

# Adaptive Filters and Multichannel Signal Processing

A.N. Venetsanopoulos  
Department of ECE  
University of Toronto  
Toronto, ON, M5S 3G4, Canada  
e-mail: anv@dsp.toronto.edu

K.N. Plataniotis  
Department of ECE  
University of Toronto  
Toronto, ON, M5S 3G4, Canada  
e-mail: kostas@dsp.toronto.edu

## Abstract

*Processing multichannel signals using digital signal processing techniques has received increased attention lately due to its importance in applications, such as multimedia technologies and telecommunications. The perspective of the topic offered here is one that comes primarily from work done in the field of multichannel image processing. The adaptive filtering techniques discussed in the paper relate to image processing with the emphasis placed primarily on filtering algorithms based on fuzzy logic concepts, multidimensional scaling, and order statistics-based designs. The strong potential of adaptive filters for multichannel image processing is illustrated with several examples.*

## 1.0. Introduction

The availability of a wide set of multichannel information sources in application areas, such as color image processing, multispectral remote sensing imagery, biomedicine, robotics, and industrial inspection have stimulated a renewed interest in developing efficient and cost effective processing techniques for multichannel signals. In recent years, significant advances have been made in the development of multichannel signal processing techniques. Such techniques are used in a variety of tasks, such as color image filtering and processing video sequences, enhancement of multispectral remote sensing data, seismic deconvolution for oil exploration, and boundary detection in vector fields. The most common signal processing task is noise filtering. Filtering is the process of estimating a signal degraded, in most cases, by additive random noise. This task is an essential part of any signal processing system, especially when the final product is used for human interpretation, such as visual inspection or for automatic analysis [1], [2]. Several filtering techniques have been proposed over the years. Among them are linear processing techniques, whose mathematical simplicity and the existence of a unifying theory

make their design and implementation easy. Their simplicity, in addition to their satisfactory performance in a variety of practical applications, has made them methods of choice for many years. However, most of these techniques operate assuming a Gaussian model for the statistical characteristics of the underlying process, and thus they try to optimize the parameters of a system suitable for such a model. Many signal processing problems cannot be efficiently solved by using linear techniques. For example, an area where linear processing techniques fail is in image processing, where conventional linear techniques cannot cope with the nonlinearities of the image formation model and do not take into account the nonlinear nature of the human visual system. Image signals are composed of flat regional parts and abruptly changing areas, such as edges, which carry important information for visual perception. Filters having good edge and image detail preservation properties are highly suitable for image filtering and enhancement. Unfortunately, most of the linear signal processing techniques tend to blur edges and to degrade lines, edges, and other fine image details [3].

Recently fuzzy techniques have been investigated to provide a bridge between linear and nonlinear techniques. This paper summarizes efforts devoted in the development of fuzzy systems suitable for filtering multichannel signals. The approach discussed here is an adaptive one. It integrates well-known theories in the areas of nonlinear filtering, multidimensional scaling, robust statistics, and fuzzy sets so as to form a new composite model. At the heart of the approach is a filter whose weights (parameters) are adaptively determined on the basis of local signal context. The filtering procedure is seen as the estimation of the prototype for a given set of input signals in a processing window. Thus, filtering is the process of replacing a noisy signal by a prototypical value, such that some metric which is a function of the filter output and its neighbors is minimized. The organization of the paper is as follows. In Section 2.0, we present the problem under consideration and we introduce the general framework for an adaptive solution. In

Section 3.0 we review membership functions based on the distance and similarity measures used in multichannel signal processing applications [4]. Properties, characteristics, and implementation issues are also discussed in detail. Section 4.0 deals with the problem of color image processing, an important area of multichannel signal processing, where we present experimental results and comparisons with other standard methods. Finally, Section 5.0 summarizes the conclusions and describes future research.

