A FUZZY-LOGIC APPROACH TO THE DESIGN OF MIXED-NORMS ADAPTIVE FILTERS ALGORITHMS

Mohiy Mohamed Hadhoud * and A. S. Ibrahim**

* Electrical Communications Department, ** Automatic Control Department
Faculty of Electronic Engineering, Menoufia University
Menouf, 32952-EGYPT

ABSTRACT

A Fuzzy Logic approach to the design of adaptive algorithms is presented. The proposed scheme easily combines the benefits of several adaptive algorithms into one algorithm. This is achieved by soft transition among different algorithms based on local context. A function is defined to map the input space into a normalized consistency domain. A fuzzy measure of the local signal properties is obtained using membership functions defined on the consistency universe of discourse. To increase the algorithm decision capability, a procedure for on-line tuning of the membership functions is provided. The proposed technique is applied to the design of a class of mixed-norm stochastic gradient adaptive algorithms. This combines the required advantage of each algorithm into one algorithm. The benefits of the Least absolute Difference (LAD) or (sign algorithm (SA)), the Least Mean Square (LMS), and the least mean Fourth (LMF) are utilized. The resultant algorithm provides more accurate final solution, very low misadjustment, and high robustness in the presence of impulsive noise. The performance of the proposed algorithm is demonstrated in a system identification simulation. Results show the superiority of the proposed fuzzy-based algorithm in the presence of Gaussian and impulsive measurement noise.

Keywords: Adaptive algorithms, Fuzzy-logic-based, Mixed-Norm, nonlinear, identification.

INTRODUCTION

Adaptive algorithms have found many applications in the field of signal processing [1]. Many adaptation strategies were proposed in the literature based on the minimization of a specific error norm. Each strategy inherently possesses its advantages and disadvantages. Recently, mixed-norm techniques were proposed to combine the benefits of two algorithms [1-2]. In Reference 1, a probabilistic approach is used to estimate a mixing parameter of the two algorithms. Reference 2 proposed a generalized regularized approach aimed to include the Least Square (LS) and the Least Absolute deviation (LAD) (the sign algorithm (SA)) in an adaptive descent algorithm in the context of Bayesian theory.

This paper proposes a fuzzy-logic approach to the design of mixed norms adaptive algorithms. The proposed approach can actually be used for a variety of applications. For example, soft transitions among parallel algorithms, variable structure algorithms or algorithms with different window sizes. Illustrative examples are given through the design of mixed-norm adaptive algorithms.

The paper is organized as follows: the next section describes a generalized mixed-norm adaptive algorithm. The architecture of a fuzzy rule-based system is given in the following section. Next, a consistency function is defined which gives a measure of the desired signal properties. The definition of fuzzy sets on the consistency universe of discourse is also given, together illustrative examples followed by the results. The conclusions are presented at the end of the paper.

FILTER DESIGN

Traditionally, filter design is based on the minimization of a cost function of the form:

$$J=F(e) \quad (1)$$

where e is the error between the filter output and a certain desired input.

Let us consider the following combination of error norms:

$$J(k) = \sum_{i=1}^{N} \alpha_i f_i(e_k) \quad (2)$$

where

$f_i(e_k)$, i = 1, 2, 3, ... , N are functions of the error. Each function reflects a desirable action of the filter. For example, $f_j(e_k) = E\{e_k^2\}$ means that the filter during convergence tries to minimize the LMS error.

α_i , i=1,..,N

are in general, real positive scalars, called mixing parameters, and should satisfy the following two conditions:

$$\sum_{i=1}^{N} \alpha_i = 1 \quad (3\text{-a})$$

$$\alpha_i \, \varepsilon[0, 1] \; i=1,.., N \quad (3\text{-b})$$

The filter update equation takes the form:

$$W(k+1)W(k)+\Delta W(k+1) \quad (4)$$

Where $\Delta W(k+1) = \mu \nabla J(k)$ and $\nabla J(k)$ is the gradient of $J(k)$ and can be written as:

$$\nabla J(k) = \sum_{i=1}^{N} \nabla \alpha_i f_i(e_k) = \sum_{i=1}^{N} \alpha_i \nabla f_i(e_k) \quad (5)$$

