Color Image Processing Using Adaptive Multichannel Filters

Konstantinos N. Plataniotis, *Member, IEEE,* Dimitrios Androutsos, *Student Member, IEEE,*
Sri Vinayagamoorthy, and Anastasios N. Venetsanopoulos, *Fellow, IEEE*

Abstract—New adaptive filters for color image processing are introduced and analyzed in this paper. The proposed adaptive methodology constitutes a unifying and powerful framework for multichannel signal processing. Using the proposed methodology, color image filtering problems are treated from a global viewpoint that readily yields and unifies previous, seemingly unrelated, results. The new filters utilize Bayesian techniques and nonparametric methodologies to adapt to local data in the color image. The principles behind the new filters are explained in detail. Simulation studies indicate that the new filters are computationally attractive and have excellent performance.

I. INTRODUCTION

MULTICHANNEL signal processing has been the subject of extensive research during the last ten years, primarily due to its importance to color image processing. The amount of research published to date indicates a great interest in the areas of color image filtering and analysis. It is widely accepted that color conveys information about the objects in a scene and that this information can be used to further refine the performance of an imaging system. Color images are studied in this paper using a vector approach. The value at each image pixel is represented by a three-channel vector, transforming the color image to a vector field in which each vector's direction and length is related to the pixel's chromatic properties [1]. Being a two-dimensional (2-D), three-channel signal, a color image requires increased computation and storage, as compared to a grey-scale image during processing.

In particular, the most common processing tasks are noise filtering and enhancement, since these are essential functions of any image processing system, regardless of whether the processed image is utilized for visual interpretation or automatic analysis [2].

Consider the following commonly used model for a color image corrupted by additive noise [3]:

$$\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{n} \tag{1}$$

where the notation $\boldsymbol{x}$ is used for simplicity to denote the three-channel uncorrupted image vector located at position (i, j) of the 2-D color image, $\boldsymbol{y}$ is the corresponding noisy vector to be filtered, and $\boldsymbol{n}$ is an additive noise vector at position

(i, j) of the image. In our analysis, it is assumed that the color image vectors are unknown and that the noise vectors are uncorrelated and signal independent at all image locations [3], [4].

The type and degree of noise corruption depends on many factors, such as the measurement apparatus, the signal levels, and the experimental procedure under consideration. The appearance of the noise and its effect on the image is related to its frequency characteristics, which is commonly assumed to be Gaussian. However, the noise encountered in image processing applications cannot always be described in terms of the Gaussian model, rather, it can be characterized in terms of "impulsive" sequences, which occur in the form of short-duration, high-energy spikes attaining large amplitudes with probability higher than that predicted by a Gaussian density model [5].

Numerous filtering techniques have been proposed to date for color image processing. Specifically, nonlinear filters applied to color images have been designed to preserve edges and details, and remove impulsive and Gaussian noise. It has been recognized by many researchers that vector processing of color images is probably a more effective way (as compared to scalar) to filter out noise and to enhance color images. A number of different vector processing filters based on order statistics (OS) have been introduced recently for color image processing [5]–[13]. By implementing a temporal window which moves across the image, the filtered estimate $\hat{\boldsymbol{x}}_l$ at each location index l is typically obtained by processing the noisy image vectors

$$\begin{aligned} \boldsymbol{Y} &= (\boldsymbol{y}_{(l-N)}, \boldsymbol{y}_{(l-N-1)}, \cdots, \boldsymbol{y}_{(l+N)}) \\ &= (\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_n) \end{aligned} \tag{2}$$

that are present in a window W of size $n = 2N + 1$ centered around the image location l. The best known filter from this class is the vector median filter (VMF) developed as a direct extension of the corresponding single-channel median filter [5]. It is well known that the VMF discards impulses and preserves edges. However, its performance in the suppression of Gaussian noise is inferior to that of the arithmetic (linear) mean filter (AMF). If an image is corrupted by additive Gaussian noise and impulsive noise, another member of the *OS* family can be used, namely the α-*trimmed* mean filter (α-TMF), which replaces the noisy input vector-valued signal with a linear combination of the $(n - 2\alpha)$ sample vectors closest to the median value. The trimming operation causes this nonlinear filter to have good performance in the presence of

impulsive noise, while the linear (mean) operation causes the filter to also perform well in the presence of Gaussian noise. Depending on the situation at hand, a different percentage of the inputs can be trimmed out, resulting in an adaptive, data-dependent filter. Recently, a new family of color image filters called vector directional filters (VDF) has been introduced. The new filters use the angle between the image vectors as an ordering criterion [7], [8]. The basic vector directional filter (BVDF) minimizes the sum of angles with all the other vectors within the processing window. The BVDF output can be obtained using the maximum likelihood (ML) principle and it is, in essence, the spherical median of the population [9]. This family also includes the distance directional filter (DDF) [8], which utilizes a hybrid structure to combine the VMF and the VDF. Finally, another member of the family, the generalized vector directional filter (GVDF), combines the directional processing of the BVDF with an adaptive magnitude processing module [9], and thus can be considered as a modified chromaticity trimmed filter, analogous to the α-trimmed mean filter. The GVDF selects a subset of the input vectors and then processes them with the optimal magnitude processing module, depending on the noise characteristics, to generate the filter output. Similar to the VMF and BVDF, which were derived using the ML principle, a number of filters have been derived using the maximum *a posteriori* approach and the Gibbs random field model [10], [11]. These filters can be viewed as weighted median filters [10]. In such a filter, the weights are based on the "clique energies" observed within the neighbor structure. Such designs require *a priori* decisions regarding the form of the "clique" function, which best matches the data. Although good results can be obtained, there is no methodology available for the on-line (adaptive) modification of the assumed clique function.

Other weighted mean filters with adaptively determined coefficients have recently been proposed for robust multi-channel estimation. In [12], an adaptive filter, known as the adaptive nearest neighbor filter (ANNF) utilizes nearest-neighbor techniques to adaptively calculate data dependent weights. Similarly, a heuristically derived multichannel filter, the distance dependent multichannel filter (DDMF), has been proposed in [13]. This adaptive filter uses the inverse of the *Euclidean* distance to weight the vectors in the final output. These adaptive designs deliver good results in a variety of application scenarios, and they are much simpler and more computationally efficient than the weighted medians of [10] and [11]. However, it is not clear, as yet, how to select their distance-based weights in an optimal way.

The plethora of filters available poses some application difficulties. Since most of them are designed to perform well in a specific application, their performance deteriorates rapidly under different operation scenarios. Thus, a nonlinear adaptive filter, which performs equally well in a wide variety of applications, is of great importance. Our goal is to devise computationally efficient and reliable filter structures, which will deliver acceptable results without requiring specific *a priori* information about the signal or the noise characteristics.

This paper introduces such filters using both an adaptive Bayesian inference procedure and nonparametric techniques

to adaptively update the filter parameters in a data-dependent way. For the image degradation model described in (1), the objective of the adaptive filter is to perform as well as the optimum estimator, which has been constructed with the knowledge of the actual noise density function $f(\boldsymbol{n})$. Since, in the case of color image processing some impartial *a priori* knowledge regarding the noise characteristics is available, we aim to derive adaptive filters that are efficient over a certain family of probability densities. This family consists of the Gaussian density, the double exponential density, and all the densities "in between" that naturally arise in image processing applications [14].

The proposed adaptive filter employs a number of different multichannel filters operating in parallel, and Bayesian adaptive procedures are utilized to provide the overall filtered output. The Bayesian approach to the development of multichannel filters for color image processing is discussed in Section II, along with motivation, design characteristics, and implementation details. In Section III, a different adaptive filter is introduced where nonparametric techniques are utilized in order to derive an efficient, data-adaptive nonlinear filter. A generalized multiple filter structure, which can be used to adaptively combine different filters operating in parallel, is also studied in this section. The behavior of the new filters for color image filtering is the subject of Section IV. Extended simulation results are included in order to demonstrate the effectiveness of the proposed adaptive methodologies. Comparisons, in terms of performance with other commonly used filters, are also provided. Finally, conclusions are drawn in Section V.

II. THE BAYESIAN PARAMETRIC APPROACH

In all the adaptive schemes defined in this paper, a "loss function," which depends on the noiseless color vector and its filtered estimate, is used to penalize errors during the filtering procedure. It is natural to assume that if one penalizes estimation errors through a loss function, then the optimum filter is that function of the measurements that minimizes the expected or average loss. If (1) represents the noise degradation model, the optimal estimator that minimizes the average or expected quadratic loss is defined as [15]

$$E(\boldsymbol{x}|\boldsymbol{y}) = \hat{\boldsymbol{x}}(\boldsymbol{y})_{mv} = \int_{-\infty}^{\infty} \boldsymbol{x} f(\boldsymbol{x}|\boldsymbol{y})\, d\boldsymbol{x} \qquad (3)$$

or

$$\hat{\boldsymbol{x}}(\boldsymbol{y})_{mv} = \int_{-\infty}^{\infty} \frac{\boldsymbol{x} f(\boldsymbol{y}, \boldsymbol{x})}{f(\boldsymbol{y})}\, d\boldsymbol{x} = \frac{\int_{-\infty}^{\infty} \boldsymbol{x} f(\boldsymbol{y}, \boldsymbol{x})\, d\boldsymbol{x}}{f(\boldsymbol{y})} \qquad (4)$$

with

$$f(\boldsymbol{y}) = \int_{-\infty}^{\infty} f(\boldsymbol{y}, \boldsymbol{x})\, d\boldsymbol{x}. \qquad (5)$$

As in the case of OS filters, we assume a sliding window of size n is available. Through this, a set of noisy multivariate vectors $\boldsymbol{Y}$, as described in (2), become available. By assuming that the actual image vectors remain constant within the filter window, determination of the $\hat{\boldsymbol{x}}_{mv}$ at the window center

corresponds to the problem of estimating the constant signal from n noisy observations present in the filter window

$$\hat{x}(Y)_{mv} = E(x|Y) = \int_{-\infty}^{\infty} x f(x|Y) \, dx. \tag{6}$$

Central to the solution discussed above is the determination of the probability density function of the image vectors conditioned on the available noisy image data. If this *a posteriori* density function is known, then the optimal estimate, for the performance criterion selected, can be determined. Unfortunately, in a realistic application scenario, such *a priori* knowledge about the process is usually not available. In our adaptive formulation, the requested probability density function is assumed to be of a known functional form but with a set of unknown parameters. This "parent" distribution provides a partial description, where the full knowledge of the underlying phenomenon is achieved through the specific values of the parameters. Given the form of (1) and the assumptions about the noise, knowledge of the actual noise distribution is sufficient for the parametric description of the image vectors conditioned on the observations.