## 2.0. The General Framework

Consider the following commonly used model for a multichannel signal corrupted by additive noise:  $\mathbf{x} = \mathbf{y} + \mathbf{n}_o$ , where  $\mathbf{y}$  denotes the  $m$ -channel uncorrupted signal vector,  $\mathbf{x}$  is the corresponding noisy vector to be filtered and  $\mathbf{n}_o$  an additive noise vector. The signal processing literature has been dominated by the assumption of the Gaussian model for the statistical characteristics of the noise. Optimal filtering solutions for the problem at hand can be devised based on this assumption. Although the Gaussian model is often justified in practice by the central limit theorem, we often encounter noise processes that exhibit impulsive behavior and are more accurately modeled by heavy-tailed, non-Gaussian distributions [5], [6]. Impulsive behavior can be characterized in terms of short duration, high-energy spikes attaining large amplitudes with probability higher than that predicted by a Gaussian density model. Even if the level of non-Gaussian noise contamination is small, the performance of a filtering system optimized under the Gaussian assumption can suffer from drastic degradation. In such a case, the performance of classical adaptive filtering schemes, such as recursive least squares (RLS) [7] or simple weighted average filters, is seriously degraded. Thus, there is a need for a flexible and efficient filter class for non-Gaussian environments that can appear in practice. Let us define  $\mathbf{x}(l)$  as the multichannel sample to be processed at time index  $l$  and let  $W$  be a set of  $(n = 2N + 1)$  neighboring samples which belong to the window  $W$  centered on  $\mathbf{x}_l$ :  $W = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] = [\mathbf{x}_{l-N}, \dots, \mathbf{x}_{l-1}, \mathbf{x}_l, \mathbf{x}_{l+1}, \dots, \mathbf{x}_{l+N}]$ . Since the most commonly used method to decrease the level of random noise present in the signal is smoothing, an averaging operation is required in order to replace the noisy vector at the window center with a suitable representative vector (prototypical value). The general form of the system presented here is given as a weighted average of the input vectors inside the window  $W$ . Thus, the uncorrupted multichannel signal is estimated by determining the centroid as follows [4], [8]-[9]:

$$\hat{\mathbf{y}} = \sum_{j=1}^n w_j \mathbf{x}_j = \sum_{j=1}^n \frac{f(\mu_j)}{\sum_{j=1}^n f(\mu_j)} \mathbf{x}_j \quad (1)$$

where  $f(\mu_j) = \mu_j^\lambda$  is a function adaptively determined on

the basis of local context with  $\mu_j$  the membership function of input  $\mathbf{x}_j$  and  $\lambda$  a parameter such that  $\lambda \in [0, \infty)$ . In this adaptive design the weights provide the degree to which an input vector contributes to the output of the filter. The relationship between the signal at the window center (vector under consideration) and each signal within the window should be reflected in the decision for the weights of the filter. Through the normalization procedure, two constraints necessary to ensure that the output is an unbiased estimator are satisfied. Namely, (i) each weight is a positive number,  $w_j \geq 0$ , and (ii) the summation of all the weights is equal to unity:  $\sum_{j=1}^n w_j = 1$ . The weights can be considered to be a membership function with respect to the specific window component. The adaptive algorithm evaluates a membership function based on a given vector signal and then uses the membership values to calculate the final filtered output. To explain the concepts behind the filter only the definition of the fuzzy set is required. Other definitions, such as fuzzy rule bases and fuzzy control are not essential to this work and thus are omitted. Assume that  $\mathbf{X}$  is a universe of discourse with elements  $\mathbf{x}$ . Then, a fuzzy set  $\mathbf{A}$  in  $\mathbf{x}$  is a set of ordered pairs  $\mathbf{A} = \left\{ \frac{\mu_{\mathbf{A}}(\mathbf{x})}{\mathbf{x}} \mid \mathbf{x} \in \mathbf{X} \right\}$ , where  $\mu_{\mathbf{A}}(\mathbf{x})$  is the membership function or grade of membership of  $\mathbf{x}$  in  $\mathbf{A}$  which maps  $\mathbf{X}$  into a membership space  $\mathbf{M}$ . The range of the membership function is a subset of the non-negative real numbers whose supremum is finite. In practical applications  $\mathbf{M}$  is normalized to the interval  $[0, 1]$ . The design summarized here qualifies as an adaptive fuzzy system since it utilizes sample input data and inference procedures (here in the form of transformed distance metrics), to define a fuzzy system at each time instant. Through the adaptation mechanism utilized, the system structure changes over time resulting in a time-varying mapping between input values and filtered output. This temporal mapping defines an adaptive fuzzy system. As was argued in [10], adaptation, or learning, is essentially parameter changing. Thus, by changing the weights in (1), we have developed an adaptive fuzzy system capable of learning new associations between input patterns and new functional dependencies. The noise smoothing problem is seen as a problem of prototype estimation given a set of signal inputs. In this sense, filtering is the process of replacing the noise-corrupted multichannel signal at the window center by a prototype signal, such that the differences between this prototype and all its neighbors inside the window are minimized in some sense. This operation is, essentially, a defuzzification procedure. It determines the most appropriate signal value (a vector signal in the case of multichannel inputs), to represent a collection of elements whose membership functions have been constructed over a universe of discourse. Although a number of different defuzzification strategies exist, the centroid defuzzification approach, known as the *Center of Gravity* (CoG), is often utilized in practice. The CoG method

