties (correlation and uncorrelation among the three intensity data). The results are, in general, not optimal and may be misleading because the method ignores the correlation among the data (tristimuli).

The noise removal problem is ill-conditioned and most of the filters bring artifacts into the images. For example, the mean filters redistribute the noise error to the surrounding and the median filters may disorder the intensity levels [17]. For grey images, the introduced artifacts may not affect too much the image appearance and structure, but for color images, the same amount of artifacts may intolerantly distort the image appearance and structure to the point that subsequent image processing may be further complicated. These artifacts are obviously sharp in color images, i.e., a small disturbance in one of the three intensity data may cause a big jump in color perception.

This paper considers mainly the median type filtering in spatial domain which possesses some attractive properties for impulse noise removal in color images, and investigates the applicability of a 'rental scheme' for noise removal in color image processing. By the nature of the filters, the 'rental' scheme adopts original filters developed for grey scale images to color images. This is part of the general problem that deals with the applications of nonlinear monochrome image processing operators to color image data, in an attempt to minimize color artifacts. In reality, there will always be color artifacts, therefore, the objective is to minimize them. Although the basic idea may not be new, the application of median filters (edge preservation, smoothing) to color images is new. A set of criteria is derived and accounted for in order to apply existing 'grey' filters for color image noise removal. As an alternative, a median vector filter is studied. Thus, the results from both filters, given the same inputs, are obtained and compared. In addition, a mean type filtering operation is proposed by incorporating and utilizing the correlation information between color image data channels. A conclusion is made that is not trivial to apply techniques devised for grey scale images to color images without considerable modifications. The effectiveness of the criteria is determined both subjectively and experimentally.

Applications of the proposed method to robotics include: robot navigation systems, automated guided vehicles (AGVs) for undersea and space exploration and assembly line inspection operations. Nonrobotic applications include prepressing industry, medical imaging and marine science and biology. The proposed method has already been applied to underwater photographic analysis to monitor population structure and growth of living organisms [21].

Section 2 of the paper describes the conventonal use of median filters and justifies the proposed modifications. The third and fourth sections present scalar and vector median filters, respectively. Two algorithms are outlined and results one included and compared in Section 5. The sixth section introduces mean type filtering. Section 7 concludes the paper.

2. The Conventional Use of Median Filters: Justifications for the Proposed Modifications

Median filters are widely used for noise suppression in early stages of a vision system due to the following properties [9]:

(1) They preserve the ramp edges and boundaries of the objects,
(2) They suppress impulses of short duration without significantly modifying other components, and,
(3) They may be implemented easily and fast.

The first two properties are most important. Preservation of object edges or boundaries in a scene is important because all subsequent operations in a vision system depend critically on it. Preservation of information related to object edges or boundaries requires that noise removal occurs without smearing out the surrounding. Due to those properties, and because color is primarily used as an inherent object feature in a scene (thus the preservation of the surface color on objects of interest is highly desirable), it is natural to investigate the efficiency and applicability of median filters to remove noise of short duration in color images. The conventional use of median type filtering in color image processing is shown in Figure 1.

Since three monochromatic signals are digitized simultaneously, at any point (pixel in image) color is determined by the intensities of reflected illumination at that point

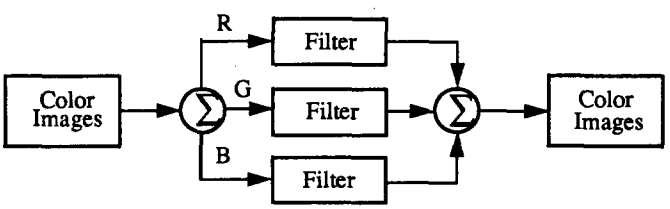

Fig. 1. The conventional use of filters in color image processing.

in terms of the RGB intensity signals, and is expressed as follows:

$$c(i,j) = \begin{bmatrix} r(i,j) \\ g(i,j) \\ b(i,j) \end{bmatrix}, \tag{1}$$

where $c(i,j)$ is the vector pixel at (i,j) in a color image C of size $N \times N$, $r(i,j)$, $g(i,j)$ and $b(i,j)$ are the pixel values in the corresponding spectrum image obtained through spectral integration and sampling. A frame of color image may be expressed as

$$C = \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \tag{2}$$

where

$$C = \{c(i,j);\ i,j \in N,\ c(i,j) \in \Re^3\};$$

$$R = \{r(i,j);\ i,j \in N,\ r(i,j) \in \Re\};$$

$$G = \{g(i,j);\ i,j \in N,\ g(i,j) \in \Re\};$$

$$B = \{b(i,j);\ i,j \in N,\ b(i,j) \in \Re\};$$

In monochrome image analysis, noise from different sources (electronic sensor, film grain, thermal sources) is always assumed to be random and uncorrelated from pixel to pixel. But, in chromatic image analysis, there may be noise correlation among the three monochromatic images (although they remain uncorrelated from pixel to pixel in a single intensity image due to the way the images are formatted). The noise correlation indicates that the correlation matrix R_n may not be necessarily a diagonal matrix. Generally, noise introduced before the electron guns is correlated among the three monochromatic images (for example noise from film grain, spurious effects and random scattering in a scene). Noise introduced from the electron guns or circuitry boards may be treated as in grey images, mostly interpreted as additive independent random noise. Such noise may be removed in a similar way to processing grey images or by using Kalman filters for correlated noise.

Consider that $r_m(i_r, j_r)$, $g_m(i_g, j_g)$ and $b_m(i_b, j_b)$ represent the medians of each intensity (channel) data. Then $c(i_r, j_r)$, $c(i_g, j_g)$ and $c(i_b, j_b)$ are vectors with only one median element in the corresponding RGB channels. They can be expressed as

$$c(i_r, j_r) = \text{med}_r\{c(i_r, j_r),\ i,j \in W\}, \tag{3}$$

$$c(i_g, j_g) = \text{med}_g\{c(i_g, j_g),\ i,j \in W\}, \tag{4}$$

$$c(i_b, j_b) = \text{med}_b\{c(i_b, j_b),\ i,j \in W\}, \tag{5}$$

where W is a selectable odd-size image window of any shape, (i_r, j_r), (i_g, j_g), and (i_b, j_b) are the coordinates of median values of the corresponding channels within the

window W, $\text{med}_r\{\ldots\}$, $\text{med}_g\{\ldots\}$ and $\text{med}_b\{\ldots\}$ are the median functions to process the RGB signals, respectively. For example, the output with the median value $r_m(i_r, j_r)$ through the red channel must satisfy by definition the probability relation

$$\int_{-\infty}^{r_m} f(r)\,\mathrm{d}r \;=\; \int_{r_m}^{\infty} f(r)\,\mathrm{d}r \;=\; \tfrac{1}{2}, \tag{6}$$

where r is considered as a variable and $f(r)$ is the distribution function. Equation (6) may be simply rewritten as

$$P(r \leqslant r_m) \;=\; P(r \geqslant r_m) \;=\; \tfrac{1}{2}. \tag{7}$$

Alternatively, in the discrete case, a similar expression may be written as

$$r_m(i_r, j_r \;=\; \min_{j \in N} \left\{ \sum_i |r_i - r_j| \right\}, \tag{8}$$

where $\min\{\ldots\}$ is a function of choosing the minimum of N values of distance measurement. Since each time only one element of a vector is produced from a single data channel, combination of the three vectors (median element and two associated elements) forms a 'median matrix' represented as

$$\begin{aligned} M \;&=\; [c(i_r, j_r) \quad c(i_g, j_g) \quad c(i_b, j_b)] \\ &=\; \begin{bmatrix} r(i_r, j_r) & r(i_g, j_g) & r(i_b, j_b) \\ g(i_r, j_r) & g(i_g, j_g) & g(i_b, j_b) \\ b(i_r, j_r) & b(i_g, j_g) & b(i_b, j_b) \end{bmatrix}. \end{aligned} \tag{a}$$

Each column in M represents a real vector pixel in the original image with $1 \leqslant \text{rank}\,(M) \leqslant 3$. The conventional use of median filters in color image processing (individual R, G, B channels), gives a new vector median $[r(i_r, j_r) \quad g(i_g, j_g) \quad b(i_b, j_b)]^T$, the diagonal elements of the median matrix M. By looking at M, observe the following:

Observation 1: The vector $[r(i_r, j_r) \quad g(i_g, j_g) \quad b(i_b, j_b)]^T$ is the real vector median that represents one of the vector pixels to be filtered in the window W only under the condition that each pixel value in the vector is simultaneously the median value of all corresponding three data channels. Thus, rank $(M) = 1$, hence, $i_r = i_g = i_b$; $j_r = j_g = j_b$; a real vector median which is one of the inputs is extracted. Therefore, it may be argued that the pixel values in each data channel are strictly balanced, or that the number of values less than the vector median and the number of values greater than the vector median are the same in all different data channels.