In image processing, a certain family of noise models are often encountered. Thus, we can introduce a symmetric parent distribution, which includes the most commonly encountered noise distributions as special cases [16]. This distribution function can be characterized by a *location* parameter, a *scale* parameter, and a third parameter γ that measures the degree of nonnormality of the distribution.

The multivariate "generalized Gaussian" function, which can be viewed as an extension of the scalar distribution introduced in [17], is defined as

$$f(m|\theta, \sigma, \gamma) = k^M \exp\left(-0.5\beta\left(\frac{|m - \theta|}{\sigma}\right)^{2/1+\gamma}\right) \tag{7}$$

where M is the dimension of the measurement space, σ, the variance, is an $M \times M$ matrix, which can be considered as diagonal with elements σ_c with $c = 1, 2, \cdots, M$, while the rest of the parameters are defined as

$$\beta = \left(\frac{\Gamma(1.5(1+\gamma))}{\Gamma(0.5(1+\gamma))}\right)^{1/1+\gamma}$$

$$k = \left(\frac{(\Gamma(1.5(1+\gamma)))^{0.5}}{(1+\gamma)(\Gamma(0.5(1+\gamma)))^{0.5}}\right)\sigma^{-1}$$

with

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} \, dt$$

and $x > 0$.

This is a two-sided symmetric density that offers great flexibility. By altering the "shape" parameter γ different members of the family can be derived. For example, a value of $\gamma = 0$ results in the Gaussian distribution. If $\gamma = 1$, the double exponential is obtained, and as $\gamma \to -1$ the distribution tends to the rectangular. For $-1 \leq \gamma 1$, intermediate symmetrical distributions can be obtained [16].

Based in this generalized "parent" distribution, an adaptive estimator can be devised utilizing Bayesian inference techniques. Let us assume that (1) describes the image degradation

process and that the noise density function belongs to the generalized family of (7). Assuming that the shape parameter γ and the location and scale parameters of this function are independent, $f(x, \sigma, \gamma) \propto f(x, \sigma) f(\gamma)$, the adaptively filtered result for a "quadratic loss function" is given as

$$E(x|Y) = \frac{\iiint x f(Y|x, \sigma, \gamma) f(x, \sigma) f(\gamma) \, dx \, d\sigma \, d\gamma}{\iiint f(Y|x, \sigma, \gamma) f(x, \sigma) f(\gamma) \, dx \, d\sigma \, d\gamma} \tag{8}$$

$$E(x|Y) = \int \left(\frac{\iint x f(Y|x, \sigma, \gamma) f(x, \sigma) \, dx \, d\sigma}{\iint f(Y|x, \sigma, \gamma) f(x, \sigma) \, dx \, d\sigma}\right)$$
$$\cdot \left(\frac{f(\gamma) f(Y|\gamma)}{f(Y)}\right) d\gamma \tag{9}$$

$$E(x|Y) = \int (E(x|Y, \gamma) f(\gamma|Y)) \, d\gamma \tag{10}$$

with

$$E(x|Y, \gamma) = \int x f(x|Y, \gamma) \, dx. \tag{11}$$

The computational complexity of the adaptive filter depends on the information available about the shape parameter γ. In applications such as image processing, where γ is naturally discrete, the exact realization of the adaptive estimator can be obtain in a computationally efficient way. If the number of shape values is finite $\gamma_1, \cdots, \gamma_\Phi$, then it is possible to obtain the overall adaptive filtered output by combining the conditional filtering results with the Bayesian learning of the unknown shape parameters. The form of the adaptive filter therefore becomes that of a weighted sum

$$E(x|Y) = \sum_{\phi=1}^{\Phi} E(x|Y, \gamma_\phi) f(\gamma_\phi|Y). \tag{12}$$

In cases where a continuous parameter space for the shape parameter is assumed, the *a priori* density function can be quantized using the form

$$f(\gamma) = \sum_{\phi=1}^{\Phi} f(\gamma_\phi) \delta(\gamma - \gamma_\phi)$$

to obtain discrete values. Using the quantized values of the shape parameter, the approximate adaptive algorithm takes the form of (12).

Assume that for a given image location, a window W consisting of n noisy image vectors is available. Let us further assume that based on these Y measurements, intermediate estimates, conditioned on various γ, are available. For example, conditioned on $\gamma = 0$, the mean value of the Y measurements can be considered as the best estimate of the location. Alternatively, if $\gamma = 1$, the median value of the Y set is essentially accepted as the best estimator. In such a scenario, the main objective of the adaptive procedure is the calculation of the posterior densities that arise for different shape parameters. Assuming a uniform reference prior in the

range of $-1 < \gamma \leq 1$ for $f(\gamma)$ the conditional densities are calculated through the following rule:

$$f(\gamma_\phi|\boldsymbol{Y}) = \frac{f(\boldsymbol{y}_n|\gamma_\phi, \boldsymbol{Y}_{n-1})f(\gamma_\phi|\boldsymbol{Y}_{n-1})}{\sum\limits_{\mu=1}^{\Phi} f(\boldsymbol{y}_n|\gamma_\mu, \boldsymbol{Y}_{n-1})f(\gamma_\mu|\boldsymbol{Y}_{n-1})} \qquad (13)$$

with

$$f(\boldsymbol{y}_n|\gamma_\phi, \boldsymbol{Y}_{n-1}) = k^M \exp\left(-0.5\left(\frac{\boldsymbol{y} - \hat{\boldsymbol{x}}_\phi}{\boldsymbol{\sigma}_\phi}\right)^{2/1+\gamma}\right) \tag{14}$$

where

$$\boldsymbol{Y} = (\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_{n-1}, \boldsymbol{y}_n)$$
$$\boldsymbol{Y}_{n-1} = (\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_{n-1})$$

are the observations obtained from the window and $\hat{\boldsymbol{x}}_\phi$ is the conditional filtered result for the image vector at the window center using a specific value of the shape parameter $\gamma = \gamma_\phi$.

The above result was obtained using Baye's rule

$$f(\gamma_\phi|\boldsymbol{Y}) = \frac{f(\gamma_\phi, \boldsymbol{Y})}{f(\boldsymbol{Y})} = \frac{f(\gamma_\phi, \boldsymbol{y}_n, \boldsymbol{Y}_{n-1})}{f(\boldsymbol{Y}|\boldsymbol{Y}_{n-1})f(\boldsymbol{Y}_{n-1})}$$
$$= \frac{f(\gamma_\phi, \boldsymbol{y}_n|\boldsymbol{Y}_{n-1})}{\sum\limits_{\mu=1}^{\Phi} f(\gamma_\mu, \boldsymbol{y}_n|\boldsymbol{Y}_{n-1})}. \qquad (15)$$

Further application of Baye's rule results with

$$f(\gamma_\phi, \boldsymbol{y}_n|\boldsymbol{Y}_{n-1})$$
$$= \frac{f(\boldsymbol{y}_n, \gamma_\phi, \boldsymbol{Y}_{n-1})}{f(\boldsymbol{Y}_{n-1})} = \frac{f(\boldsymbol{y}_n|\gamma_\phi, \boldsymbol{Y}_{n-1})f(\gamma_\phi, \boldsymbol{Y}_{n-1})}{f(\boldsymbol{Y}_{n-1})} \tag{16}$$

or

$$f(\gamma_\phi, \boldsymbol{y}_n|\boldsymbol{Y}_{n-1}) = f(\boldsymbol{y}_n|\gamma_\phi, \boldsymbol{Y}_{n-1})\left(\frac{f(\gamma_\phi, \boldsymbol{Y}_{n-1})}{f(\boldsymbol{Y}_{n-1})}\right)$$
$$= f(\boldsymbol{y}_n|\gamma_\phi, \boldsymbol{Y}_{n-1})f(\gamma_\phi|\boldsymbol{Y}_{n-1}). \qquad (17)$$

To complete the adaptive determination of the *a posteriori* density $f(\gamma_\phi|\boldsymbol{Y})$ in (13) the predictive density $f(\boldsymbol{y}_n|\gamma_\phi, \boldsymbol{Y}_{n-1})$ must be computed. From (1)

$$f(\boldsymbol{y}_n|\gamma_\phi, \boldsymbol{Y}_{n-1}) = f(\boldsymbol{y}_n|\gamma_\phi, \boldsymbol{x}) = f_{\boldsymbol{n}|\boldsymbol{x}}(\boldsymbol{y}_n - \boldsymbol{x}|\gamma_\phi)$$
$$= f_{\boldsymbol{n}}(\boldsymbol{y}_n - \boldsymbol{x}|\gamma_\phi) \qquad (18)$$

where $f_{\boldsymbol{n}|\boldsymbol{x}}(\cdot)$ denotes the conditional probability density function (pdf) of $\boldsymbol{n}$ given $\boldsymbol{x}$ and $f_{\boldsymbol{n}|\boldsymbol{x}}(\cdot) = f_{\boldsymbol{n}}(\cdot)$ when $\boldsymbol{n}$ and $\boldsymbol{x}$ are independent. Thus, the density $f(\boldsymbol{y}_n|\gamma_\phi, \boldsymbol{Y}_{n-1})$ can be considered to be generalized Gaussian with shape parameter γ_ϕ and location estimate the conditional filter output.