From (4) and (5),

$$W(k+1) = W(k) + \mu \sum_{i=1}^{N} \alpha_i \nabla f_i(e_k) \quad (6)$$

Therefore the filter response depends on the gradient of $f_i(e_k)$ and the mixing parameters α_i . i =1,.., N

The scalars α_i need to be carefully chosen, which is not an easy task. The choice of the mixing parameters depends on the properties of currently input data to the filter. Such properties can not be sharply defined and can be evaluated only vaguely. (e.g. is a specific data point near an outlier?, very near ?)

There are some strategies for calculating the parameters α_i. Reference 1 suggested a probability test at each point to determine one mixing parameter λ. This method is based on the probability that a data point is greater than a threshold level, which is again unknown quantity.

However, fuzzy rule-based systems solve this problem easily and effectively by combining the consequent of the rules. Simply, we make soft decisions based on each condition, aggregate the decisions made, and finally make a decision based on the aggregation. This gives smoother and more effective result compared to a method that makes hard decision at each point.

FUZZY RULE -BASED SYSTEM

Let us consider the filter update given by Equation 4. Define the rule base system consisting of N-rules. The i^{th} rule is given by:

Rule (i): *If* s_k *is* A_j *Then* $\Delta W_j = H_j$ (7)

note that A_j is the linguistic value associated with the lingustic variable s_k (k is the sampling index)

e.g. A_j = " the local data point is near an outlier"

and H_j is the value of increment ΔW_j calculated according to the definition of $f_j(e_k)$ in Equation 2.

i.e. $H_j = \mu_i \nabla f(e_k)$ (8)

The total increment ΔW_{k+1} can be computed from the output of the fuzzy rule-base as [3];

$$\Delta W_{k-1} = \frac{\sum_{i=1}^{N} t_i \Delta W_i}{\sum_{i=1}^{N} t_i} \quad (9)$$

where α_j is replaced by t_j., which is the degree of satisfaction of the antecedent in the i^{th} rule.

CONSISTANCY FUNCTION

The procedure given in the previous section is used to design a class of mixed-norm stochastic gradient adaptive algorithms for system identification applications, in which the desired response is corrupted with an additive impulses and Gaussian noise. The system identification structure is shown in Figure 1.

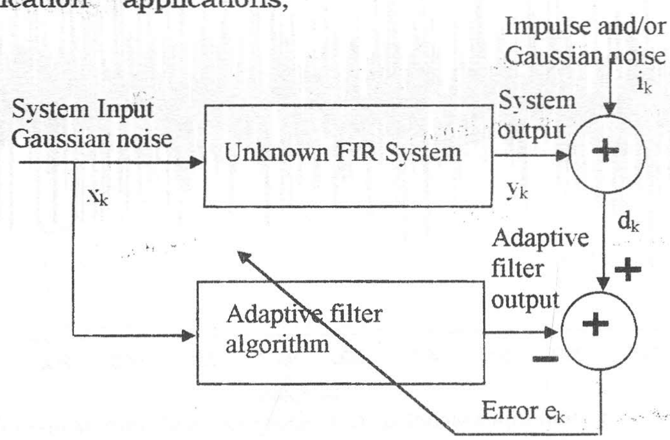


Figure 1 Adaptive system identification problem

In Figure 1, the adaptive filter has finite impulse response. The input x_k is a zero-mean Gaussian noise, y_k is the system output, i_k is an additive impulsive measurement noise, and d_k is the adaptive filter algorithm desired input. Figure 2 illustrates part of the input signal to the system and the desired signal with impulse noise.