As an example, visualize a two-channel median filtering process. Figure 2 shows 9 vector pixels to be median-filtered individually. Apparently pixel 5 is the median value of both data channels, which means that the output of the filter is one of the input data. A close look at the data structure shows that the set of data is strictly balanced, namely the elements of the output are medians in all data channels.

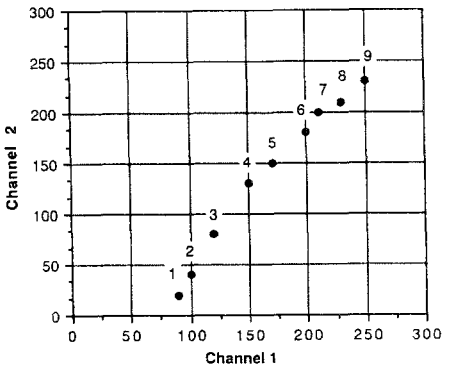

Fig. 2. (Observation 1). A two channel median filtering process.

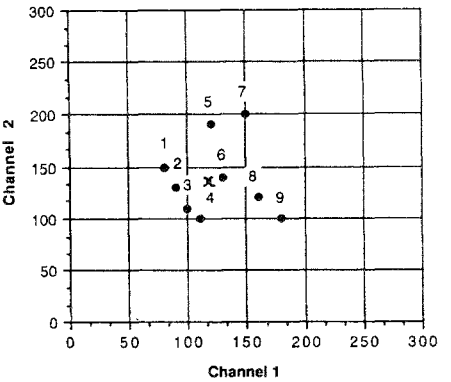

Fig. 3. (Observation 2). A two channel median filtering process. x is the new pixel.

Observation 2: If the vector pixel sequence in one data channel is not strictly balanced, i.e., $2 \leqslant \text{rank}\,(M) \leqslant 3$, the diagonal elements $[r(i_r, j_r) \quad g(i_g, j_g) \quad b(i_b, j_b)]^T$ in the median matrix do not represent one of the original vectors. The output becomes unpredictable, hence artifacts may be introduced which sometimes distort the color appearance if the output is not restricted within a certain range. For example, the created median vector may lie somewhere in between if there are two clusters. Consider again a two-channel median filtering process as shown in Figure 3 in which pixel 5 is the median value of channel 1, but pixel 2 is the median value of channel 2. A new vector pixel is introduced labeled by x in Figure 3. The output of the filter is no longer a member of the input family.

To be more specific, consider a median matrix with the following entries:

$$M = \begin{bmatrix} 100 & 85 & 80 \\ 90 & 100 & 95 \\ 80 & 80 & 100 \end{bmatrix}$$

Any column in M represents a color with a valid hue. Individual R, G, B channel processing creates a new vector pixel $[100 \quad 100 \quad 100]^T$ which obviously distorts the color appearance and makes it impossible to process color information in subsequent operations because the hue is totally lost (hence the undefined saturation). An achromatic spot is introduced and shows an ugly grey spot in a color image. On the other hand, there is hardly a way to recover the original color information from this pixel, and the side effects of this singular spot may complicate further processing. If vector pixel $[85 \quad 95 \quad 80]^T$ is used instead, and the reconstructed vector pixel does not bring too much offset in chromaticity, then the problems with the diagonal selection may be vanished. *This suggests that hue preservation is highly desirable.*

Observation 3: If $\text{rank}\,(M) \neq 1$, besides introducing artifacts, conventional processing may shift edges, called edge jitter, thus distorting the scene structure and

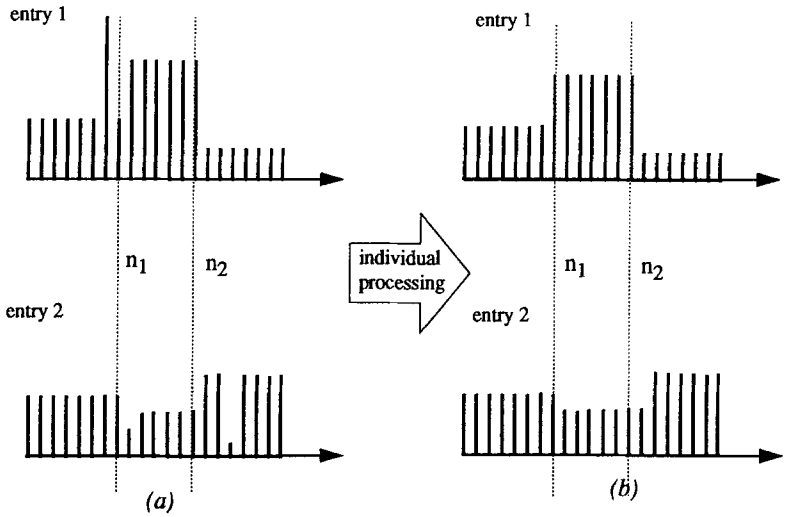

Fig. 4. (Observation 3). The edge jitter problem.

making pixel or edge-based measurements inaccurate. Figure 4 illustrates this problem using a one-dimensional vector signal with only two entries. Since there are two distinctions at n_1 and n_2, respectively, the original signal represents three segments, left band bounded-right by n_1, the middle band bounded-left by n_1 and bounded-right by n_2, right band bounded-left by n_2. Both entries are contaminated by impulse noise, for example, one at $n_1 - 1$ and the other at $n_2 + 3$, as shown in Figure 4(a). The overall picture shows two spots. The result of individual processing with a window size of 5 pixels shows that a new-painted impulse at n_1 is generated and the distinction which separates two segments in the original image is shifted by one pixel position. The same thing happened at $n_2 + 1$, too. From this observation, one may conclude that individual processing does not remove the impulse noise, instead, it moves the noise position to contaminate its neighbor values in the image.

3. Scalar Median Filters

Color image smoothing by monochromatic-like processing has the drawback that a vector pixel may be painted unexpectedly due to the order of intensity values of its neighbors. Pixel classification of the processed image (spectral information) may produce significant errors. It is therefore necessary to devise a strategy to minimize the color distortion.

Observe that the median matrix M consists of at most three different candidates for vector medians through median filtering of the three data channels. Each of the vector medians is a real vector in the image to be filtered, though only some entries of the vector may be degraded due to noise. However, the nondegraded entries contain the partial faithful information of original color which is the clue one should use to predict or reconstruct a vector pixel to replace the degraded one. Additional

information may be also extracted from the neighbors because the pixels are spatially and spectrally correlated.

From Observation 2, it may be deduced that a two-channel median filter gives two new combinations of processed pixels plus the two original vector pixels, to form a median matrix for a possible output vector median. The diagonal vector median is the output of individual signal/channel processing. Accordingly, if we median-filter a color image through the three data channels, there are 27 candidates for a vector median computed as follows:

$$\binom{3}{1} \times \binom{3}{1} \times \binom{3}{1} = 27 \tag{10}$$

where

$$\binom{n}{m} = \frac{n!}{m! \times (n - m)!}$$

is read as 'n choose m'. Three are the original, one is the output from monochromatic-like processing, and the other 23 are the remaining possible vector medians. The focus of removing noise utilizing the 'rental scheme' is to reconstruct an optimum output from the median matrix with respect to both human perception (color fidelity) and subsequent operations (edge/boundary preservation). One way to do so is to use the concept of median filters along with a set of criteria derived from human visual experiences, such that the result satisfies the human visual perception. That is, we are looking for the *heuristic perceptually optimum output with minimum artifacts.*

However, since the noise removal problem itself is an ill-conditioned problem as such, it lacks a unique solution. There may exists an extremely large set of candidate solutions to the problem. All possible solutions may sound mathematically valid. But a valid mathematical solution may not give the same results compared to the perceptually optimum results. The criterion of choice for a candidate solution is not only the mathematical validity of the solution (which is *necessary but not sufficient* for visual quality of the result), but also the extent to which a solution meets some appropriately chosen optimum criteria which is felt to affect the visual quality of the result.