The Bayesian inference procedure described above allows for the selection of the appropriate density from the family of densities considered. If the densities corresponding to the different shape values assumed are representative of the class of densities encountered in image processing applications, then the Bayesian procedure should provide good results regardless

of the underlying density, resulting in a robust adaptive estimation procedure.

The adaptive filter described in this section can be viewed as a linear combination of specified, elemental filtered values. The weights in the adaptive design are nonlinear functions of the difference between the measurement vector and the elemental filtered values determined by conditioning on various γ. In this context, the Bayesian adaptive filter can be viewed as a generalization of radial basis neural networks [18] or fuzzy basis functions networks [19].

If it is desired, the minimum mean square error (MMSE) of the unknown scalar shape parameter can be determined as

$$E(\gamma|\boldsymbol{Y}) = \hat{\gamma}_{\mathrm{mmse}}(\boldsymbol{Y}) = \sum_{\phi=1}^{\Phi} \gamma f(\gamma_\phi|\boldsymbol{Y})$$

with the error in the shape parameter estimation calculated through

$$E((\gamma - \hat{\gamma}_{\mathrm{mmse}}(\boldsymbol{Y}))^2|\boldsymbol{Y}) = \sum_{\phi=1}^{\Phi}(\gamma - \hat{\gamma}_{\mathrm{mmse}}(\boldsymbol{Y}))^2 f(\gamma_\phi|\boldsymbol{Y}). \tag{19}$$

In a similar manner, the maximum *a posteriori* likelihood estimate of the shape parameter $\hat{\gamma}_{\mathrm{map}}(\boldsymbol{Y}) = \hat{\gamma}_\mu(\boldsymbol{Y})$ can be obtained through the adaptive filter.

The following comments can be made regarding the adaptive filter.

1) The adaptive filter of (12) is optimum in the Bayes sense every time it is used inside the window and its optimality is independent of the convergence. The weights that regulate the contribution of the elemental filters are not derived heuristically. Rather, the weights are determined through Bayes theorem using the assumptions on the noise density functions. The adaptive filter weights are dependent on the local image information and thus, as the filter window moves from one pixel to the next, a different adaptive filter output is obtained.

2) Through the adaptive design, the problem of determining the appropriate distribution for the noise is transformed into the problem of combining a collection of admissible distributions. This constitutes a problem of considerably reduced complexity, since specific noise models, such as additive Gaussian noise, impulsive noise, or a combination of both are often encountered in image processing applications.

3) The proposed adaptive design is a scalable one. The designer controls the complexity of the procedure by determining the number and form of the individual filters. Depending on the problem specification and the computational constraints imposed by the design, an appropriate number of elemental filters can be selected. The filter requires no *a priori* training signals or test statistics and its parallel structure makes it suitable for real-time image applications.

The proposed Bayesian adaptive procedure is simple, computationally efficient, easy-to-use, and reasonably robust. In the proposed approach, the posterior probabilities are more important than the manner in which the designer can obtain the elemental estimates that are used in the procedure. Different methodologies can be utilized to obtain these estimates. Filters derived using the ML principle, (e.g. VMF, BVDF), robust estimators, (e.g. α-TMF), and estimators based on adaptive designs, such as the ANNF or the DDMF, can all be used to provide these needed elemental estimates.

From the large number of filters that can be designed using the adaptive procedure, a filter of great practical importance is the one that combines a VMF with an AMF. Extensive experimentation in the past has proven that, in the homogeneous regions of the image, a mean filter is probably the most suitable estimator, whereas in areas where edges or fine details are present, a median filter is preferable. Through the adaptive design in (12), these two filters can be combined. By using local image information in the current processing window, such an adaptive filter, which we call BFMA, can switch between the two elemental filters in a data dependent manner, offering enhanced filtering performance.

III. THE NONPARAMETRIC APPROACH

The adaptive formulation presented in the previous section was based on the assumption that a certain class of densities can be used to describe the noise corrupting color images. Thus, a Bayesian adaptive procedure has been utilized to determine on line the unknown parameters that are used to describe the noise density function. However, in a more general formulation, we may also not know the functional form of the noise density. In such a case, the densities involved in the derivation of the optimal estimator of (6) can not be determined through a parametric technique, such as the one described in Section II. Rather, we have to estimate them from available sample observations using a nonparametric technique.

Among the plethora of the different nonparametric schemes, we adopt the kernel estimator. The notation of nonparametric estimation remains relatively unknown, therefore a brief overview is included.

If the objective is the nonparametric determination of an unknown multivariate density $f(z)$ from a set of independent samples $\mathbf{Z} = z_1, z_2, \cdots, z_n$ drawn from the unknown underlying density, the form of a data-adaptive nonparametric kernel estimator is

$$\hat{f}(\mathbf{z}) = (n^{-1}) \sum_{l=1}^{n} (h_l)^{-M} K\left(\frac{\mathbf{z} - \mathbf{z}_l}{h_l}\right) \qquad (20)$$

where $z_l \in \mathbf{R}^M$, M is the dimensionality of the measurement space ($M = 3$ for color images), $K: \mathbf{R}^M \mapsto \mathbf{R}^1$ is a function centered at zero that integrates to one and h_l is the smoothing term [20]–[23].

The form for the data-dependent smoothing parameter is of great importance for the nonparametric estimator. To this end, a new smoothing factor suitable for multichannel estimation is proposed in this paper. For the sample point defined in (20), we define a smoothing factor that is a function of the aggregate distance between the local observation under consideration and all the other vectors inside the $\mathbf{Z}$ set, excluding the point at which the density is evaluated. The smoothing parameter is therefore given by

$$h_l = n^{-k/M} A_l = n^{-k/M} \left(\sum_{j=1}^{n} |z_j - z_l|\right) \qquad (21)$$

where $z_j \neq z_l$ for $\forall z_j, j = 1, 2, \cdots, n, |z_j - z_l|$ is the absolute distance (L_1 metric) between the two vectors, and k is a parameter to be determined. The resulting variable kernel estimator exhibits local smoothing that depends both on the point at which the density estimate is taken and information local to each sample observation in the $\mathbf{Z}$ set.

In addition to the smoothing parameter discussed above, the form of the kernel selected also affects the result. Usually, positive kernels are selected for the density approximation. The most common choices are kernels from symmetric distribution functions, such as the Gaussian or the double exponential. For the simulation studies reported in this paper, the multivariate exponential kernel $K(z) = \exp(-|z|)$ and the multivariate Gaussian kernel $K(z) = \exp(-0.5z^\tau z)$ were selected [22].

As for any estimator, the behavior of the nonparametric estimator of (20) is determined through the study of its statistical properties. Certain restrictions should apply to the design parameters, such as the smoothing factor, in order to obtain an asymptotically unbiased and consistent estimator. According to the analysis introduced in [22], if the conditions $(\lim_{n\to\infty}(nh_l^{2M}(n)) = \infty$ (asymptotic consistency), $(\lim_{n\to\infty}(nh_l^M(n)) = \infty$ (uniform consistency), and $(\lim_{n\to\infty}(h_l^M(n)) = 0$ (asymptotic unbiasedness) are satisfied, then $\hat{f}(z)$ becomes an asymptotically unbiased and consistent estimate of $f(z)$. The multiplier $n^{-k/M}$ in (21) with $(0.5 > k > 0)$ guarantees the satisfaction of the conditions for an asymptotically unbiased and consistent estimator [22]. The selection of the A_l for the same design parameter does not affect the asymptotic properties of the estimator in (20). However, for a finite number of samples, as in our case, the function A_l is the dominant parameter that determines the performance of the nonparametric estimator.

After this brief introduction let us return to the problem of nonparametric evaluation of the densities involved in the derivation of the optimal estimator in (6). This time, no assumption regarding the functional form of the noise present in the image is made.

It is only assumed that n pairs of image vectors $(\mathbf{x}_l, \mathbf{y}_l), l = 1, 2, \cdots, n$ are available through a sliding window of length n centered around the noisy observation $\mathbf{y}$. Based on this sample, the densities $f(\mathbf{y})$ and $f(\mathbf{y}, \mathbf{x})$ will be approximated using sample point adaptive nonparametric kernel estimators.

Our first task is to approximate the joint density $f(\mathbf{y}, \mathbf{x})$. As a nonparametric density approximation we may choose the following:

$$\hat{f}(\mathbf{x}, \mathbf{y}) = n^{-1} \sum_{l=1}^{n} (h_{lx})^M (h_{ly})^{-M}$$
$$\cdot K\left(\left(\frac{(\mathbf{x} - \mathbf{x}_l)}{h_{lx}}\right), \left(\frac{(\mathbf{y} - \mathbf{y}_l)}{h_{ly}}\right)\right). \qquad (22)$$

Assuming a product kernel estimator [24], the nonparametric approximation of the joint density $f(\boldsymbol{y}, \boldsymbol{x})$ has as follows:

$$\hat{f}(\boldsymbol{x}, \boldsymbol{y}) = n^{-1} \sum_{l=1}^{n} (h_{lx})^{-M} (h_{ly})^{-M}$$
$$\cdot K\left(\frac{(\boldsymbol{x} - \boldsymbol{x}_l)}{h_{lx}}\right) K\left(\frac{(\boldsymbol{y} - \boldsymbol{y}_l)}{h_{ly}}\right). \qquad (23)$$

The marginal density $f(\boldsymbol{y})$ in the denominator of (4) can then be approximated using the results in (23) as follows:

$$\int \hat{f}(\boldsymbol{y}, \boldsymbol{x}) \, d\boldsymbol{x} = n^{-1} \sum_{l=1}^{n} (h_{ly})^{-M} K\left(\frac{(\boldsymbol{y} - \boldsymbol{y}_l)}{h_{ly}}\right)$$
$$\cdot \left(\int (h_{lx})^{-M} K\left(\frac{(\boldsymbol{x} - \boldsymbol{x}_l)}{h_{lx}}\right) d\boldsymbol{x}\right)$$
$$= n^{-1} \sum_{l=1}^{n} (h_{ly})^{-M} K\left(\frac{(\boldsymbol{y} - \boldsymbol{y}_l)}{h_{ly}}\right) \qquad (24)$$

since $\int K(\boldsymbol{z}) \, d\boldsymbol{z} = 1$ assuming that the kernel results from a real density.