To get a measure of the property of the adaptive filter desired input d_k, a consistency function $C(d_k)$ is defined. This function shows how a certain data point d_k is consistent with the previous data points. It also gives a measure of how far is the data point from an outlier. $C(d_k)$ is defined such that:

$$C(d_k) \ : \ d_k \rightarrow [0,1] \tag{10}$$

One possible choice is

$$C(d_k) = e^{\frac{-r_k}{\gamma}} \tag{11}$$

where

$$r_k = \sum_{i=1}^{M} (d_k - d_{k-i})^2 \tag{12}$$

and γ is a scalar. The function $C(d_k)$ is an indicator of the noise corruption of the data point d_k. The smaller the value of the function $C(d_k)$, the more likely the data point is near an outlier

MEMBERSHIP FUNCTION

Theoretically, any number of fuzzy sets can be defined on the consistency universe of discourse. In this work the number of sets is equal to the number of algorithms to be mixed. The fuzzy sets representing "Very small", "Small", "Medium", ... consistency with corresponding membership functions μ_{VS}, μ_S, μ_M, ... can be used.

FUZZY FILTER ALGORITHM A

This filter algorithm underlies the LMS and the SA algorithms. The filtering scheme is based on the minimization of the cost function:

$$J(k) = \alpha_1 E\{e_k^2\} + \alpha_2 E\{| e_k |\} \tag{13}$$

Only two membership functions representing "Small" and "Large" consistency are defined on the C-domain, as shown in Figure 3-a.
The membership functions are chosen as bell-shaped function defined by:

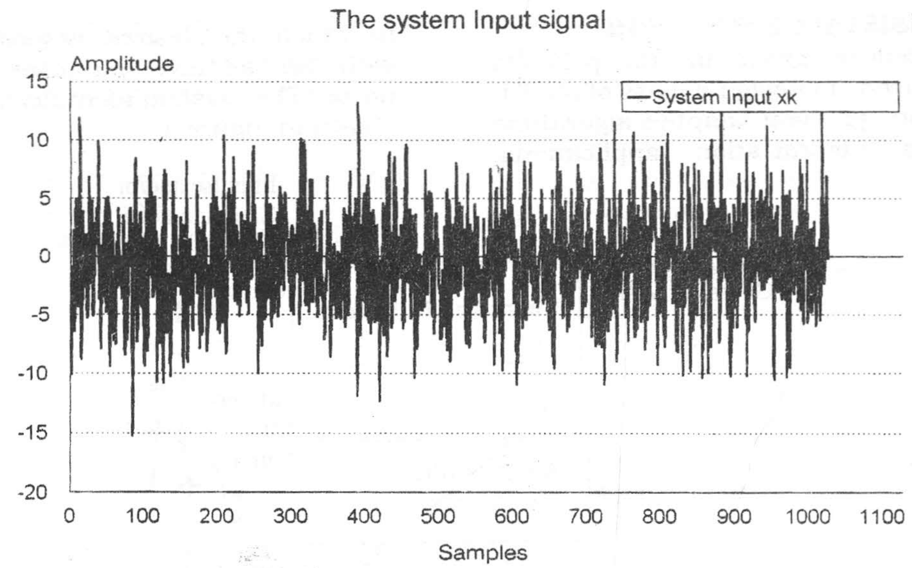
$$\mu(C) = 1 - e^{-\{\frac{\beta}{|\varphi - C|}\}^{2.5}} \tag{14}$$

where φ and β are scalars controlling the center and width of the membership function, respectively. Hence

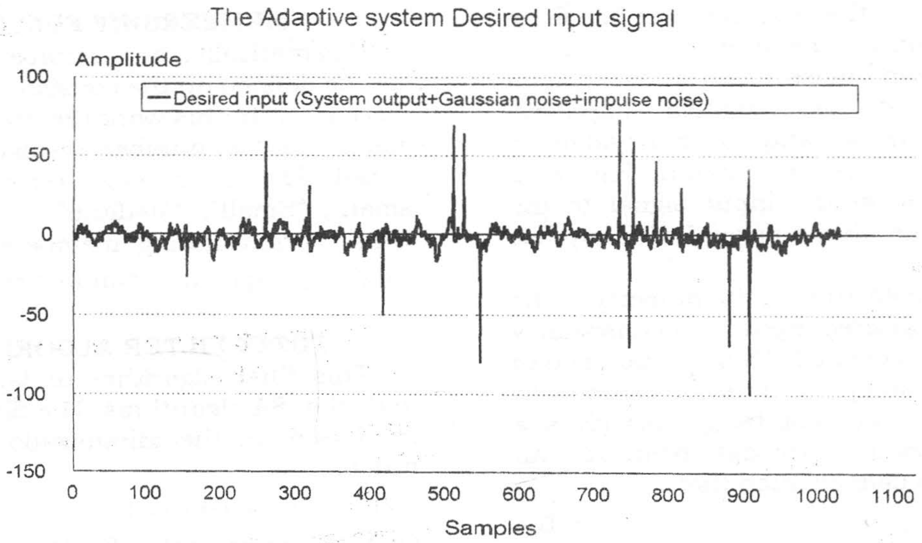
$$\mu_s(C) = 1 - e^{-\{\frac{\beta}{|C|}\}^{2.5}} \tag{15}$$