Hue (H), intensity (I) and saturation (S) may be used to describe a color. These terms are abstracted from complete visual experiences and used to represent dimensions along which color may vary. They are functionally related to many factors, for example, the spectral characteristics of the stimulating energy E, the spectral matching functions of a particular observer $(\bar{r}, \bar{g}, \bar{b})$, the observer's memory M for similar objects, the surround S, adaptive state of the observer A, neighboring objects O, the observer's attitude at the moment T and so on. In general, a color may be described by

$$\begin{aligned} \text{Color} &= (H, I, S) \\ &= \mathfrak{I}(E, M, S, A, O, T, \bar{r}, \bar{g}, \bar{b}, \ldots), \end{aligned} \tag{11}$$

where $\Im(\ )$ is an abstract function used to represent a color. The three dimensions of hue, intensity and saturation are the summary of these aspects which, although may not necessarily comprise a unique coordinate system, represent a widely accepted one. To reconstruct an output from the median matrix in terms of human perceptual satisfaction, the three functions widely used in color science to judge these items are adopted:

$$h = H(r, g, b), \qquad i = I(r, g, b), \qquad s = S(r, g, b), \tag{12}$$

where H, I and S are definable functions but not necessary monotonic. h, i and s are the abstract quantitative values of each single vector pixel that combine all the effects of the original raw data. Different r, g, b values may result in identical h, i, s values since these functions are primarily human oriented. The advantage of using hue, intensity and saturation, defined in an empirically derived color space, is that the color judgement on which the space is based, is performed in perceptual terms; a natural aligning of perceptual variables, therefore, results. However, in practice, is is noted that the transformation from the image data represented in RGB to the human oriented function h, i and s is ill-conditioned [12]. The transformation has a non-removable singularity and spurious gaps. For example, no hue and saturation can be defined when the values in red, green and blue channels are identical. This situation is called perturbations which are annoyingly significant and not easily avoidable [12]. In order to reduce the possibility of perturbation occurrence, the evaluation of h, i and s may be obtained as follows:

$$h = H(\bar{r}, \bar{g}, \bar{b}), \qquad i = I(\bar{r}, \bar{g}, \bar{b}), \qquad s = S(\bar{r}, \bar{g}, b), \tag{13}$$

where $\bar{r}$, $\bar{g}$ and $\bar{b}$ are the average values for red, green and blue channel within a window, respectively. Note that

$$H(\bar{r}, \bar{g}, \bar{b}) \neq \bar{h} = \int\int h(x, y)\, dx\, dy,$$

$$I(\bar{r}, \bar{g}, \bar{b}) \neq \bar{i} = \int\int i(x, y)\, dx\, dy,$$

$$S(\bar{r}, \bar{g}, \bar{b}) \neq \bar{s} = \int\int s(x, y)\, dx\, dy.$$

In order or minimize the distortions in color appearance when an image is processed, the following criteria are proposed:

(1) The hue changes should be minimized,
(2) The shift of saturation is made as small as possible and it is better to increase than to decrease the saturation, and,
(3) It is desirable to maximize the luminance contrast.

The criteria are listed in order of importance with respect to the sensitivities of the human perceptual system. Due to the fact that the three signal components have different effects on the image appearance, it is not a wise idea to combine all three criteria to remove the spurious effects in a color image. Therefore, while reconstructing

an output with minimum artifacts from the median matrix, the initial candidates are initially filtered in accordance with the first criterion:

$$[r_l \quad g_m \quad b_n]^T = [r_i \quad g_j \quad b_k]^T; \quad \text{if } \min\{H(r_i, g_j, b_k) - h\}, \tag{14}$$

where $l, m, n \in i, j, k$, a subset and $i, j, k = 1, 2, 3$. Notation $\min\{\ \}$ means the minimum value of the difference measurement. Those data which may hit the singularity in the transformation are excluded. Since hue is judged separately from the other two terms, it is likely that there are more than one possible candidate for an optimum output after this criterion. Several possible vector medians may possess the same hue and be close to the overall average hue $\bar{h}$ in the region. If the number of $[r \quad g \quad b]^T$ is larger than one, then the second criterion applies to reduce the number of candidates:

$$[r_x \quad g_y \quad b_z]^T = [r_l \quad g_m \quad b_n]^T \quad \text{if } \min\{S(r_l, g_m, b_n) - s\}, \tag{15}$$

where $x, y, z \in l, m, n$ and $l, m, n = 1, 2, 3$. Thus, those vector medians with the same hue are sorted out if other vector medians are closer to and more saturated than the mean s. If the number of the candidates is larger than one, it is still possible for very few vector medians to have identical hue and saturation. Then the application of the last criterion will produce a unique output, namely the vector median with the highest intensity value is selected. The optimization procedures in the above order to search for a *heuristic perceptual-optimum output* based on the outputs of the median filters, preserve the properties of a median filter and do not introduce very dispersive color pixels. The reconstruction of the output which is closest to the overall hue and saturation guarantees that the neighbors are not smeared out when moving the median windows.

The RGB color space used in most machine vision systems is solely machine-oriented. The transformation from the raw data to other spaces is not recommended due to the non-removable singularity and instabilities induced by the existence of noise in the raw data. Nevertheless, the human oriented color respresentations are good clues to evaluate the human visual effects of the RGB values. Local measurement using hue, saturation and intensity after the raw data is preprocessed reveals the applicability of the proposed method. The optimum vector pixel median obtained through the criteria, is perceptually acceptable in terms of chromaticity preservation and is more conformable with human color vision.

ALGORITHM IMPLEMENTATION

Step 1: Individual processing of each component.
Step 2: Form the median matrix.
Step 3: Remove the data set hitting the singularity.
Step 4: Apply Equation (13).
Step 5: Get initial candidates based on Equation (14).
Step 6: Narrow down the candidates according to Equation (15).
Step 7: Obtain the output from Equation (15).

EXPERIMENTAL DATA

Real and synthetic color images are processed with either impulse noise of short duration or Gaussian noise. It is observed empirically that under very few circumstances, the diagonal elements of the median matrix M are selected as a vector median if a noisy image is given without prior knowledge about the noise. The choices are usually among the other 26 possible vectors, including the three original vector pixels. Also, it is observed that the fixed points in raw data are more stable when the perceptual optimum criteria are followed, than in the simple diagonal selection.

The optimization of the output is entirely image-and-region-dependent, thus error analysis becomes mathematically complicated. Since median filters are nonlinear, this complicates the mathematical analysis of their performance even for a scalar image. It is even more difficult to analyze the errors when filtering the vector signals. The main difficulty lies in the fact that it is not possible to separate signal effects and noise effects as may be done in linear filtering analysis.

To illustrate the processing results, a one-dimensional sequence of vector signals, extracted from an image, is used. Since the quantitative evaluations are too complicated to by physically meaningful, we used the hue measurement with respect to the human color perceptual system to evaluate the effect of noise removal. Hue is generally considered in colorimetry and human color vision as a single measurement of color experience, and represents the most sensitive direction of the three dimensional space in color image analysis. Figure 5 shows a set of noise-free original vector signals used for experiments. The utilization of the hue measurement (normalized) shows that there are two quite different segments (regions), the left part falls in the blue region (190°) and the right part in the green region (80°), in the hue chart shown in Figure 5(b). The edge between them is apparent if the intensity and saturation are appropriate. Figure 6(a) shows the signals contaminated by noise of short duration. The noise appears unpredictably colorful and very perceivable in a color image, while

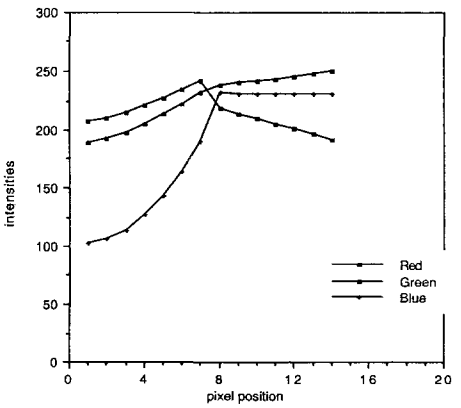

Fig. 5(a). Original vector signals.