The determination of the numerator in (4) is now feasible. The assumption that

$$\int z_1^1 \cdots z_M^1 K(\boldsymbol{z}) \, d\boldsymbol{z} = 0$$

implies [24] that

$$\int \boldsymbol{x} K(\boldsymbol{x} - \boldsymbol{x}_l) \, d\boldsymbol{x} = \boldsymbol{x}_l. \qquad (25)$$

Thus, the numerator of (6) becomes

$$\int \boldsymbol{x} \hat{f}(\boldsymbol{y}, \boldsymbol{x}) \, d\boldsymbol{x} = n^{-1} \sum_{l=1}^{n} \boldsymbol{x}_l (h_{ly})^{-M} K\left(\frac{(\boldsymbol{y} - \boldsymbol{y}_l)}{h_{ly}}\right). \quad (26)$$

Utilizing (22)–(26), the sample point adaptive nonparametric estimator of $\hat{\boldsymbol{x}}(\boldsymbol{y})$ can be defined as

$$\hat{\boldsymbol{x}}(\boldsymbol{y})_{np} = \int_{-\infty}^{\infty} \frac{\boldsymbol{x} \hat{f}(\boldsymbol{x}, \boldsymbol{y})}{\hat{f}(\boldsymbol{y})} \, d\boldsymbol{x} = \frac{\int_{-\infty}^{\infty} \boldsymbol{x} \hat{f}(\boldsymbol{x}, \boldsymbol{y}) \, d\boldsymbol{x}}{\int_{-\infty}^{\infty} \hat{f}(\boldsymbol{x}, \boldsymbol{y}) \, d\boldsymbol{x}} \qquad (27)$$

$$\hat{\boldsymbol{x}}(\boldsymbol{y})_{np} = \frac{\sum_{l=1}^{n} \boldsymbol{x}_l \left((n^{-1}) h_l^{-M} K\left(\frac{\boldsymbol{y} - \boldsymbol{y}_l}{h_l}\right)\right)}{\sum_{l=1}^{n} \left((n^{-1}) h_l^{-M} K\left(\frac{\boldsymbol{y} - \boldsymbol{y}_l}{h_l}\right)\right)} \qquad (28)$$

$$\hat{\boldsymbol{x}}(\boldsymbol{y})_{np} = \sum_{l=1}^{n} \boldsymbol{x}_l \left(\frac{h_l^{-M} K\left(\frac{\boldsymbol{y} - \boldsymbol{y}_l}{h_l}\right)}{\sum_{l=1}^{n} h_l^{-M} K\left(\frac{\boldsymbol{y} - \boldsymbol{y}_l}{h_l}\right)}\right) = \sum_{l=1}^{n} \boldsymbol{x}_l w_l(\boldsymbol{y})$$
$$(29)$$

where $\boldsymbol{y}_l \in W$ and $w_l(\boldsymbol{y})$ is a weighting function defined in the interval [0, 1].

From (29), it can be seen that the nonparametric estimator, often called the Nadaraya–Watson estimator [25], [26], is given as a weighted average of the samples in the window selected. The inputs in the mixture are the noise-free vectors

$\boldsymbol{x}_l$. This estimator is linear with respect to the $\boldsymbol{x}_l$, and can be considered therefore as a linear smoother. The basis functions, on the other hand, determined by the kernel function $K(\cdot)$ and the smoothing parameter $h(\cdot)$, can be either linear or nonlinear on the noisy measurements $\boldsymbol{y}_l$. It is easy to recognize the similarity between the Bayesian adaptive parametric filter discussed in Section II with the Nadaraya–Watson estimator. The Bayesian adaptive filter is also a linear smoother with respect to a function of the $\boldsymbol{x}_l$ (the elemental filtered results) and utilizes nonlinear basis functions, which are determined by the unknown "shape" (γ) parameter from the generalized "parent" distribution assumed.

Although the existence of a consistent estimate in mean square has been proven, there are no *a priori* guidelines on the selection of design parameters, such as the smoothing vector or the kernel, on the basis of a finite set of data. Smoothing factors, other than the aggregated distance introduced here, such as the minimum distance or the maximum distance between the $\boldsymbol{y}_l$ and the kth nearest neighbors, constitute valid solutions and can be used instead [27].

The adjustable parameters of the proposed filter are $\boldsymbol{x}, \boldsymbol{y}, K(\cdot)$ and $h(\cdot)$. The degree of the smoothness is controlled by the smoothing factor $h(\cdot)$. It can easily be seen that by appropriately modifying the smoothing factor we can force the nonparametric estimator to match a given sample arbitrarily close. To accomplish this, we modify the kernel by adjusting through an exponent the effect of the smoothing parameter $h(\cdot)$. In this case, the form of the estimator is as follows:

$$\hat{\boldsymbol{x}}(\boldsymbol{y}) = \sum_{l=1}^{n} \boldsymbol{x}_l \left(\frac{h_l^{-M} K\left(\frac{\boldsymbol{y} - \boldsymbol{y}_l}{h_l^r}\right)}{\sum_{l=1}^{n} h_l^{-M} K\left(\frac{\boldsymbol{y} - \boldsymbol{y}_l}{h_l^r}\right)}\right) \qquad (30)$$

where the parameter r regulates the smoothness of the kernel. Since the nonparametric filter is a regression estimator, which provides a smooth interpolation among the observed vectors inside the processing window, the r parameter can provide the required balance between smoothing and detail preservation. Because r is a one-dimensional (1-D) parameter, it is usually not difficult to determine an appropriate value for a practical application. By increasing the value of the r, we can force the nonparametric estimator to approximate arbitrary close any one of the vectors inside the filtering window. To this end, let us suppose that a nonparametric estimator with given value $r = r^*$ exists, given the available input set $\boldsymbol{Y}$. Then the following relation holds:

$$\hat{\boldsymbol{x}}(\boldsymbol{y}) = \frac{\boldsymbol{x}_j + \sum_{l=1 \atop l \neq j}^{n} \boldsymbol{x}_l \left(h_l^{-M} K\left(\frac{\boldsymbol{y} - \boldsymbol{y}_l}{h_l^r}\right)\right)}{1 + \sum_{l=1 \atop l \neq j}^{n} \left(h_l^{-M} K\left(\frac{\boldsymbol{y} - \boldsymbol{y}_l}{h_l^r}\right)\right)} \qquad (31)$$

with $l = 1, 2, \cdots, n$. Assuming that $\boldsymbol{x}_j \neq \boldsymbol{x}_l$ for $j \neq l$. Then for arbitrary $\epsilon > 0$ and any l, j with $l = 1, 2, \cdots, n$ and $j \neq l$ we can force $K(\boldsymbol{y} - \boldsymbol{y}_l / h_l^r) < \epsilon$ since by properly choosing a value $r^* = 1/r_v$ the kernel $K(\boldsymbol{y} - \boldsymbol{y}_l / h_l^{r^*}) \mapsto 0$ as $r_v \mapsto 0$ if

Number	Noise Model
1	Gaussian ($\sigma = 30$)
2	impulsive (4%)
3	Gaussian ($\sigma = 7.5$), impulsive (2%)
4	Gaussian ($\sigma = 15$), impulsive (2%)
5	Gaussian ($\sigma = 30$), impulsive (4%)

$\boldsymbol{y} \neq \boldsymbol{y}_l$. Thus, we can conclude that there exists some value of r such that the nonparametric regressor approaches arbitrary close to an existing vector.

To obtain the final estimate, we must assume that, in the absence of noise, the actual image vectors $\boldsymbol{x}_l$ are available. As is the case for the adaptive/trainable filters, a training record can be obtained in some cases during some calibration procedure in a controlled environment. In a real-time image processing application, however, that is not always possible. Therefore, alternative suboptimal solutions are introduced. In a first approach, each vectors $\boldsymbol{x}_l$ in (29) is replaced with its noisy measurement $\boldsymbol{y}_l$. The resulting suboptimal estimator, called adaptive multichannel nonparametric filter (hereafter AMNF), is solely based on the available noisy vectors and the form of the data-adaptive kernel selected for the density approximations. Thus, the AMNF form is as follows:

$$\hat{\boldsymbol{x}}(\boldsymbol{y})_{\text{AMNF}} = \sum_{l=1}^{n} \boldsymbol{y}_l \left(\frac{h_l^{-M} K\left(\dfrac{\boldsymbol{y} - \boldsymbol{y}_l}{h_l}\right)}{\displaystyle\sum_{l=1}^{n} h_l^{-M} K\left(\dfrac{\boldsymbol{y} - \boldsymbol{y}_l}{h_l}\right)} \right). \tag{32}$$

A different form of the adaptive nonparametric estimator can be obtained if a reference vector $\boldsymbol{x}_{lr}$ is used instead of the actual color vector $\boldsymbol{x}_l$ in (29). A robust estimate of the location, usually evaluated in a smaller subset of the input set, is utilized instead of the $\boldsymbol{x}_l$. Usually the median is the preferable choice, since it smoothes out impulsive noise and preserves edges and details. However, unlike scalars, the most central vector in a set of vectors can be defined in more than one way. Thus, the VMF or the marginal median filter (MAMF) operating in a 3 × 3 window centered around the current pixel can be used to provide the requested reliable reference. In this paper, the VMF evaluated in a 3 × 3 window was selected to provide the reference vector. The new adaptive multichannel nonparametric filter, (hereafter AMNF2), has the following form:

$$\hat{\boldsymbol{x}}(\boldsymbol{y})_{\text{AMNF}} = \sum_{l=1}^{n} \boldsymbol{x}_l^{\text{VM}} \left(\frac{h_l^{-M} K\left(\dfrac{\boldsymbol{y} - \boldsymbol{y}_l}{h_l}\right)}{\displaystyle\sum_{l=1}^{n} h_l^{-M} K\left(\dfrac{\boldsymbol{y} - \boldsymbol{y}_l}{h_l}\right)} \right). \tag{33}$$

The AMNF2 can be viewed as a double-window (DW), two-stage estimator. First, the original image is filtered by a multichannel median filter in a small processing window in order to reject possible outliers, and then the adaptive filter of (33) is utilized to provide the final filtered output. The AMNF2 filter can be viewed as an extension to the multichannel case of the DW filtering structures extensively used for gray-scale image processing. As in gray-scale processing, with this adaptive filter we can distinguish between two operators: i) the computation of the median in a smaller window, and ii) the adaptive averaging in a second processing window.