generates a defuzzified value which is at the center of the values of a fuzzy set. Its defuzzified output actually corresponds to the membership-graded weighted mean of the square (Euclidean) distance. To clarify this, let us consider a fuzzy set  $A$  that is defuzzified as:  $\mathbf{A} = (\frac{\mu_1}{x_1}, \frac{\mu_2}{x_2}, \dots, \frac{\mu_n}{x_n})$  where  $\mu_n$  is the membership function associated with the input value  $x_n$ . If a quadratic cost function is considered:  $K(\bar{x}) = \sum_{i=1}^n |x_i - \bar{x}|^T \mu_i |x_i - \bar{x}|$  the CoG defuzzified value is obtained when  $K(\bar{x})$  is minimized by differentiation:  $\bar{y} = \frac{\sum_{i=1}^n x_i \mu_i}{\sum_{i=1}^n \mu_i}$ . Simple inspection of the CoG defuzzified value obtained reveals the similarity with the adaptive filtered output of (1). We can therefore claim that the output of our adaptive filter can be considered as the output of the CoG defuzzification strategy with the noisy multi-channel signals as members of a fuzzy set and the membership functions  $\mu_i$ ,  $i = 1, 2, \dots, n$  defined over them. In such a design, the overall performance of the processing system is determined by the defuzzification procedure selected. The quadratic cost function discussed above can be generalized to include any arbitrary function of  $\mu$ . Under such a scenario, we assume that the cost function associated with the selection of the defuzzified value to represent the fuzzy set  $A$  is:  $K(\bar{x}) = \sum_{i=1}^n |x_i - \bar{x}|^T f(\mu_i) |x_i - \bar{x}|$  where  $f(\mu_i)$  is a function of the associated membership function. By minimizing the above quadratic form, a defuzzified (crisp) value can be obtained as:  $\bar{y} = \frac{\sum_{i=1}^n x_i f(\mu_i)}{\sum_{i=1}^n f(\mu_i)}$  which is identical to the form used to generate the filtered output in the adaptive design of (1). If the power function  $f(\mu_i) = \mu_i^\lambda$  with  $\lambda \in [0, \infty)$  is used, the defuzzified value can be obtained through the following equation:  $\bar{y} = \frac{\sum_{i=1}^n x_i \mu_i^\lambda}{\sum_{i=1}^n \mu_i^\lambda}$ . It can easily be seen that, in the generalized defuzzification rule, if  $\lambda = 1$  the widely used CoG strategy can be obtained. The defuzzified vector-valued signal obtained through the CoG strategy is a vector-valued signal, which was not part of the original set of input vectors. However, there are signal processing applications, such as image filtering, where it is desirable for the filter output to be one of the samples in the input window. As an example, the vector median filter (VMF) [11] is always constrained, by definition, to be one of the input samples. Thus, if the output of the adaptive fuzzy system is required to be a member of the original input set, a different defuzzification strategy should be used. By defining  $\mu_{(max)}$  to be the largest membership value, the adaptive weights in (1) can

be rewritten as follows:  $w_j = \frac{\mu_j^\lambda}{\sum_{j=1}^n \mu_j^\lambda} = \frac{(\frac{\mu_j}{\mu_{(max)}})^\lambda}{\sum_{j=1}^n (\frac{\mu_j}{\mu_{(max)}})^\lambda}$ . Given that  $\mu_j < \mu_{(max)}$ , as  $\lambda \rightarrow \infty$  then:

$$w_j = \begin{cases} 1 & \text{if } \mu_j = \mu_{(max)} \\ 0 & \text{if } \mu_j \neq \mu_{(max)} \end{cases} \quad (2)$$