$$\mu_L = 1 - \mu_s$$

A suitable value of β is found to be 0.5.



(a) The identification system input signal x_k, Gaussian noise with zero mean and variance 25



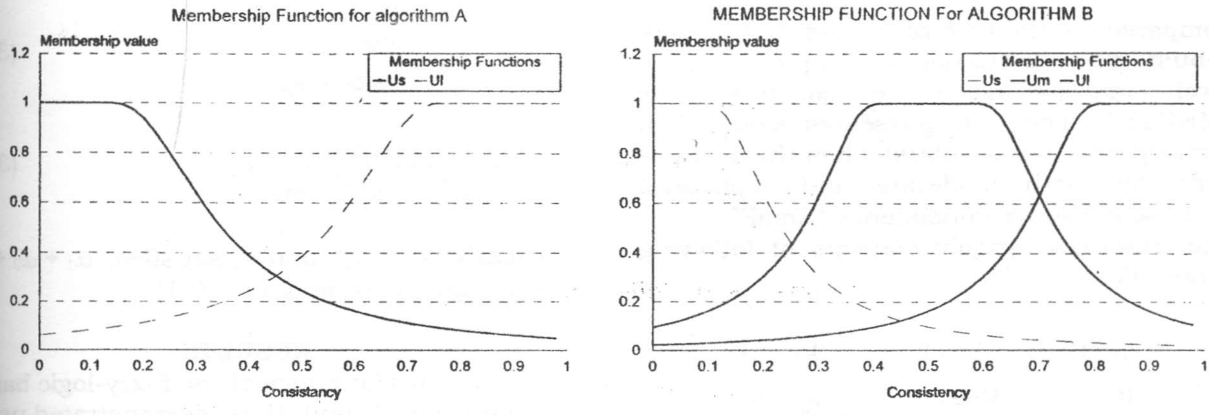
(b) Adaptive filter algorithm desired input d_k, consists of the system output + Gaussian noise of variance 9 + Impulsive noise

Figure 2 System input and the noisy adaptive filter algorithm desired input of the adaptive algorithm

The width of the membership function can be adapted according to the consistency of the data points by setting

$$\beta=\beta_o + 0.1C \qquad (16)$$

It is worth mentioning that the on-line adjustment of the membership functions by changing the value of β according to equation (16) highly improves the algorithm's response.



(a) Two partitions (b) Three partitions

Figure 3 Consistency versus membership function.

The LMS algorithm provides faster initial convergence, faster time variation, and accurate final solution, depending on the convergence factor μ, in the absence of impulse noise. i.e. when data points are far from outlier or equivalently for "large" consistency. The SA possesses slow initial convergence, but is robust when the signal is corrupted with impulse noise, i.e. consistency "Small".

The inclusion of the membership functions μ_S and μ_L in the rule-base provides effective soft mixing of the two algorithms. The rule-base simply consists of two rules:

$R^{(1)}$: If C is small $\Delta W_1 = H_1$ (17)

else:

$R^{(2)}$: If C is Large $\Delta W_2 = H_2$ (18)

The total incremental ΔW is given by:

$$\Delta W = \frac{\mu_s(C^*)H_1 + \mu_L(C^*)H_2}{\mu_s(C^*) + \mu_L(C^*)} \quad (19)$$

where

$$H_1 = \mu_1 \operatorname{sign}(e_k)\,\overline{x}(k) \quad (20)$$

$$H_2 = 2\,\mu_2\, e_k\, \overline{x}(k) \quad (21)$$

and $\mu_s(C^*) = t_1$ the degree of satisfaction of rule number 1.