In the derivation of the data-adaptive nonparametric estimator presented in (29), (32), and (33), a number of design parameters have been introduced. They are as follows:

- the window size, and therefore the number of noisy measurement vectors, available for the evaluation of the approximate density;
- the form of the smoothing factor h_l where decisions about the multiplier and the distance measure utilized can greatly affect the performance of the density estimator;
- the type of kernel used in (29);
- the vectors used instead of the actual, unavailable color vectors $\boldsymbol{x}_l$ in the derivation of (29).

All these elements affect the filtering process, since they determine the output of the estimator. In an adaptive formulation, we can define $\boldsymbol{\Theta}$ as the parameter vector, which is the abstract representation of all elements listed above. It is not necessary that all these elements be treated as parameters in a specific design. Problem specifications and design objectives can be used to determine the elements included in the parameter vector $\boldsymbol{\Theta}$.

By varying the different parameters on the design of the nonparametric kernel, different results $\hat{\boldsymbol{x}}_{\boldsymbol{\Theta}}(Y) = \hat{m}(\boldsymbol{x})$ can be obtained. Let us suppose that $\hat{m}(\boldsymbol{x}_i), i = 1, 2, \cdots, P$ are different nonparametric estimators, all based on the same sample record $Y = (\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_n)$ but possibly with different kernels $K_1, K_2, \cdots, K_p$ and different smoothing factors $h_1, h_2, \cdots, h_p$. An overall estimator based on these values can be obtained as the expected value $\hat{\boldsymbol{x}}_{np} = E(\hat{m}(\boldsymbol{x})|\boldsymbol{y})$ calculated over the given nonparametric values determined by the different techniques.

Assuming that the different estimated values $\hat{m}(\boldsymbol{x})$ are available and that they are related to the observed sample through the model

$$\boldsymbol{y} = \hat{m}(\boldsymbol{x}) + \boldsymbol{\xi} \tag{34}$$

with $\boldsymbol{\xi}$ additive corruption noise, we can claim that the minimization of the expected error leads to a solution for $\hat{\boldsymbol{x}}_{np}$ as

$$\hat{\boldsymbol{x}}_{np} = \frac{\displaystyle\sum_{i=1}^{P} \hat{m}_i(\boldsymbol{x}) f_{\xi}(\boldsymbol{y} - \hat{m}_i(\boldsymbol{x}))}{\displaystyle\sum_{i=1}^{P} f_{\xi}(\boldsymbol{y} - \hat{m}(\boldsymbol{x}))} = \sum_{i=1}^{P} \hat{m}_i(\boldsymbol{x}) w_{npi} \tag{35}$$

with

$$w_{npi} = \frac{f_{\boldsymbol{\xi}}(\boldsymbol{y} - \hat{m}_i(\boldsymbol{x}))}{\displaystyle\sum_{j=1}^{P} f_{\boldsymbol{\xi}}(\boldsymbol{y} - \hat{m}_j(\boldsymbol{x}))}. \tag{36}$$

TABLE II
FILTERS COMPARED

Notation	Filter	Reference
AMF	Arithmetic (Linear) Mean Filter	[2]
VMF	Vector Median Filter	[6]
BVDF	Basic Vector Directional Filter	[1], [7]
GVDF	Generalized Vector Directional Filter	[7], [9]
	with an α-trimmed magnitude module, ($\alpha = 1.5$)	
DDF	Directional-Distance Filter	[8]
ANNF	Adaptive Nearest Neighbor Filter	[12]
DDMF	Distance Dependent Multichannel Filter	[13]
α-TMF	α-trimmed Mean Filter ($\alpha = 1.5$)	[5]
AMNFG	Adaptive Non-parametric Filter, eq. 32, M=3, k=0.33, Gaussian kernel	
AMNFE	Adaptive Non-parametric Filter, eq. 32, M=3, k=0.33, Exponential kernel	
AMNFE2	AMNF2 filter with an Exponential kernel, eq. 33, VMF reference point, M=3, k=0.33	
AMNFG2	AMNF2 filter with a Gaussian kernel, eq. 33, VMF reference point, M=3, k=0.33	
BFMA	Bayesian Adaptive Filter (Median & Arithmetic Mean) subfilters	
MAMNFEG	Multiple Non-parametric Filter, 2 subfilters with Gaussian & Exponential kernels	
MAMNFEG	Multiple AMNF2 Filters, 2 subfilters with Gaussian & Exponential kernels, eq. 35	
MAMNF35	Multiple Non-parametric Filter, 2 subfilters with different window sizes, eq. 35	

TABLE III
EXISTING FILTERS: NMSE 10^{-2} FOR THE LENA IMAGE, 3×3

Noise	$BVDF$	$GVDF$	DDF	VMF	AMF	$ANNF$	$DDMF$	$\alpha - TMF$
1	2.8962	1.46	1.524	1.60	0.6963	0.851	0.833	1.272
2	0.3848	0.30	0.3255	0.19	0.8186	0.261	0.221	0.186
3	0.463	0.341	0.439	0.295	0.527	0.2532	0.285	0.288
4	1.1354	0.6238	0.6483	0.5404	0.6160	0.3878	0.3927	0.467
5	3.8515	1.982	2.1646	1.6791	1.298	1.086	1.1044	1.39

To calculate the exact value of the multiple nonparametric estimator, the function $f_\xi(\cdot)$ must be evaluated. Since it is generally unknown, we approximate it in a nonparametric fashion based on the set of the elemental values $\hat{m}(x)$ available. If P elemental estimates $\hat{m}_i(x)$ are available, with $i = 1, 2, \cdots, P$ we introduce the nominal parameter $\xi_i = y - \hat{m}_i(x)$. Therefore, our objective is the nonparametric evaluation of the density $f_\xi(\cdot)$ using the set of the available data points $\Xi = \xi_1, \xi_2, \cdots, \xi_P$. The approximation task can be carried out by using any standard nonparametric approach, such as the different kernel estimators discussed in (29)–(33). For the simulation studies discussed in the next section, we decided to utilize our sample point adaptive kernel estimator of (20). Thus, the following estimate of the density $f_\xi(y - \hat{m}_i(x))$ is used:

$$\hat{f}_\xi(y - \hat{m}_i(x)) = \hat{f}_\xi(\xi) = (P^{-1}) \sum_{l=1}^{P} (h_l)^{-M} K\left(\frac{\xi - \xi_l}{h_l}\right)$$

$$(37)$$

with the smoothing parameter calculated as

$$h_l = P^{-k/M} A_l = P^{-k/M} \left(\sum_{j=1}^{P} |\xi_j - \xi_l| \right), \quad (38)$$

where $\xi_j \neq \xi_l$ for $\forall \xi_j$ $j = 1, 2, \cdots, P$, and $|\xi_j - \xi_l|$ is the absolute distance (L_1 metric) between the two vectors.

From (37), we can claim that $\hat{f}_\xi(\xi)$ integrates to one, given the form of the approximation and the fact that the kernel $K(\cdot)$ results from a real density. Thus, the set of weights w_{npi} has the following properties: i) Each weight is a positive number, $w_{npi} \geq 0$, and ii) the summation of all the weights is equal to one, $\Sigma_{j=1}^{P} w_{npi} = 1$, and can be interpreted as posterior probabilities used to incorporate prior information concerning local smoothness. Thus, each weight in (35) regulates the contribution of the associated filter by its posterior component density.

The following comments can be made regarding the multiple nonparametric filter.

TABLE IV
NEW FILTERS: NMSE 10^{-2} FOR THE LENA IMAGE, 3×3

Noise	$BFMA$	$AMNFE$	$AMNFE2$	$AMNFG$	$AMNFG2$	$MAMNFEG$	$MAMNFEG2$	$MAMNF35$
1	0.7286	0.87	0.565	0.8417	0.546	0.8534	0.546	0.6196
2	0.3067	0.1836	0.171	0.2006	0.221	0.1915	0.208	0.2821
3	0.229	0.221	0.232	0.206	0.231	0.214	0.232	0.242
4	0.4284	0.3746	0.302	0.3578	0.312	0.3660	0.314	0.3838
5	1.0718	1.0438	0.699	1.007	0.672	1.0234	0.679	0.7152

TABLE V
EXISTING FILTERS: NMSE 10^{-2} FOR THE LENA IMAGE, 5×5

Noise	$BVDF$	$GVDF$	DDF	VMF	AMF	$ANNF$	$DDMF$	$\alpha - TMF$
1	2.8819	1.08	1.0242	1.17	0.5994	0.626	0.639	0.664
2	0.7318	0.54	0.5126	0.58	0.6656	0.421	0.369	0.312
3	0.685	0.401	0.553	0.503	0.536	0.377	0.332	0.386
4	1.3557	0.459	0.6913	0.5172	0.5702	0.436	0.418	0.409
5	4.1237	1.1044	1.3048	1.0377	0.8896	0.758	0.814	0.826

- The general form of the filter is given as a linear combination of nonlinear basis functions. The weights in the above mixture are the elemental filtered values obtained by the different nonparametric estimators applied to the problem. The nonlinear basis function is determined by the form of the approximate density $\hat{f}_\xi$, and can take many different forms, such as Gaussian, exponential, or triangular. It is not hard to see that in the case of a Gaussian kernel the multiple estimator of (35) can be viewed as a radial basis function (RBF) neural network.