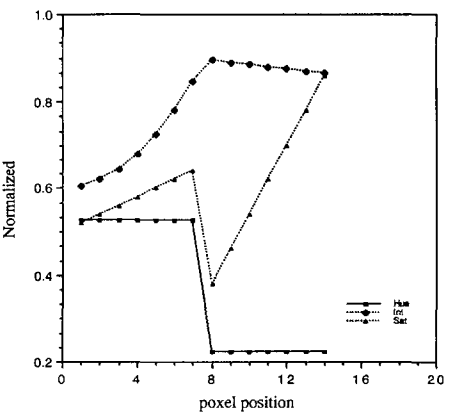

Fig. 5(b). The hue measurement of the original signals.

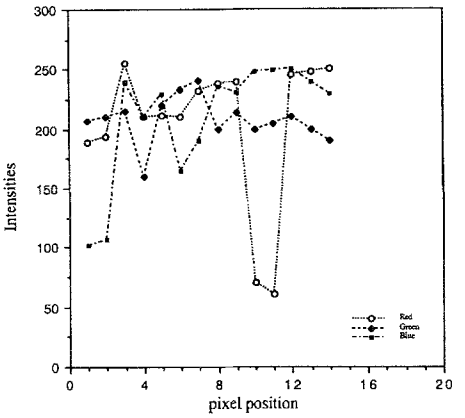

Fig. 6(a). The noise contaminated vector signals.

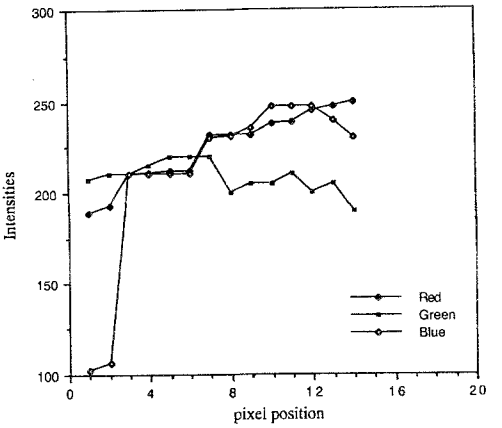

Fig. 6(b). The outputs of the conventional median processing.

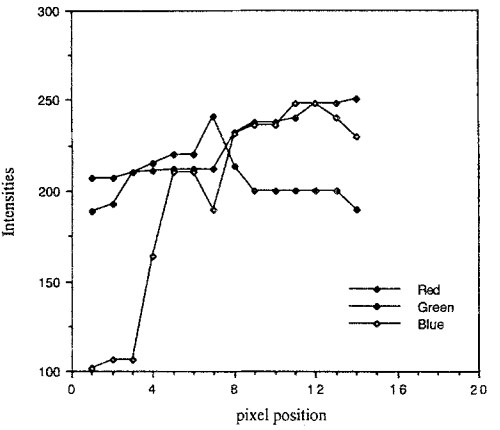

Fig. 6(c). The heuristic perceptually optimum output.

in grey images the noise affects only the grey levels and some of them may not be visible at all. Figure 6(b) shows the results performed by the conventional simple method in which each element of the vector is processed individually. The output comes out from the diagonal elements of the median matrix which ignores the interdependent information. Figure 6(c) shows the result of optimizing the median matrix using the proposed criteria. The output is the best reconstruction of the elements of the median matrix with respect to human color perception. Since it is difficult to visualize the significance from the one-dimensional RGB data, we adopt the hue measurement which is shown in Figure 7. Observe that the measurement of the noisy signals indicates that two color regions break into several unstructured pieces. The simple median processing scheme does suppress part of the noise but, meanwhile, introduces some fatal effects which make subsequent operations impossible.

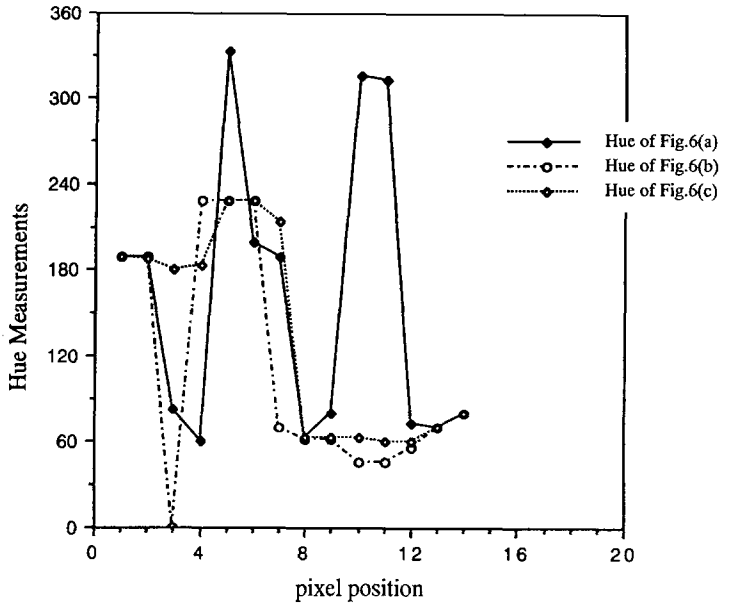

Fig. 7. The hue measurement.

For example, all color information that vector pixel at 3 carries, is virtually lost because the hue at that pixel is undefined (hence the saturation). The effect may be clearly seen in Figure 6(b) in which the vector pixel contains the identical elements caused by the diagonal selection form the median matrix. But the optimization procedure avoids the 'grey' spot, selecting the best elements to form a vector median according to the surroundings which theoretically eliminates the possibility of losing chromaticity information.

4. Vector Median Filters

One of the most important properties of median filters is that they do not introduce any sample values that are not present in the input data and the output is always one of the inputs. However, the method proposed in the previous section cannot generally guarantee that the reconstructed output is one of the inputs. Therefore, an alternative design complying with the original idea of median filters is introduced.

In the scalar case, the search for the median values is based on order statistics of the measured samples. This idea cannot be (directly) extended to vector cases, because when dealing with vector functions concepts such as 'greater than' or 'less than', necessary for finding the sample order statistics, do not have the same meaning. Moreover, it is known that the concepts of the mean and mode may be extended without any ambiguities to vector signals. However, this is not the case for the median. Observe that Equations (7) and (8) are used to define a median, but they have a physical difference. Since they happen to be identical in values in the scalar case, they

are used indistinguishably. To be more specific, Equation (7) defines an arithmetic median while Equation (8) defines a geometric median. When one considers the vector case, these two definitions arer apparently no longer equivalent. It is essential to distinguish the two different median operations for the vector case. The arithmetic definition cannot guarantee that the output be one of the inputs, the geometric definition does. The search is to find an output which is one of the inputs. The vector median operation is defined as a geometric vector median. The definition follows directly:

Definition: Given a sequence of vectors $c_1, c_2, \ldots, c_N$, a vector median c_m must be an element of the same sequence and satisfy the condition:

$$c_m = \min_{j \in N} \left\{ \sum_i \| c_i - c_j \|_n \right\}, \tag{16}$$

where c_m is the vector median to be found. The norm subscript n is a nonnegative constant which is normally chosen from the constants [1, 2, ∞] to achieve certain objectives. For instance, if c_m is unreliable, $n = 1$ tends to put relatively small weight on the so-called data outlier, namely outlying observations in statistics. On the other hand, minimization of the largest model residual requires $n = \infty$. In the sequel, $n = 2$ is chosen; it represents the Euclidean distance. This choice has been made to explain the vector median graphically, known as 'the point of minimum aggregate distance (travel)'. Thus, Equation (16) becomes an extended version of a scalar median; a vector in a multidimensional space. The search for the vector median is a *minisum* problem.