Eq. (2) represents the *maximum defuzzifier* strategy. If the

maximum value occurs at a single point only, the maximum defuzzifier strategy coincides with the *mean of maxima* (MOM) defuzzification process. Through the *maximum defuzzifier*, the output of an adaptive fuzzy system is defined as:  $\hat{y} = x_j$   $\mu_j = \mu_{(max)}$ . In this case, the fuzzy adaptive filter behaves as a mode-like selection filter since by construction its output is always one of the samples inside the processing window. This selection property is shared by well-known nonlinear filters, such as the VMF and the *myriad* filter [6]. However, unlike these filters, which are optimized for specific noise models (the Laplacian and Cauchy model respectively), the fuzzy filter can be optimized for any noise model by tuning its membership function. Thus existing selection filters can be generalized.

### 3.0. Determining the Parameters

The most crucial step in the design of the adaptive fuzzy system lies in determining the membership function to be used in the construction of its weights. The difficulties associated with the meaning and measurement of the membership function hinder the applicability of fuzzy techniques to many practical applications. Our analysis considers the membership function as a function of similarity. Viewing membership values as similarity indicators is often used in prototype theory where membership is a notion of being similar to a representative of a category [12]. Thus a membership function value can be used to quantify the degree of similarity of an element to the set in question. The assumption behind this approach is that there exists a perfect (ideal) example of the set which belongs to the set to the full degree. The valuation of membership for the rest of the elements in the set can be regarded as the comparison of a given input  $x_i$  with the ideal input  $x_r$ , which results in a distance  $d(x_i, x_r)$ . The generic form of our function was given in [13] as:  $\mu_i = \frac{1}{1+f(d_i)}$ , where  $f(\cdot)$  is a function of the distance between the vector signal  $x_i$  and the ideal prototype  $x_r$ . Membership functions are either monotonically increasing functions from 0 to 1, monotonically decreasing from 1 to 0, or can be divided into monotonically increasing or decreasing parts. Each increasing or decreasing part is specified by a cross-over or dispersion point. The particular function  $f(d_i)$  used in the equation will determine the actual shape of the membership function. The approach of [13] suggests that since the relationship between distances measured in physical units and perception is generally exponential, an exponential type of function should be used in the generic membership function. The resulting type of a sigmoidal function deduced from this proposition can be defined as:

$$\mu_i(x_i) = \frac{1}{1 + \exp\left(\frac{-c(s(x_i, x_r) - a_1)}{\sigma_1}\right)} \quad (3)$$

with  $\lim_{s \rightarrow \infty} \mu_i = 1$  and  $\lim_{s \rightarrow 0} \mu_i = 0$  for a monotonically increasing function, where  $s(\mathbf{x}_i, \mathbf{x}_r)$  is any similarity function between two vectors and

$$\mu_i(\mathbf{x}_i) = \frac{1}{1 + \exp\left(\frac{-c(d_p(\mathbf{x}_i, \mathbf{x}_r) - a_2)}{\sigma_2}\right)}, \quad (4)$$