- The adaptive procedure can be used to combine a variety of nonparametric estimators, each one of them developed for a different value set of the parameter vector Θ. For example, such a structure can be used to combine elemental nonparametric filters derived for different window sizes W. The number of color vector samples utilized in the development of the nonparametric estimator depends on the window W centered around the pixel under consideration. Usually a square 3×3 or 5×5 window is selected. However, such a decision affects the filter's performance. In smooth areas or when Gaussian noise is anticipated, a larger window (e.g., 5×5) is preferable. On the other hand, near edges or when impulsive noise is assumed a smaller window (usually 3×3 window) is more appropriate. An adaptive filter that utilizes elemental filters with different window sizes (hereafter, MAMNF35), is most probably a better choice in an unknown or mixed noise environment. Using the same approach other practical adaptive filters, such as the MAMNFEG, which utilizes two elemental nonparametric filters with an exponential and a Gaussian kernel, respectively, can be devised. Due to the specific form of the kernel, it is anticipated that a nonparametric filter with a Gaussian kernel is probably a better choice for Gaussian noise smoothing. Similarly, a filter with an exponential kernel will provide better filtering results when impulsive or long-tailed noise is present. An adaptive design that allows for both filters to be utilized simultaneously is of paramount importance in an unknown noise environment. Such examples emphasize the versatility of the proposed adaptive approach, which can provide a wide range of different practical filters.

- Although the filter in (35) has been derived as a generalization of a nonparametric estimator, it can be used to combine different heterogeneous estimators applied to the same task. Specifically, the designer can utilize a number of different elemental filters, such as order statistic-based filters, the Bayesian adaptive filter, nearest-neighbor filters, and nonparametric estimators, and then combine all the different results using (35) and (36). The effectiveness of the adaptive scheme is determined by the validity of the elemental filtered results and the approximation required in (36). However, due to the different justification of the elemental filters, extensive simulation results are the only way to examine the performance of the filter in practical applications. Experimentation with color images will be used in the next section to demonstrate the effectiveness of the multiple filter and to access the improvement in terms of the performance achieved using a multiple nonparametric filter, vis-a-vis a simple nonparametric filter. We believe that the multiple filter can be a powerful design tool since it allows the combination of filters designed using different methodologies and different design parameters. This is of paramount importance in practical applications, since it allows for the development of efficient adaptive filters when no indication for the selection of a suitable filtering approach is available.

IV. APPLICATION TO COLOR IMAGES

In this section the performance of the proposed adaptive filters is evaluated in the most important area of multichannel processing, namely color image filtering. The noise attenuation properties of the different filters are examined by utilizing the

TABLE VI
New Filters: NMSE 10^{-2} for the Lena Image, 5×5

Noise	$BFMA$	$AMNFE$	$AMNFE2$	$AMNFG$	$AMNFG2$	$MAMNFEG$	$MAMNFEG2$	$MAMNF35$
1	0.5810	0.5908	0.518	0.5807	0.514	0.5866	0.519	0.5725
2	0.3146	0.3096	0.301	0.3110	0.307	0.3106	0.307	0.2744
3	0.231	0.331	0.324	0.321	0.314	0.213	0.316	0.295
4	0.3799	0.3936	0.371	0.3847	0.362	0.3899	0.368	0.3371
5	0.6634	0.635	0.583	0.6258	0.581	0.6306	0.588	0.6452

Fig. 1. The Lena image corrupted with Gaussian noise $\sigma = 30$ mixed with 4% impulsive noise.

Fig. 2. GVDF of (1) with a 3×3 window.

color image Lena. The test image has been contaminated using various noise source models in order to assess the performance of the filters under different scenarios (see Table I). Gaussian noise implies corruption by zero mean additive noise with

standard deviation σ. The impulsive noise has been simulated in two steps. In the first step, each image channel is corrupted independently using spikes with magnitude 255. The program assumes that both negative and positive spikes are equally likely to occur. Data clipping is used to limit the samples' value in the range [0, 255]. In a second step, a correlation factor $\rho = 0.5$ is used to determine the corruption of the pixel (i, j) in channel $\boldsymbol{C}$ if the same pixel is corrupted in any of the other two channels [7], [9]. The original image, as well as its noisy versions, are represented in the red–green–blue (RGB) color space. This color coordinate system is considered to be objective, since it is based on the physical measurements of the color attributes. The filters operate on the images in the RGB color space.

Since it is impossible to discuss all the adaptive filters resulting from the theory introduced in this paper, we construct six different filters based on our parametric and nonparametric designs. These filters are compared, in terms of performance, with other widely used nonadaptive and adaptive color image filters (see Table II).

A number of different objective measures can be utilized to assess the performance of the different filters. All of them provide some measure of closeness between two digital images by exploiting the differences in the statistical distributions of the pixel values [28]. The most widely used measure is the normalized mean square error (NMSE) defined as

$$\text{NMSE} = \frac{\sum\limits_{i=0}^{N1} \sum\limits_{j=0}^{N2} \|(\boldsymbol{y}(i,j) - \hat{\boldsymbol{y}}(i,j)\|^2}{\sum\limits_{i=0}^{N1} \sum\limits_{j=0}^{N2} \|(\boldsymbol{y}(i,j)\|^2} \qquad (39)$$

where $N1, N2$ are the image dimensions, and $\boldsymbol{y}(i,j)$ and $\hat{\boldsymbol{y}}(i,j)$ denote the original image vector and the estimation at pixel (i, j) respectively. Tables III–VI summarize the results obtained for the color test image Lena.

In many application areas, such as multimedia, telecommunications (e.g., high-definition television), motion picture production, the printing industry, and graphic arts, greater emphasis is given to perceptual image quality. Consequently, the perceptual closeness (alternatively perceptual difference or error) of the filtered image to the uncorrupted original image is ultimately the best measure of the efficiency of any color image filtering method. There are basically two major approaches used for assessing the perceptual error between two color images. In order to make a complete and thorough assessment

TABLE VII
EXISTING FILTERS: NCD FOR THE LENA IMAGE, 3×3

Noise	$BVDF$	$GVDF$	DDF	VMF	AMF	$ANNF$	$DDMF$	$\alpha - TMF$
1	0.05081	0.046824	0.04996	0.057356	0.033382	0.0373	0.0377	0.0504
2	0.0195	0.017023	0.017288	0.014703	0.028431	0.0178	0.0184	0.0175
3	0.0272	0.0219	0.0279	0.0215	0.0234	0.0196	0.0223	0.0197
4	0.03215	0.028157	0.029729	0.033228	0.029526	0.0263	0.0273	0.0298
5	0.04979	0.048793	0.053260	0.060989	0.041983	0.0413	0.0411	0.0396

TABLE VIII
NEW FILTERS: NCD FOR THE LENA IMAGE, 3×3

Noise	$BFMA$	$AMNFE$	$AMNFE2$	$AMNFG$	$AMNFG2$	$MAMNFEG$	$MAMNFEG2$	$MAMNF35$
1	0.036005	0.038947	0.0311	0.038169	0.030	0.038735	0.031	0.031552
2	0.02014	0.015047	0.0151	0.01886	0.0169	0.020266	0.016	0.021061
3	0.0179	0.01784	0.01701	0.0174	0.0171	0.0176	0.0173	0.0179
4	0.025051	0.023980	0.0213	0.024321	0.0213	0.024311	0.0215	0.022612
5	0.040409	0.040423	0.0331	0.040058	0.0325	0.040459	0.033	0.032657

Fig. 3. DDMF of (1) with a 3×3 window.

of the performance of the various filters, both approaches are used in this paper.

The first approach is to make an objective measure of the perceptual error between two color images. This leads us to the question of how to estimate the perceptual error between two color vectors. Precise quantification of the perceptual error between two color vectors is one of the most important and open research problem. RGB is the most popular color space used conventionally to store, process, display, and analyze color images. However, the human perception of color cannot be described using the RGB model [29]. Therefore, measures such as the normalized mean square error (NMSE) defined in the RGB color space are not appropriate to quantify the

perceptual error between images. Thus, it is important to use color spaces, which are closely related to the human perceptual characteristics and suitable for defining appropriate measures of perceptual error between color vectors. A number of such color spaces are used lately in areas such as computer graphics, motion pictures, graphic arts, and the printing industry. Among these, perceptually uniform color spaces are the most appropriate to define simple yet precise measures of perceptual error. The Commission Internationale de l'Eclairage (CIE) standardized two color spaces, $L^*u^*v^*$ and $L^*a^*b^*$, as perceptually uniform. The $L^*u^*v^*$ color space is chosen for our analysis because it is simpler in computation than $L^*a^*b^*$ color space, without any sacrifice in perceptual uniformity.

The conversion from nonlinear RGB color space (the nonlinear RGB values are the ones stored in the computer and applied to the CRT of the monitor to generate the image) to the $L^*u^*v^*$ color space is explained in detail in [30]. Nonlinear RGB values of both the uncorrupted original image and the filtered image are converted to corresponding $L^*u^*v^*$ values for each of the filtering method under consideration. In the $L^*u^*v^*$ space, the L^* component defines the lightness, and the u^* and v^* components together define the chromaticity. In a uniform color space, such as $L^*u^*v^*$, the perceptual color error between two color vectors is defined as the Euclidean distance between them, given by

$$\Delta E_{Luv} = [(\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2]^{1/2} \qquad (40)$$

where ΔE_{Luv} is the color error and $\Delta L^*, \Delta u^*,$ and Δv^* are the difference in the $L^*, u^*,$ and v^* components, respectively, between the two color vectors under consideration. Once the ΔE_{Luv} for each pixel of the images under consideration is computed, the normalized color difference (NCD) is estimated

TABLE IX
EXISTING FILTERS: NCD FOR THE LENA IMAGE, 5 × 5

Noise	$BVDF$	$GVDF$	DDF	VMF	AMF	$ANNF$	$DDMF$	$\alpha - TMF$
1	0.04087	0.032077	0.037725	0.043390	0.027598	0.0287	0.030	0.0325
2	0.01895	0.017996	0.018716	0.018032	0.027032	0.020	0.0197	0.0197
3	0.0242	0.0188	0.0227	0.0214	0.023	0.021	0.0191	0.0213
4	0.02675	0.023224	0.025749	0.028125	0.025296	0.0225	0.0219	0.0229
5	0.04018	0.033921	0.039696	0.045736	0.033816	0.0314	0.0327	0.0332