The vector median filtering is also a nonlinear and shift invariant operation. Without loss of generality, assume that α and β are two scalars which scale the elements of a set of vectors $c_1, c_2, \ldots c_N$. Then, the vector median after being scaled, may be defined, respectively, as

$$c_m^{\alpha} = \min_{j \in N} \left\{ \sum_i \| \alpha c_i - \alpha c_j \| \right\} = |\alpha| \min_{j \in N} \left\{ \sum_i \| c_i - c_j \| \right\}, \tag{17}$$

$$c_m^{\beta} = \min_{j \in N} \left\{ \sum_i \| \beta c_i - \beta c_j \| \right\} = |\beta| \min_{j \in N} \left\{ \sum_i \| c_i - c_j \| \right\}, \tag{18}$$

while

$$\begin{aligned} c_m^{\alpha+\beta} &= \min_{j \in N} \left\{ \sum_i \| (\alpha + \beta) c_i - (\alpha + \beta) c_i \| \right\} \\ &= |\alpha + \beta| \min_{j \in N} \left\{ \sum_i \| c_i - c_j \| \right\} \\ &\leqslant (|\alpha| + |\beta|) \min_{j \in N} \left\{ \sum_i \| c_i - c_j \| \right\}. \end{aligned} \tag{19}$$

The result is based on the fact that $|\alpha + \beta| \leqslant |\alpha| + |\beta|$ for any constants α and β. Clearly, the shift of a set of data does not affect the median, hence, shift-invariance.

Vector	c_1	c_2	...	c_N	S_{norm}
c_1	d_{11}	d_{12}	...	d_{1N}	$\sum d_{1i}$
c_2	d_{21}	d_{22}	...	d_{2N}	$\sum d_{2i}$
⋮	⋮	⋮		⋮	⋮
c_N	d_{N1}	d_{N2}	...	d_{NN}	$\sum d_{Ni}$

Fig. 8. The search of vector median.

Based on the distance measurement, the search for vector median becomes straightforward. Figure 8 illustrates the search for such a vector median. The norm-2 distance $d_{ii}(i = 1, 2, \ldots, N)$ must be zero and $d_{ij} = d_{ji}(i, j = 1, 2, \ldots, N)$. The minimum value S_{norm} corresponds to the vector median:

$$c_m = \min\left(\sum_{i=1}^{N} d_{ji}, j = 1, 2, \ldots, N\right). \tag{20}$$

Although the above definition shows that there should be at least one solution (existence), it does not guarantee that the output of the vector median filter produces a unique answer (non-uniqueness). Very often, there are more than one vectors satisfying the vector median conditions and all may be a valid vector median. Figure 9 shows a possible distribution of a sequence of two-dimensional vector signals. Both vectors c_1 and c_2 may be the outputs of the filter. If the candidates are distributed symmetrically, either one may be selected as the output; but preference is given to the one whose original is closer to the position of the vector median. If the distribution is not symmetric, the best solution may be the output of an enlarged window. In other words, the output may be determined from a window of size 5×5 if the output of a window size 3×3 is not unique.

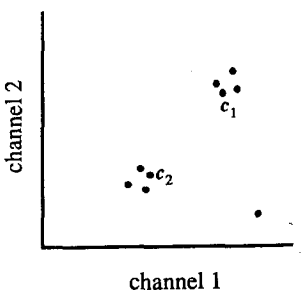

Fig. 9. A possible symmetric distribution.

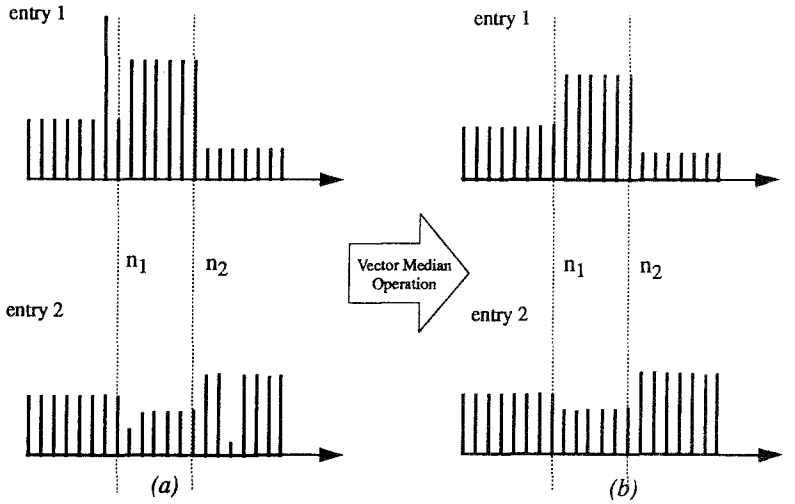

Fig. 10. The result of vector median operation.

ALGORITHM IMPLEMENTATION

Step 1: Distance computation between all the pairs.
Step 2: Store all results in a vector.
Step 3: Sort the elements.
Step 4: Find the minimum distance value.
Step 5: Determine the vector median based on Equation (20).

EXPERIMENTAL DATA

The vector median filter is implemented in the same way as the scalar median filter. A window of certain size is placed on each vector sample and the sample remains the same if its median is within the window, or is replaced by a new value which is a median in the window (note that the definition of vector median does not require odd window size). The data for observation 3 (Figure 4) is used for testing the vector median operation. For a complete filtering, we append to the beginning of the signal 2 additional constants equal in value to the first sample of the signal. Similarly, 2 constant samples are appended to the end of the signal. By doing so, we assure that when the initial signal's first or last sample is in the center of the window, the median filter output equals this sample value. When a signal passes through the median filter unchanged, this means that the central sample value for each window position is itself the median of the samples within the window. If all the signals are assigned a numerical value, the distance can be calculated and the minimum one which corresponds to the vector median is searched within a one-dimensional window of 5 elements. The result is shown in Figure 10. It may be observed that the edges which separate the two segments are preserved and there is no artifact introduced using the vector median operation.

5. Comparisons of the Two Median Type Filters

The performance of median filters varies with the shapes of the chosen window. It also depends on the contents of the images as well as the sizes of the operator windows. These variations make the comparisons between the results of the two proposed median type filters difficult. In the following experiments we use both synthetic and real images with 8 bits per pixel processed through a conventional 3×3 square window. A mean squared error measurement is used for comparison purposes. To get an accurate picture of the proposed techniques, results are also obtained via mean filters (on top of median filters).

Figure 11 shows a synthetic image of a color cube. The image contaminated by impulse noise is displayed in Figure 12. For monochromatic-like processing the color images are decomposed into the RGB individual intensity signals and then the mean and median filter is applied to each one of them. The final results are obtained by combining the three individual outputs. Figures 13(a) and (b) show the results of individual processing in the RGB space. Both, mean and median filters, are used to process each component individually. Then the output is formed directly from the three medians; no correlation information between the data channels is utilized. Observe that the error pixels are redistributed by the averaging operations while some 'crisp' noise in the original image has not been removed but 'blurred' by individual median operations. This phenomenon is the result of edge jitter. For illustration purposes only, the results in other two color descriptive spaces are shown in Figures 13(c), (d), (e) and (f). These two spaces are the HIS and CIE $L*a*b*$. Although there are many other popular spaces used in different applications, these two seem to be of interest to visual signal processing. One is the analogic description of human color perception and the other one is the approximation to perceptual uniform color space. The transformation from the HIS space to the RGB space or vice versa is neither stable nor a one-to-one conversion. There are non-removable singularities (notably when $R = G = B$) as well as spurious effects. Different combinations of RGB may give the same hue, intensity or saturation. Nevertheless, this is one of the

Fig. 11. Synthetic Image.

Fig. 12. Noise-contaminated image.

Fig. 13(a). Result of mean filter in the RGB space.