with  $\lim_{d \rightarrow \infty} \mu_i = 0$  and  $\lim_{d \rightarrow 0} \mu_i = 1$  for a monotonically decreasing function, where  $d_p(\mathbf{x}_i, \mathbf{x}_r)$  is any member of the generalized Minkowski family of metrics. In cases where a distance measure is used to quantify dissimilarity between the vector under consideration and the ideal prototype, the decreasing form of the function is utilized. If a similarity measure is used instead, we consider the monotonically increasing version of the membership function. The model of (3)-(4) satisfies the requirement imposed by the adaptive fuzzy framework. However, the membership function in its present form is computationally expensive since it involves the evaluation of the exponential function, and more importantly its parameters cannot be evaluated easily in practical applications. Therefore, other functions which can retain the same characteristics and are easier to implement are needed. Such a membership form was proposed in [12]. This function is continuously increasing (decreasing), satisfies the same boundaries conditions, complies with the generic membership function, and retains the properties of the S-shaped membership function. However, unlike the function in (3)-(4) it can be written as a rational function of two polynomials. In the new formulation, for any input value  $z$ , the membership function, by construction, can be completely characterized by only four parameters: (i-ii) the interval  $[a, b]$  of the input parameter  $z$ , (iii) the sharpness  $\lambda$  of the membership function, and, (iv) the inflection point  $\nu$  of the S-shaped function. Based on these parameters a membership function can be defined as:

$$\mu(z) = \frac{(1 - \nu)^{\lambda-1}(z - a)^\lambda}{(1 - \nu)^{\lambda-1}(z - a)^\lambda + \mu^{\lambda-1}(b - z)^\lambda}, \quad (5)$$

for a monotonically increasing function, and

$$\mu(z) = \frac{(1 - \nu)^{\lambda-1}(b - z)^\lambda}{(1 - \nu)^{\lambda-1}(b - z)^\lambda + \mu^{\lambda-1}(z - a)^\lambda}, \quad (6)$$

for a monotonically decreasing function, with the inflection point  $\nu$  defined via  $\mu(z_\nu) = \nu$  and  $z_\nu = (b - a)\nu + a$  or  $z_\nu = (a - b)\nu + b$  for the case of monotonically increasing or decreasing functions respectively. The sharpness of the function (an indicator of increasing/decreasing membership) can be defined respectively as:  $\lambda = f'(z - \nu)(b - a)$ . This membership function, which is universally applicable, can be utilized by considering the distance or similarity value as the input to the membership function. Assume that  $d_i = d_p(\mathbf{x}_i, \mathbf{x}_r)$  and  $s_i = s_i(\mathbf{x}_i - \mathbf{x}_r)$  are appropriate

distance or similarity measures between the vector under consideration and the ideal prototype, similar to those discussed in the previous section, we may rewrite the membership function needed in the fuzzy system of (1) as follows:

$$\mu_i = \frac{(1 - \nu)^{\lambda-1}(d_{max} - d_i)^\lambda}{(1 - \nu)^{\lambda-1}(d_{max} - d_i)^\lambda + \mu^{\lambda-1}(d_i - d_{min})^\lambda}, \quad (7)$$

with  $d_i \in [d_{max}, d_{min}]$ . Alternatively a monotonically increasing function can be defined based on a similarity measure  $s_i = s_i(\mathbf{x}_i - \mathbf{x}_r)$  as follows:

$$\mu_i = \frac{(1 - \nu)^{\lambda-1}(s_i - s_{min})^\lambda}{(1 - \nu)^{\lambda-1}(s_i - s_{min})^\lambda + \mu^{\lambda-1}(s_{max} - s_i)^\lambda}, \quad (8)$$

with  $s_i \in [s_{max}, s_{min}]$ . For the case of  $\lambda = 1$  the linear form of the membership function is obtained:  $\mu_i = \frac{s_i - s_{min}}{s_{max} - s_{min}}$  for a monotonically increasing function, and  $\mu_i = \frac{d_{max} - d_i}{d_{max} - d_{min}}$  for a monotonically decreasing function, which corresponds to the nearest-neighbor rule used in [3], [4] to define membership functions. The membership function is critically dependent on the similarity measure and the reference point selected. The ideal reference signal is the actual value of the multidimensional signal in the specific location under consideration. This signal, however, is not available. In addition, the noisy vector at the same location is not the appropriate choice since any input vector inside the window can be an outlier. Therefore, to make the procedure more robust and to ensure that the fuzzy system will provide accurate results, we eliminate the need for a reference point by evaluating the membership function used to weight each input vector  $\mathbf{x}_i$  in (1) on the aggregate distance between the vector  $\mathbf{x}_i$  under consideration and all the other vectors inside the processing window. Thus, the vector with the smallest overall distance (or maximum similarity) is now assigned the maximum membership value. Needless to say the membership function selected is now evaluated on the aggregated distances and not on the distance between the vector and the ideal prototype. It is obvious that such a design does not depend on a reference point and thus is more robust to occasional outliers. However, the computational complexity of the algorithm increases as a result of the need to evaluate a number of distances (similarities) in the processing window. Many distance metric or similarity functions [14], [3] can be used in the formulation of the aggregate distance. For example, assume that the noisy vector  $\mathbf{x}_i$  inside the processing window  $W$  is considered. Its aggregated distance from all other vectors inside the window is given as:  $d_i = \sum_{j=1}^n d_2(\mathbf{x}_i, \mathbf{x}_j)$  with  $d_2(i, j) = (\sum_{k=1}^m (x_i^k - x_j^k)^2)^{\frac{1}{2}}$  if the Euclidean metric has been selected to measure dissimilarity between two vector signals. This aggregate distance value is used as an input to the membership functions of (4) or (7) that will