TABLE X
NEW FILTERS: NCD FOR THE LENA IMAGE, 5 × 5

Noise	$BFMA$	$AMNFE$	$AMNFE2$	$AMNFG$	$AMNFG2$	$MAMNFEG$	$MAMNFEG2$	$MAMNF35$
1	0.030970	0.028851	0.0261	0.028721	0.0279	0.02816	0.0277	0.029890
2	0.019244	0.01774	0.0173	0.017919	0.0177	0.017903	0.0179	0.016421
3	0.0198	0.0184	0.0191	0.0186	0.0181	0.0176	0.0184	0.0173
4	0.022851	0.021526	0.0212	0.021669	0.0216	0.021602	0.0215	0.021216
5	0.033989	0.029548	0.0281	0.029476	0.0294	0.029540	0.0294	0.031008

according to the following formula:

$$\text{NCD} = \frac{\sum_{i=0}^{N1}\sum_{j=0}^{N2}\|\Delta E_{Luv}\|}{\sum_{i=0}^{N1}\sum_{j=0}^{N2}\|E^*_{Luv}\|} \qquad (41)$$

where

$$E^*_{Luv} = [(L^*)^2 + (u^*)^2 + (v^*)^2]^{1/2}$$

is the *norm* or *magnitude* of the uncorrupted original image pixel vector in the $L^*u^*v^*$ space.

Tables VII–X summarize the results obtained for the test color image Lena.

Although quantitative measures, such as ΔE_{Luv} and NCD, are close approximations to the perceptual error, they cannot exactly characterize the complex attributes of human perception. Therefore, an alternative subjective approach is commonly used by researchers [31] for estimating the perceptual error.

The second evaluation approach is the subjective evaluation of the two images to be compared in which both images are viewed, simultaneously, under identical viewing conditions by a set of observers. A set of color image quality attributes can be defined for the subjective evaluation [31]. The evaluation must take into consideration three important factors in regard to color image filtering, namely, noise attenuation, edge preservation, and detail retention. Thus, the overall quality of the color image is determined from the following attributes:

- *detail preservation*, which corresponds to edge and fine detail preservation;
- *sharpness*, which refers to the distinctness of boundaries among colors;
- *color uniformity*, which refers to the consistency of the color in uniform areas;

Fig. 4.   α-TMF of (1) with a 3 × 3 window.

- *defects*, classified as any imperfections such as blocking artifacts that were not present in the original (noise-free) image.

In this study, the color images under consideration were viewed in parallel, on a Sun SparcStation 5 with a 24-b color monitor, and the observers were asked to mark scores on a printed evaluation sheet following the guidelines summarized in Table XI [32]. To subjectively evaluate the noise-removal capabilities of the algorithms, a similar procedure was followed. Observers were instructed to assign a lower number if noise was still present in the filtered output (see Table XI).

The results of the subjective evaluation is summarized in Tables XII and XIII, and is used to complement the results obtained from the objective evaluation using quantitative measures such as NMSE and NCD.

TABLE XI
SUBJECTIVE IMAGE EVALUATION GUIDELINES

Score	Overall Evaluation	Noise Removal Evaluation
1	Very Disruptive distortion	poor
2	Disruptive distortion	fair
3	Destructive but not disruptive distortion	good
4	Perceivable but not destructive distortion	very good
5	Imperceivable distortion	excellent

Fig. 5.   AMNFG of (1) with a 3 × 3 window.

Fig. 7.   BFMA of (1) with a 3 × 3 window.

Fig. 6.   MAMNFEG of (1) with a 3 × 3 window.

Fig. 8.   The RGB color image of peppers.

Now, having defined the measures for the assessment of performance by various filters, we proceed to the assessment itself. One of the obvious observations from the results in Tables III–XIII is the effect of window size on the performance of the filter. In the case of nonadaptive filters, the bigger window size (5× 5) gives considerably better results for the

Fig. 9. Peppers image corrupted with Gaussian noise $\sigma = 30$ mixed with 4% impulsive noise.

Fig. 10. BVDF of (9) with a 3×3 window.

TABLE XII
FIGURE OF MERIT

a	Overall Subjective Evaluation
b	Additive Gaussian Noise
c	Impulsive Noise
d	Moderate Mixed (Gaussian/Impulsive) Noise
e	Mixed (Gaussian/Impulsive) Noise

removal of Gaussian noise (noise model 1), while decreasing the performance for the removal of impulsive noise (noise model 2). Although a similar pattern follows for the adaptive filters, the effect of window size on performance is less dramatic as compared to the nonadaptive filters.

TABLE XIII
SUBJECTIVE EVALUATION

Filter	Figure of Merit				
Filter	a	b	c	d	e
BVDF	3	3	3	2	1
GVDF	3	3	4	3	3
DDF	3	3	3	3	2
VMF	4	2	5	3	3
AMF	2	5	1	3	4
ANNF	4	3	3	3	3
DDMF	3	3	3	3	3
α-TMF	3	3	3	3	3
BFMA	4	5	4	4	4
AMNFE	4	4	5	5	4
AMNFE2	5	4	5	5	4
AMNFG	4	4	5	5	5
AMNFG2	4	4	5	5	5
MAMNFEG	5	4	5	5	5
MAMNFEG2	5	4	5	5	5
MAMNF35	4	5	5	4	5

Fig. 11. GVDF of (9) with a 3×3 window.

The AMF is theoretically the best nonadaptive filter for the removal of pure Gaussian noise (noise model 1). The results in Tables III, V, VII, and IX confirms this. In other words, the NMSE, NCD, and the subjective measure all indicate best performance by the AMF. So, the performance of the AMF filter is used as a benchmark to compare the performance of the new adaptive filters in the same noise environment. The results in Tables IV, VI, VII, and X indicate that the new adaptive filters perform better or close enough to the AMF

Fig. 12. α-TMF of (9) with a 3×3 window.

Fig. 14. BFMA of (9) with a 3×3 window.

Fig. 13. AMNFE2 of (9) with a 3×3 window.

and outperform existing adaptive filters, such as the ANNF or DDMF in NMSE, NCD and in the subjective sense. Clearly, the new AMNFG2 adaptive filter is the best for Gaussian noise and performs exceptionally well, outperforming the existing filters (both adaptive and nonadaptive), with respect to all three error measures and for both window sizes.

Similarly, in the case of pure impulsive noise (noise model 2), the VMF gives the best performance among nonadaptive filters according to the results, as well the theory, and is thus used as a benchmark to evaluate the adaptive filters. As for the Gaussian noise case, the new adaptive filters perform better or close enough to the VMF and outperform the existing adaptive filters with respect to all three error measures and for both window sizes. For a 3×3 square window the proposed AMNFE2 adaptive filter emerges as the best contender for the

removal of impulsive noise, outperforming other nonadaptive and adaptive filters. For the case of a 5×5 square processing window our MAMNF35 filter clearly outperforms all other filters.

For the mixture of Gaussian and impulsive noise (noise models 3–5), the new adaptive filters consistently outperform any of the existing listed filters, both adaptive and nonadaptive, with respect to NMSE, NCD, and subjective measures, for both window sizes (see Figs. 1–7). This is demonstrated by the simple fact that, for noise models 3–5 (see Table I), the highest error among the new adaptive filters is lower than the lowest error among the existing filters, both adaptive and nonadaptive. Herein lies the real advantage of the new adaptive filters. As mentioned before, in real applications, the noise model is unknown *a priori*. Nevertheless, the most common noise types encountered in real situations are Gaussian or impulsive or a mixture of both. Therefore, the use of these new adaptive filters guarantees near optimal performance for the removal of any kind of noise encountered in practical applications. On the contrary, application of a "noise-mismatched" nonadaptive filter, such as a VMF for Gaussian noise or AMF for impulsive noise, can have profound consequences leading to unacceptable results.

In a final example, we compare the performance of the different filters in noise attenuation using a different color image. The test RGB image of peppers (see Fig. 8) is corrupted by a mixture of additive Gaussian noise (standard deviation, $\sigma = 30$) and 4% impulsive noise (see Fig. 9). All the filters considered operate using a square 3×3 processing window. Filtering results using different estimators are depicted in Figs. 10–14. A comparison of the images clearly favors our adaptive designs over existing techniques.

The following important conclusions can be drawn from the results listed above, regarding our adaptive designs.

- The proposed adaptive filters attenuate both impulsive and additive Gaussian noise with or without outliers present

Fig. 15.   BFMA of (8) with a 3 × 3 window.

Fig. 16.   MAMNFEG of (8) with a 3 × 3 window.

in the test image. It must be noted that if no assumption about the noise characteristics is made, the new adaptive filters provide results better than the results obtained by any other filter under consideration. The proposed adaptive filters can effectively remove impulses, smooth out nominal noise and keep edges and details unchanged as can be seen from the results obtained through the application of the BFMA and MAMNFEG filters on the noise-free peppers image (see Figs. 15 and 16).

• Results also indicate that the adaptive multiple filters are less sensitive to the window length as compared to the GVDF or the VMF. The performance of our adaptive filters decrease less as compared with that of the nonadaptive filter as window size increases. In ad-

dition, the adaptive filters do not suffer from VMF's inefficiency in a nonimpulsive noise scenario and small filtering window. More than that, the generalized adaptive framework allows for the utilization of elemental filters that utilize different window sizes balancing between the detail preservation property of the small processing window with the enhanced noise attenuation obtained through a window of size 5 × 5 .

• Considering the number of computations required for the implementation of the adaptive filters, it should be noted that it is comparable to those of any other multichannel filter. The adaptation procedure needed to weight the different elemental filters does not introduce significant additional computational cost. In addition, the different elemental filters can be run in parallel reducing the execution time and making the adaptive filters suitable for real-time implementation with digital signal processors. To the best of the authors' knowledge the adaptation mechanism introduced in this work is the only one capable of providing this form of parallel processing capability.