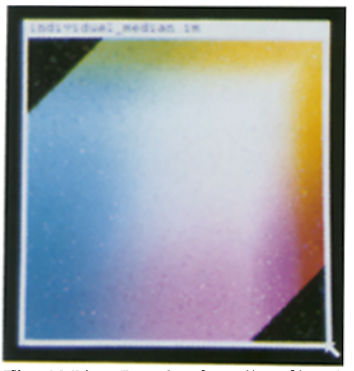

Fig. 13(b). Result of median filter in the RGB space.

Fig. 13(c). Result of mean filter in the HIS space.

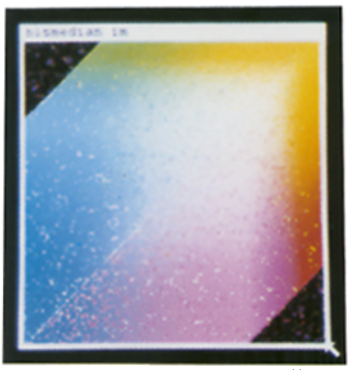

Fig. 13(d). Result of median filter in the HIS space.

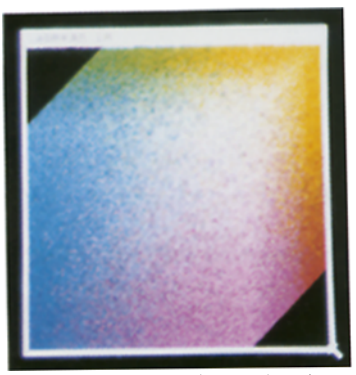

Fig. 13(e). Result of mean filter in the $L*a*b*$ space.

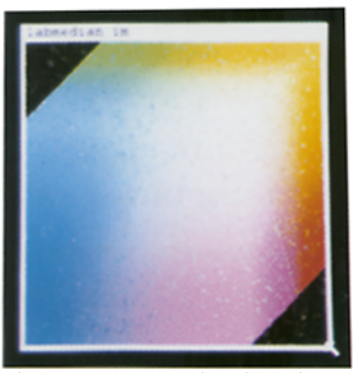

Fig. 13(f). Result of median filter in the $L*a*b*$ space.

very few color spaces representing the way the human visual system perceives color. Figures 13(c) and (d) show the results which are obtained from the individual processing of the hue, intensity and saturation component. As the transformation from RGB to HIS extracts all the color information into one component (hue), the newly formed component inevitably becomes extremely sensitive to any data changes, the error redistribution from the averaging operation, therefore, actually changes the

color appearance as observed in Figure 13(c). Due to the data arrangement in generating the color cube, the error changes in the hue component, happen to alter abruptly the color variations; a diagonal 'line' along the data R = B can be perceived. The phenomenon is much alleviated in the result of median filter because the error is not redistributed to the surrounding pixels. The color appearance is altered only by the wrong combination of the data due to the individual processing. Comparison to the results obtained in the RGB space as well as the consideration of the problems with the transformation, may conclude that the HIS space should not be used for noise removal problems, especially for mean operations, because the result in the hue cannot be predicated.

Figure 13(e) and (f) show the results in CIE $L*a*b*$ uniform color space. Since the transformation from the RGB representation to $L*a*b*$ representation is better defined, there is no distortion in color appearance. As far as the effects of noise removal are concerned, it may be seen that there is almost no difference between the results obtained in the RGB and $L*a*b*$ spaces. This proves that *there is no need* to transform the RGB data into other spaces where the images are filtered and transform the data back to the RGB format. *An efficient approach to the noise removal in color images is to process the data in the RGB space.*

Figure 14 shows the results using the 'rental' scheme which is apparently better than that of all the individual processing. But there are still some spots which have not been removed using only one pass of the filter due to the limited candidates for reconstruction of a final result. The searching of vector median shows the dramatic improvement as demonstrated in Figure 15.

Since the comparison among various techniques is performed on unbiased images, with *a priori* knowledge of the input, it is possible to quantitatively measure the noise cleaning power using a MSE measurement as follows:

$$J(C, C') = \frac{1}{s} \int \int [C(x, y) - C'(x, y)]^2 \, dx \, dy, \tag{21}$$

where s is the image area or the total number of pixels, C and C' are the original and processed images, respectively. The measurement is by no means accurate but gives

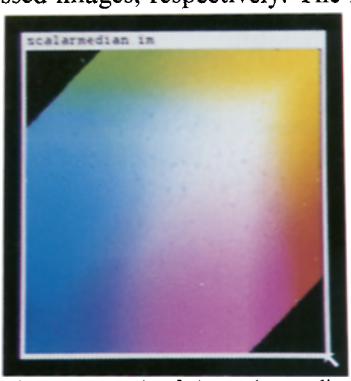

Fig. 14. Result of the scalar median filter.

Fig. 15. Result of the vector median filter.

Table I. The mean squared errors for the synthetic image

Images	MSE values
Original Image (Fig. 11)	0.00
Noisy Image (Fig. 12)	14334.08
Mean Result (Fig. 13.a)	4665.40
Median Result (Fig. 13.b)	1434.38
Mean Result (Fig. 13.c)	8991.26
Median Result (Fig. 13.d)	1629.71
Mean Result (Fig. 13.e)	4719.23
Median Result (Fig. 13.f)	1435.83
Scalar Median (Fig. 14)	946.01
Vector Median (Fig. 15)	292.67

a comparison benchmark. This is a global measure of degradation that treats all spatial frequencies and intensity levels in the image uniformly.

Table I shows the mean squared errors of all filtered images compared to the original image. Obviously the noise-contaminated image has the highest errors. Although it is difficult, from the pictures, to compare the performance of the scalar median filter to the vector median filter, the quantitative measure on both filters shows that the result of the vector median filters produce the best/closest to the original, for this particular image.

Figures 16 to 20 show the results of a real image. Similar comparisons may be observed. The MSE comparison is listed in Table II. The quantitative measurement of MSE again indicates that the results in the $L*a*b$ space is quite similar to that obtained in the RGB space.

Figures 21 to 25 show another set of results of a real image taken under the water. Similar comparisons may be observed. The MSE comparison is listed in Table III.

It may be noticed that based on the MSE criterion, the result of using a vector median filter (in the previous two examples) seems to be better than the scalar median filter. However, this example shows the opposite, namely the scalar median filter

Fig. 16. The original image of Mandrill face.

Fig. 17. The noise-contaminated image of Mandrill face.

Fig. 18(a). Result of using individual mean filter in the RGB space.

Fig. 18(b). Result of using individual median filter in the RGB space.

Fig. 18(c). Result of using individual mean filter in the HIS space.

Fig. 18(d). Result of using individual median filter in the HIS space.

Fig. 18(e). Result of using individual mean filter in $L*a*b*$ space.

Fig. 18(f). Result of using individual median filter in $L*a*b*$ space.

Fig. 19. Result of using scalar median filter.

Fig. 20. Result of using vector median filter.

Table II. The mean squared errors for the Mandrill-face image

Images	MSE values
Original Image (Fig. 16)	0.00
Noisy Image (Fig. 17)	12200.66
Mean Result (Fig. 18.a)	3829.84
Median Result (Fig. 18.b)	2424.06
Mean Result (Fig. 18.c)	6050.44
Median Result (Fig. 18.d)	3224.26
Mean Result (Fig. 18.e)	3902.24
Median Result (Fig. 18.f)	2435.12
Scalar Median (Fig. 19)	2005.47
Vector Median (Fig. 20)	2001.29

seems to be quantitatively better than the median vector filter for this picture. This reflects the fact that there is no universally acceptable best filter. *The results are not only noise-dependent but also signal content-dependent.*

6. On Mean Type Filtering

A mean type filter is generally considered to be better when removing additive white noise (Gaussian or non-Gaussian). To extend from the median vector operation to the mean vector operation, a mean vector filtering technique is proposed using a linear combination of inputs. Although Wiener filters can be used to restore the original images, the proposed approach requires much less computation (for particular applications). The idea of using linear combination of inputs may not be new, but it is believed that the use for a vector signal as an input is challenging.