be used to determine the fuzzy weights in the multichannel filter. Similarly, the *vector angle criterion* defines the scalar measure:  $d_{ai} = \sum_{j=1}^n A(\mathbf{x}_i, \mathbf{x}_j)$  as the distance associated with the noisy vector  $\mathbf{x}_i$  inside the processing window of length  $n$ , when the angle between two vectors  $A(\mathbf{x}_i, \mathbf{x}_j) = \cos^{-1}(\frac{\mathbf{x}_i \cdot \mathbf{x}_j}{\|\mathbf{x}_i\| \|\mathbf{x}_j\|})$  is used to measure dissimilarity. The approach suggested here eliminates the need for a reference point and generalizes the concept of membership function as a similarity indicator. In the suggested formulation, the valuation of membership is regarded not as a comparison with an ideal point but as a comparison to the rest of the elements to be included in the fuzzy set. To summarize, we have outlined a possible interpretation of the membership function and discussed how membership functions can be built based on similarity concepts. A generalized model for building membership functions was utilized here. The different similarity or distance measures discussed here can be used as input values to this membership function model. The possibility of tuning the design parameters, namely distance (similarity) metrics and membership functions provides the adaptive fuzzy system of (1) with a rich variety of modes of operation that range from simple selection type filters, such as the VMF to hybrid FIR/nonlinear filters, such as the  $\alpha$ -trimmed filter.

#### 4.0. Color Image Processing

The adaptive filters discussed here can be used to process multichannel signals in a variety of practical applications, such as color image processing, medical imaging, remote sensing applications, geophysical signal processing and military communications. Due to numerous practical applications, color images comprise an important class of multichannel signals and thus they can serve as an excellent illustrative application. We have conducted a set of experiments in order to evaluate the adaptive designs and compare their performance against the performance of other filters. The noise attenuation properties of the different filters are examined by utilizing the widely used 512X480 RGB color image 'Peppers'. The test image has been contaminated using various noise source models in order to assess the performance of the filters under different scenarios. The test image is contaminated with: (1) Gaussian noise ( $\sigma = 30$ ), (2) impulsive noise (4%), (3) Gaussian ( $\sigma = 15$ ) impulsive (2%), and (4) Gaussian ( $\sigma = 30$ ) impulsive (4%). The normalized mean square error (NMSE) has been used as quantitative measure for evaluation purposes. Since it is impossible to discuss all the adaptive filters resulting from the theory introduced here, we instead construct five different filters based on our designs. These filters are compared in terms of performance with other widely used multichannel filters, such as Basic vector directional filter (BVDF), Generalized vector directional