V. CONCLUSION

This paper has introduced adaptive algorithms for filtering color image data. These filters utilize Bayesian learning techniques and nonparametric methodologies to adapt to local image data. The behavior of the adaptive filters is analyzed and their performance is compared with that of the most commonly used nonadaptive filters. Simulation results and subjective evaluation of the filtered images indicate that the adaptive filters outperform all other filters under consideration. Moreover, as seen from the attached images, the adaptive filters preserve the chromaticity component, which is very important in the visual perception of color images.

REFERENCES

[1] P. E. Trahanias, I. Pitas, A. N. Venetsanopoulos, "Color image processing," in *Advances In 2D and 3D Digital Processing (Techniques and Applications)*, C. T. Leondes, Ed.   New York: Academic, 1994.

[2] I. Pitas and A. N. Venetsanopoulos, *Nonlinear Digital Filters: Principles and Applications.*   Norwell, MA: Kluwer, 1990.

[3] N. Nikolaidis and I. Pitas, "Multichannel L filters based on reduced ordering" in *Proc. IEEE Workshop on Nonlinear Signal Processing,* 1995.

[4] M. E. Zervakis and A. N. Venetsanopoulos, "Linear and nonlinear image restoration under the presence of mixed noise," *IEEE Trans. Circuits Syst.,* vol. 38, pp. 258–271, 1991.

[5] I. Pitas and A. N. Venetsanopoulos, "Order statistics in digital image processing," in *Proc. IEEE,* vol. 80, pp. 1893–1923, 1992.

[6] J. Astola, P. Haavisto, and Y. Neuvo, "Vector median filter," *Proc. IEEE*, vol. 78, pp. 678–689, 1990.

[7] P. E. Trahanias and A.N. Venetsanopoulos, "Vector directional filters: A new class of multichannel image processing filters," *IEEE Trans. Image Processing,* vol. 2, pp. 528–534, Apr. 1993.

[8] D. G. Karakos, P. E. Trahanias, "Combining vector median and vector directional filters: The directional-distance filters," in *Proc. IEEE Conf. on Image Processing,* ICIP-95, pp. 171–174.

[9] P. E. Trahanias, D. G. Karakos, and A.N. Venetsanopoulos, 'Directional processing of color images: Theory and experimental results," *IEEE Trans. Image Processing,* vol. 5, pp. 868–880, June 1996.

[10] L. Onural, M. Bigle Alp, and M.I. Gurelli, "Gibbs random field model based weight selection for the 2-D weighted median filter," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 16, pp. 831–837, 1994.

[11] R. J. Stevenson and S. M. Schweizer, "Nonlinear filtering structure for image smoothing in mixed-noise environments," *J. Math. Imaging Vis.*, vol. 2, pp. 137–154, 1992.

[12] K. N. Plataniotis, D. Androutsos, V. Sri, and A. N. Venetsanopoulos, "A nearest neighbor multichannel filter," *Electron. Lett.*, pp. 1910–1911, 1995.

[13] S. Fotopoulos and G Economou, Multichannel filters using composite distance metrics, *Proc. IEEE Workshop on Nonlinear Signal Processing*, pp. 503–506, 1995.

[14] P. J. Bickel, "On adaptive estimation," *Ann. Stat.*, vol. 10, pp. 647–671, 1982.

[15] A. P. Sage and J. L. Melsa, *Estimation Theory with Applications to Communication and Control.* Huntington, NY: R. E. Krieger, 1979.

[16] G. E. Box and G. C. Tiao, "A note on criterion robustness and inference robustness," *Biometrika,* vol. 51, pp. 169–173, 1964.

[17] ———, *Bayesian Inference in Statistical Analysis*. Reading, MA: Addison-Wesley, 1973.

[18] K. N. Plataniotis, "Distributed parallel processing state estimation algorithms," Ph.D dissertation, Florida Inst. Technol., Melbourne, FL, 1994.

[19] H. M. Kim and J. M. Mendel, "Fuzzy basis functions: Comparisons with other basis functions," *IEEE Trans. Fuzzy Syst.*, vol. 3, pp. 158–169, 1995.

[20] T. Cacoullos, "Estimation of a multivariate density," *Ann. Stat. Math.*, vol. 18, pp. 179–189, 1966.

[21] V. K. Epanechnikov, "Non-parametric estimation of a multivariate probability density," *Theory Prob. Appl.*, vol. 14, pp. 153–158, 1969.

[22] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. New York: Academic, 1990.

[23] L. Breiman, W. Meisel, and E. Purcell, "Variable kernel estimates of multivariate densities," *Technometrics*, vol. 19, pp. 135–144, May 1977.

[24] B. L. S. Prasaka Rao, *Non-parametric Functional Estimation.* New York: Academic, 1983.

[25] E. A. Nadaraya, "On estimating regression," *Theory Prob. Appl.*, vol. 15, pp. 134–137, 1964.

[26] G. S. Watson, "Smooth regression analysis," *Sankhya Ser. A*, vol. 26, pp. 359–372, 1964.

[27] T. J. Wagner, "Nonparametric estimates of probability density," *IEEE Trans. Inform. Theory,* vol. IT-21, no. 4, pp. 438–440, 1975.

[28] A. M. Eskicioglu, P. S. Fisher, and S. Chen, "Image quality measures and their performance," *IEEE Trans. Commun.*, vol. 43, pp. 2959–2965, 1995.

[29] W. K. Prat, *Digital Image Processing*, 2nd ed. New York: Wiley, 1991.

[30] C. A. Poynton, *Poynton's Color FAQ*, electronic preprint, 1995.

[31] P. G. Engeldrum, "A framework for image quality models," *J. Imaging Sci. Technol.*, vol. 39, pp. 312–318, 1995.

[32] N. Narita, "Consideration of subjective evaluation method for quality of image coding," *Electron. Commun. Jpn.*, part 3, vol. 77, pp. 84–97, 1994.

Konstantinos N. Plataniotis (S'88–M'95) received the B. Eng. degree in computer engineering and information science from the University of Patras, Patras, Greece, in 1988, and the M.S and Ph.D degrees in electrical engineering from the Florida Institute of Technology, Melbourne, in 1992 and 1994, respectively.

He was with the Computer Technology Institute, Patras, Greece, from 1989 to 1991. Currently, he is with the Digital Signal and Image Processing Laboratory, Department of Electrical and Computer Engineering, University of Toronto, Toronto, Ont., Canada. His research interests include image processing and analysis, multimedia, stochastic estimation, communication systems, fuzzy logic, and neural networks.

Dr. Plataniotis is a member of the Technical Chamber of Greece.

Dimitrios Androutsos (S'90) was born in Toronto, Ont., Canada, in 1969. He received the B.A.Sc. degree and the M.A.Sc. degree in electrical engineering from the University of Toronto in 1992 and 1994, respectively.

He is currently with the Digital Signal and Image Processing Laboratory, University of Toronto, and is working toward the Ph.D. degree in electrical engineering, in the area of content-based image retrieval. His research interests lie in color image filtering, adaptive filtering, perceptually lossless compression, signal processing for multimedia applications, and content-based image retrieval.

Sri Vinayagammoorthy received the B.E. degree in electrical and electronics engineering from the University of Madras, India, in 1987, and the M.Eng. degree from the Asian Institute of Technology, Bangkok, Thailand, in 1990.

He was with the UNEP-GRID, Asian Institute of Technology, from 1990 to 1994. He is currently with the Communication Group, University of Toronto, Toronto, Ont., Canada, working toward the M.A.Sc. degree. His research interest is nonlinear and adaptive processing of color and multispectral images.

Anastasios N. Venetsanopoulos (SM'79–F'88) received the Dipl. Engineering degree from the National Technical University of Athens (NTU), Greece, in 1965, and the M.S., M.Phil., and Ph.D. degrees in electrical engineering from Yale University, New Haven, CT, in 1966, 1968, and 1969, respectively.

He joined the University of Toronto, Toronto, Ont., Canada, in September 1968, where he has been a Professor in the Department of Electrical and Computer Engineering since 1981. He has served as Chairman of the Communications Group and Associate Chairman of the Department of Electrical Engineering. He was on research leave at the Federal University of Rio de Janeiro, Brazil, the Imperial College of Science and Technology, London, U.K., the National Technical University of Athens, Swiss Federal Institute of Technology, Lausanne, Switzerland, and the University of Florence, Italy, and was Adjunct Professor at Concordia University, Montreal, P.Q., Canada. He has served as Lecturer in 130 short courses to industry and continuing education programs, and as Consultant to several organizations. His general research interests include linear M-D and nonlinear filters, processing of multispectral (color) images and image sequences, telecommunications, and image compression. In particular, he is interested in the development of efficient techniques for multispectral image transmission, restoration, filtering, and analysis. He is a contributor to 24 books, and is co-author of *Nonlinear Filters in Image Processing: Principles and Applications* (Boston: Kluwer) and *Artificial Neural Networks: Learning Algorithms, Performance Evaluation and Applications* (Boston: Kluwer), and has published over 500 papers on digital signal and image processing and digital communications.

Dr. Venetsanopoulos has served as Chairman on numerous boards, councils, and technical conference committees including IEEE committees, such as the Toronto Section (1977–1979) and the IEEE Central Canada Council (1980–1982). He was President of the Canadian Society for Electrical Engineering and Vice President of the Engineering Institute of Canada (EIC) (1983–1986). He has been a Guest Editor or Associate Editor for several IEEE journals, and the Editor of the *Canadian Electrical Engineering Journal* (1981–1983). He is a member of the IEEE Communications, Circuits and Systems, Computer, and Signal Processing Societies, as well as a member of Sigma Xi, the Technical Chamber of Greece, the European Association of Signal Processing, the Association of Professional Engineers of Ontario (APEO) and Greece. He was elected as a Fellow of the IEEE "for contributions to digital signal and image processing," is a Fellow of EIC, and was awarded an Honorary Doctorate from the National Technical University of Athens for his "contribution to engineering" in October 1994.