Mean type filters, if not well designed, may distort crucial details needed to interpret the image (such details once lost can never be recovered). Correlation of information between data channels must be utilized in order to effectively remove noise. Color images as vector signals should not be decomposed, individually processed and then

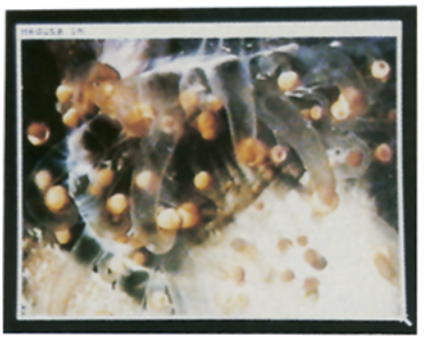

Fig. 21. Original Image of a Medusa under the sea.

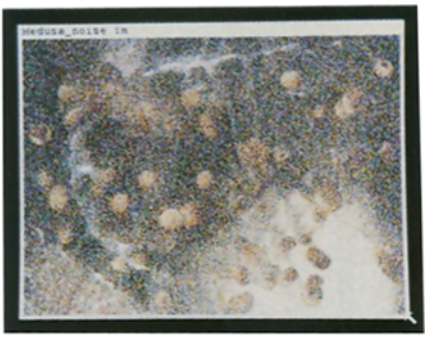

Fig. 22. A noise-contaminated image.

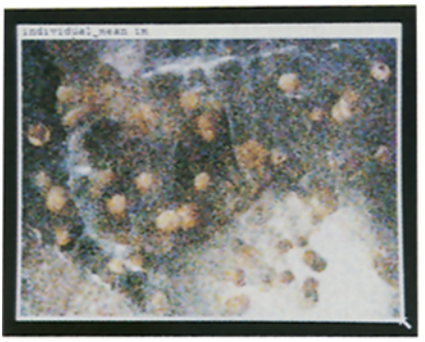

Fig. 23(a). Result of individual mean filter in the RGB space.

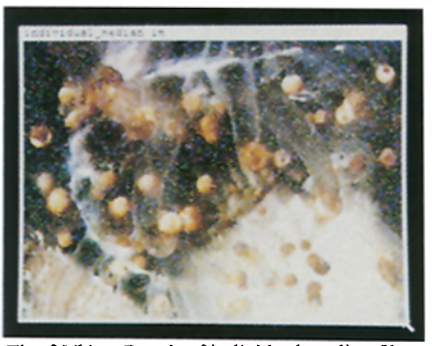

Fig. 23(b). Result of individual median filter in the RGB space.

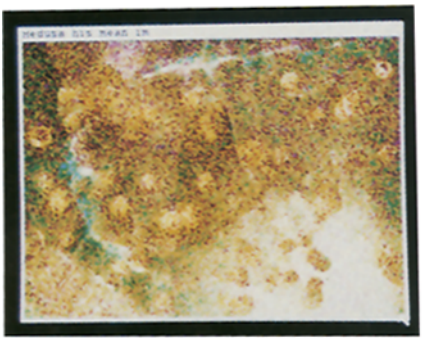

Fig. 23(c). Result of individual mean filter in the HIS space.

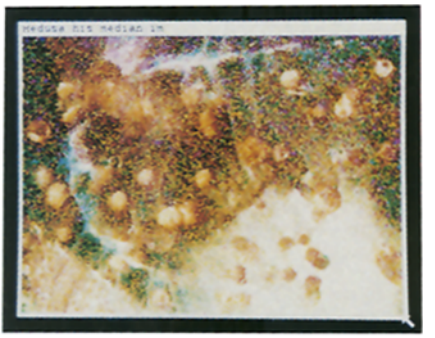

Fig. 23(d). Result of individual median filter in the HIS space.

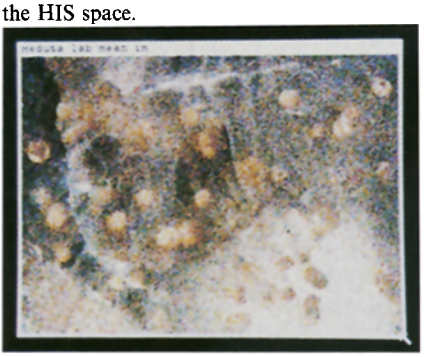

Fig. 23(e). Result of individual mean filter in the $L*a*b*$ space.

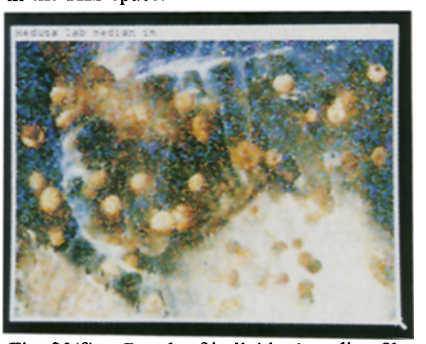

Fig. 23(f). Result of individual median filter in the $L*a*b*$ space.

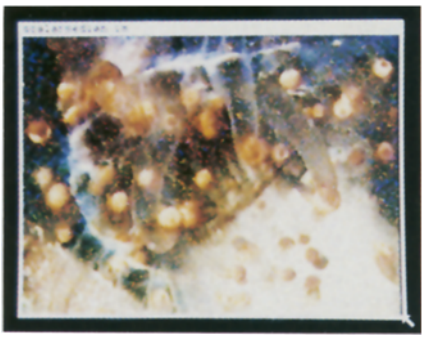

Fig. 24. Result of scalar median filter.

Fig. 25. Result of vector median filter.

Table III. The mean squared errors for the Medusa image

Images	MSE values
Original Image (Fig. 21)	0.00
Noisy Image (Fig. 22)	12162.18
Mean Result (Fig. 23.a)	3393.10
Median Result (Fig. 23.b)	1774.57
Mean Result (Fig. 23.c)	5239.34
Median Result (Fig. 23.d)	2074.27
Mean Result (Fig. 23.e)	3392.97
Median Result (Fig. 23.f)	1783.34
Scalar Median (Fig. 24)	1499.66
Vector Median (Fig. 25)	1563.17

composed. Instead, a vector signal should be the direct input to a filter in which all correlations between vector elements are utilized to produce an optimum output.

Consider a vector signal C corrupted by zero-mean additive random noise. The input data to the filter under consideration is of the form

$$c_i = \hat{c}_i + n_i, \tag{22}$$

where $\hat{c}_i$ are the original signals corrupted by noise n_i; it is assumed that n_i are random vector variables satisfying $E\{n_i\} = 0$. The correlation matrix of the noise n_i, R_n, is defined as

$$R_n = \begin{bmatrix} R_n^{11} & R_n^{12} & R_n^{13} \\ R_n^{21} & R_n^{22} & R_n^{23} \\ R_n^{31} & R_n^{32} & R_n^{33} \end{bmatrix}, \tag{23}$$

where

$$R_n^{ij} = E\{n_i n_j^T\},$$

with $(i, j = 1, 2, 3) \in (r, g, b)$. Because of symmetry, R_n^{ij} is identical to R_n^{ji}. With the windowing operation, the element vectors may be expressed as $c_i = c + n_i$ $(i = 1, 2, \ldots N_w)$ if the window size is small enough such that the original vectors are

constant. Based on the definition of mean vector, one may write the output $\bar{c}$ of a vector mean filter as

$$\bar{c} = \sum_{i=1}^{N_w} a_i c_i, \tag{24}$$

where a_i are the weight coefficients. If we choose

$$a_i = \frac{1}{N_w} \quad (i = 1, 2, \ldots, N_w),$$

we end up with individual mean filtering. The vector median filter, after the input is sorted, may be also considered as a special case in which one of the a_i is 1 with the rest equal to zero. For example: $a_1 = 1$, if a set of vectors as inputs is 'ordered' in an increasing order based on the distance measurement. Here, we intend to find a set of optimum coefficients a_i such that the output $\bar{c}$ is closest to c. The bound of a_i must satisfy the following unbiased condition for the output

$$\sum_{i=1}^{N_w} a_i = 1. \tag{25}$$

The MSE criterion to minimize the difference between c and $\bar{c}$ is

$$\text{error} = \varepsilon = E\{\| c - \bar{c} \|^2\}. \tag{26}$$

The sum

$$\bar{c} = \sum_{i=1}^{N_w} a_i c_i = a_1 c_1 + a_2 c_2 + \ldots + a_{N_w} c_{N_w}$$

is a vector in the subspace of the space defined by c_i. From the projection theorem,

$$\begin{aligned} \varepsilon &= E\left\{ \left\| c - \sum_{i=1}^{N_w} a_i c_i \right\|^2 \right\} \\ &= E\left\{ \left\| c - \sum_{i=1}^{N_w} a_i (c + n_i) \right\|^2 \right\} \\ &= E\left\{ \left\| \sum_{i=1}^{N_w} a_i n_i \right\|^2 \right\}. \end{aligned} \tag{27}$$