filter (GVDF), Arithmetic mean filter (AMF), Directional-distance filter (DDF), Vector median filter (VMF) [3] and the Hybrid directional filter (HF) and Adaptive hybrid directional filter (AHF) [15]. In particular, we introduce a simple rank-order filter (hereafter referred to as content-based rank filter (CBRF)), which can be seen as an adaptive system utilizing a maximum defuzzifier. We also include the fuzzy vector directional filter (FVDF) which is based on the CoG defuzzification rule, the membership formula of (4), and the  $d_{ai}$  aggregated distance evaluated over the filtering window  $W$ . Also included in the set, the adaptive nearest-neighbor filter (ANNF) which is based on the maximum defuzzification strategy, the linear version of the membership function formula of (7), and the distance measure of  $d_i$ . Further, we utilized the same defuzzification formula and the same membership function, along with the aggregated distance of  $d_2$  to derive the double window nearest neighbor filter ANNMF. By using the Canberra distance [3] and the distance measure of  $d_i$  instead of the angular distance, four new filters have been devised, namely the CANNF, CANNMF, CBANNF, and the CBANNMF. From the results listed in the tables, it can be easily seen that the adaptive designs provide consistently good results in all types of noise, outperforming the other multichannel filters under consideration here. The versatile design of (1) allows for a number of different filters which can provide solutions to many types of different filtering problems. Simple adaptive designs, such as the ANNF or the CANNF can preserve edges and smooth noise under different scenarios, outperforming other widely used multichannel filters. If knowledge about the noise characteristics is available, the designer can tune the parameters of the adaptive filter to obtain better results. Finally, considering the number of computations, the computationally intensive part of the adaptive fuzzy system is the required distance calculation. However, this step is common in all multichannel algorithms considered here. In summary, the design is simple, does not increase the numerical complexity of the multichannel algorithm, and delivers excellent results for complicated multichannel signals, such as real color images.

#### 5.0. Conclusions

The paper reviewed a framework for adaptive multichannel signal processing based on fuzzy concepts. The framework combines nonlinear filters, fuzzy membership functions and distance (similarity) criteria. Several filters can be considered special cases of this framework. The behavior of these adaptive designs was analyzed and their performance was compared with that of the most commonly used filters for a problem of great practical importance, namely color image processing. Multichannel signal processing is a rich and expanding field. Numerous new and advanced areas have appeared which have increased the im-

portance of the tools introduced and analyzed here. Multimedia signal processing, visual data processing and analysis multimodal signal processing telecommunications, digital audio restoration, satellite imagery, seismic deconvolution, and biomedicine are only some of the areas in which the methodologies proposed here can be applied. From a long-term research perspective, there is a need to establish a coherent theoretical foundation for nonlinear filtering algorithms. New algorithms and methodologies which may result in even more effective filtering structures suitable for intelligent processing of multimedia signal processing demand investigation. The framework presented here can serve as an initial point for further research and development in the area, and ultimately help in the development of new results and products in the near future.

**Table 1. NMSE( $\times 10^{-2}$ ) for the RGB 'peppers' image,  $3 \times 3$  window**

| Filter  | Noise | Model |      |      |  |
|---------|-------|-------|------|------|--|
|         | 1     | 2     | 3    | 4    |  |
| None    | 5.027 | 6.53  | 3.29 | 6.51 |  |
| BVDF    | 3.93  | 1.510 | 0.86 | 1.49 |  |
| CBRF    | 1.96  | 0.47  | 0.44 | 0.47 |  |
| GVDF    | 1.85  | 0.46  | 0.36 | 0.46 |  |
| DDF     | 3.51  | 0.59  | 0.56 | 0.59 |  |
| VMF     | 1.84  | 0.37  | 0.33 | 0.38 |  |
| FVDF    | 1.46  | 0.43  | 0.34 | 0.41 |  |
| ANNF    | 1.110 | 0.510 | 0.31 | 0.52 |  |
| ANNMF   | 0.91  | 0.36  | 0.31 | 0.34 |  |
| HF      | 1.59  | 0.47  | 0.36 | 0.48 |  |
| AHF     | 1.43  | 0.43  | 0.36 | 0.47 |  |
| CANNF   | 1.19  | 0.47  | 0.35 | 0.47 |  |
| CANNMF  | 0.90  | 0.47  | 0.43 | 0.46 |  |
| CBANNF  | 1.23  | 0.46  | 0.46 | 0.44 |  |
| CBANNMF | 0.89  | 0.45  | 0.43 | 0.46 |  |

## References

- [1] I. Pitas, A.N. Venetsanopoulos, *Nonlinear Digital Filters: Principles and Applications*, Kluwer Academic, Norwell Ma., 1990.
- [2] I. Pitas, A.N. Venetsanopoulos, 'Order statistics in digital image processing', *Proceedings of IEEE*, vol. 80, no. 12, pp. 1893-1923, 1992.
- [3] K.N. Plataniotis, A.N. Venetsanopoulos, *Color Image Processing and Applications*, Springer Verlag, Berlin, in press.
- [4] K.N. Plataniotis, D. Androutsos, A.N. Venetsanopoulos, 'Adaptive fuzzy systems for multichannel signal processing', *Proceedings of the IEEE*, vol. 87, no. 9, pp. 1601-1622, 1999.