This equation is in quadratic form and may be rewritten as

$$\varepsilon = E\{\| AN^T \|^2\}, \tag{28}$$

where A is a coefficient vector

$$A = [a_1 a_2 \ldots a_{n_w}] \tag{29}$$

and

$$N = [n_{i1} n_{i2} \ldots n_{iN_w}]. \tag{30}$$

Then

$$\begin{aligned} \varepsilon &= E\{|AN^T NA^T|\} \\ &= |AR_N A^T|. \end{aligned} \tag{31}$$

Equation (31) establishes the relationship between the MSE and the coefficients. The determination of the a_i, a_2, . . . a_{N_w} which minimize the error is subject to the unbiased condition Equation (25). This becomes a constrained minimization problem. In order to utilize the regular differential techniques, we may form a new unconstrained problem by appending the constraints to with a Lagrange multiplier λ. The new function becomes

$$L(A, \lambda) = |A^T R_N A| + \lambda(1 - eA^T), \tag{32}$$

where, $e = \{1 \quad 1 \ldots 1\}$, and note that

$$\sum_{i=1}^{N_w} a_i = 1 = eA^T. \tag{33}$$

Since the problem defined by Equation (32) is now unconstrained, the first derivative with respect to A yields

$$\frac{\mathrm{d}\, L(a, \lambda)}{\mathrm{d}\, A} = (2|AR_N e^T|). \tag{34}$$

Setting this expression equal to zero and factoring out λ, we obtain the root of the equation

$$\lambda = 2|AR_N e^T|, \tag{35}$$

from where

$$A = \frac{e^T R_n e^T}{R_n e^T}, \tag{36}$$

which determines the coefficients needed to linearly combine the input vectors in order to produce an optimum output.

It is not difficult to see the difference between the Weiner filter solution and the solution derived in Equation (36). To get an optimum result through the Weiner filter, requires a significant amount of prior knowledge on the degradation function, the covariance of the original image R_s and the noise R_N. To obtain an accurate estimate of the covariance, an ensemble of many samples of the ideal image (as random variables) is required. However, in practical applications it is unlikely that there will be an ensemble of ideal image samples (that is a set of prototype images) available for the estimation. Therefore, spatial averages are often used in place of ensemble averages in implementation. Hence, the covariance estimated from the single copy of the degraded image is far from the true covariance of the ideal image which is required by the MSE estimate. For these reasons, it is postulated that the Wiener filters,

in our applications, are no longer optimal because of the lack of accurate prior information.

The solution derived here requires only the estimation of the noise model which the Wiener filters require, too. Moreover, the computation involved compared to the Weiner filters is dramatically reduced. Equation (36) shows that the coefficients may be optimally determined if the noise correlation is known. To obtain the noise correlation, we have to model the noise. Theoretically, we may remove any additive nonimpulsive noise with minimum color artifacts. The practical results, however, may depend on how accurately the noise is modeled.

7. Conclusions

It has been illustrated that monochromatic-like processing fails when applied directly to resolve noise problems in color images.

To overcome noise removal problems in color image processing, two median type filters are studied: scalar and vector filter. Quantitative measurements and subjective comparisons of the outputs have been performed.

A mean type filtering technique is also introduced, incorporating vector signals as inputs.

Results are obtained and compared using, both, real and synthetic images.

References

1. Aus, H.M. *et al.*, Computer color vision, *Proc. 3rd Internat. Conf. Robot Vision and Sensory Controls*, Cambridge, MA pp. 225–229 (1983).
2. Ballard, D.H. and Brown, C.M., *Computer Vision*, Prentice-Hall, Englewood Cliffs (1982).
3. Buchsbaum, G., A spatial processor model for object color perception, *J. Franklin Institute* **310**, 1–26 (1980).
4. Cornsweet, T.N., *Visual Perception*, Academic Press, New York (1970).
5. DeValois, R.L., Color vison mechanisms in the monkey, *J. Gen. Physiol.*, **43**, 115–128 (1960).
6. Durrett, H.J., *Color and the Computer*, Academic Press, New York (1987).
7. Gallagher, N. and Wise, G.L., A theoretical analysis of the properties of median filters, *IEEE Trans. Acoust. Speech, Signal Process.* **ASSP-29**(6) 1136–1141 (1981).
8. Healey, G. and Binford, T.O., The role and use of color in a general vision system, *Proc. ARPA Image Understanding Workshop*, USC (1987).
9. Horn, B.K.P., Extract reproduction of colored images, *Computer Vision, Graphics, and Image Processing* **26**, 135–167 (1984).
10. Huang, T.S., *Two-Dimensional Digital Signal Processing II*, Springer-Verlag, New York (1981).
11. Keil, R.E., Machine vision with color detection, *Proc. 3rd Internat. Conf. Robot Vision and Sensory Controls*, Cambridge, MA, pp. 503–512 (1983).
12. Kender, J., Saturation, hue and color: Calculation, digitization effects, and use, Technical Report, Department of Computer Science, Carnegie-Mellon University, Pittsburgh, PA (1976).
13. MacAdam, D.L., *Color Measurement, Theme and Variations*, Springer-Verlag, New York (1981).
14. Nevatia, R., A color edge detector and its use in scene segmentation, *IEEE Trans. Systems Man Cybernet.* **SMC-7**(11), 820–826 (1977).
15. Ohlander, R., Price, K., and Reddy, D.R., Picture segmentation using a recursive region splitting method, *Computer Graphics and Image Processing* **8**, 313–333 (1978).
16. Ohta, Y., Kanade, T., and Sakai, T., Color information for region segmentation, *Computer Graphics and Image Processing* **13**, 222–241 (1980).

17. Pratt, W.K., *Digital Image Processing*, Wiley, New York (1977).
18. Richard, H., *Image, Object, and Illusion*, Scientific American, Inc. (1974).
19. Scollar, I., Weidner, B., and Huang, T.S., Image enhancement using and the interquartile distance, *Computer Vision, Graphics, and Image Processing* **25**, 236–251 (1984).
20. Stabell, U. and Stabell, B., Effects of rod activity on color threshold, *Vision Research* **16**, 1105–1110 (1976).
21. Zheng, J., Valavanis, K., and Gauch, J., Object extraction using color information in computer vision, *Internat. Conf. Automation, Robotics and Computer Vision, ICARCV '90*, Singapore (1990).
22. Davis, L.S., A survey of edge detection techniques, *Computer Graphics and Image Processing* **4**, (3), 248–270 (1975).
23. Pitas, I. and Venetsanopoulos, A.N., *Nonlinear Digital Filters*, Kluwer Academic Pulishers, Dordrecht (1990).
24. Lee, J., Digital image smoothing and the sigma filter, *Computer Vision, Graphics and Image Processing* **24**, 255–269 (1983).
25. Rosenfeld, A., Image analysis: Problems, progress and prospects, in M. Fischler and O. Firschein, (eds), *Computer Vision,* Morgan Kaufmann, pp. 3–12 (1987).
26. Tukey, J.W., *Exploratory Data Analysis*, Addison-Wesley, Reading, MA (1971).
27. Faugeras, O.D., *Fundamentals in Computer Vision*, Cambridge University Press (1983).