**Table 2. NMSE( $\times 10^{-2}$ ) for the RGB 'peppers' image,  $5 \times 5$  window**

| Filter  | Noise | Model |      |      |  |
|---------|-------|-------|------|------|--|
|         | 1     | 2     | 3    | 4    |  |
| None    | 5.027 | 6.53  | 3.29 | 6.51 |  |
| BVDF    | 4.28  | 2.80  | 1.70 | 4.13 |  |
| CBRF    | 1.46  | 0.71  | 0.68 | 0.72 |  |
| GVDF    | 1.24  | 0.70  | 0.66 | 0.71 |  |
| DDF     | 2.15  | 0.77  | 0.74 | 0.76 |  |
| VMF     | 1.34  | 0.68  | 0.66 | 0.68 |  |
| FVDF    | 2.11  | 0.73  | 0.70 | 0.72 |  |
| ANNF    | 1.01  | 0.53  | 0.53 | 0.62 |  |
| ANNMF   | 0.81  | 0.45  | 0.45 | 0.46 |  |
| HF      | 1.01  | 0.99  | 0.77 | 0.97 |  |
| AHF     | 1.12  | 0.91  | 0.76 | 0.98 |  |
| CANNF   | 1.03  | 0.74  | 0.67 | 0.75 |  |
| CANNMF  | 0.87  | 0.64  | 0.65 | 0.65 |  |
| CBANNF  | 1.01  | 0.73  | 0.67 | 0.74 |  |
| CBANNMF | 0.87  | 0.64  | 0.64 | 0.64 |  |

- [5] S. Haykin, P. Lee, E. Derbez, 'Optimum nonlinear filtering', *IEEE Trans. on Signal Processing*, vol. 45, no. 11, pp. 2774-2786, 1997.
- [6] S. Kalluri, G.R. Arce, 'Adaptive weighted myriad filter algorithms for robust signal processing in  $\alpha$ -stable noise environments', *IEEE Trans. on Signal Processing*, vol. 46, no. 2, pp. 322-334, 1998.
- [7] S. Haykin, *Neural Networks: A Comprehensive Foundation*, McMillan, N.Y., 1994.
- [8] K.N. Plataniotis, D. Androutsos, A.N. Venetsanopoulos, 'Fuzzy adaptive filters for multichannel image processing', *Signal Processing Journal*, vol. 55, no. 1, pp. 93-106, 1996.
- [9] K.N. Plataniotis, D. Androutsos, A.N. Venetsanopoulos, 'Multichannel filters for image processing', *Signal Processing: Image Communications*, vol. 9, no. 2, pp. 143-158, 1997.
- [10] B. Kosko, *Neural networks for signal processing*, Prentice Hall, Englewood Cliffs, N.J., 1991.
- [11] J. Astola, P. Haavisto, Y. Neuvo, 'Vector median filters', *Proceedings of the IEEE*, vol. 78, pp. 678-689, 1990.
- [12] J. Dombi, 'Membership function as an evaluation', *Fuzzy Sets and Systems*, vol. 35, pp. 1-21, 1990.
- [13] H.J. Zimmerman, P. Zysno, 'Quantifying vagueness in decision models', *European Journal of Operation Research*, vol. 22, pp. 148-154, 1996.
- [14] V. Barnett, 'The ordering of multivariate data', *J. Royal Statistical Society A*, vol. 139, Part 3, pp. 318-348, 1976.
- [15] M. Gabbouj, F.A. Cheickh, 'Vector Median-Vector Directional hybrid filter for color image restoration', *Proceedings of EUSIPCO-96*, pp. 879-881, 1996.