$\mu_L(C^*) = t_2$ the degree of satisfaction of rule number 2..

The filter update is given by:

$$W_{k+1} = W_k + \Delta W$$

$$= W_k + \frac{t_1 H_1 + t_2 H_2}{t_1 + t_2} \quad (22)$$

Equation 22 minimizes the cost function given by Equation 13 for:

$$\alpha_2 = \frac{\mu_s(C^*)}{\mu_s(C^*) + \mu_L(C^*)}$$

$$\alpha_1 = \frac{\mu_L(C^*)}{\mu_s(C^*) + \mu_L(C^*)}$$

Clearly α_1 and α_2 satisfies $\alpha_1 + \alpha_2 = 1$, and $\alpha_1, \alpha_2\ \varepsilon[0,1]$

FILTER ALGORITHM B

This filter algorithm underlies the SA, LMS and the LMF filtering algorithms. The filtering scheme is based on the minimization of the cost function:

$$J(k) = \alpha_1 E\{|e_k|\} + \alpha_2 E\{e_k^2\} + \alpha_3 E\{e_k^4\} \quad (23)$$

Three membership functions representing "Small", "Medium" and "Large" consistency are defined on the C-domain, as shown in Figure 3-b.

The LMF algorithm has faster convergence in the absence of impulse noise, i.e. when data point is far from outlier or equivalently for "Large" consistency. The LMS possesses slower initial convergence as

compared to the LMF algorithm, but is more robust when the signal is corrupted with low level impulse noise, i.e. consistency is "Medium". The SA possesses slow initial convergence, but is robust when the signal is corrupted with moderate and high level impulse noise, i.e. consistency "Small".
The rule-base simply consists of following three rules:

$R^{(1)}$: If C is Small $\Delta W_1 = H_1$
$R^{(2)}$: If C is Medium $\Delta W_2 = H_2$
else
$R^{(3)}$: If C is Large $\Delta W_3 = H_3$ (24)

The total incremental ΔW is given by:

$$\Delta W = \frac{\mu_S(C^*)H_1 + \mu_M(C^*)H_2 + \mu_L(C^*)H_3}{\mu_S(C^*) + \mu_M(C^*) + \mu_L(C^*)} \quad (25)$$

where

$$H_1 = \mu_1 \ \text{sign}(e_k)\,\bar{x}(k) \quad (26)$$

$$H_2 = 2\mu_2 \ e_k \ \bar{x}(k) \quad (27)$$

$$H_3 = \mu_3 \ e_k^3 \ \bar{x}(k) \quad (28)$$

and $\mu_S(C^*) = t_1$ is the degree of satisfaction of rule number 1.

$\mu_M(C^*) = t_2$ is the degree of satisfaction of rule number 2..

$\mu_L(C^*) = t_3$ is the degree of satisfaction of rule number 3..

The filter update is given by:

$$W_{k+1} = W_k + \Delta W = W_k + \frac{t_1H_1 + t_2H_2 + t_3H_3}{t_1 + t_2 + t_3} \quad (29)$$

Equation 29 minimizes the cost function given by Equation 23 for:

$$\alpha_1 = \frac{\mu_S(C^*)}{\mu_S(C^*) + \mu_M(C^*) + \mu_L(C^*)} \quad (30)$$

$$\alpha_2 = \frac{\mu_M(C^*)}{\mu_S(C^*) + \mu_M(C^*) + \mu_L(C^*)} \quad (31)$$

$$\alpha_3 = \frac{\mu_L(C^*)}{\mu_S(C^*) + \mu_M(C^*) + \mu_L(C^*)} \quad (32)$$

Clearly α_1 , α_2 and α_3 satisfies, $\alpha_1 + \alpha_2 + \alpha_3 = 1$, and α_1 , α_2 and α_3 ε [0,1]

RESULTS

The performance of the fuzzy-logic based algorithms A and B is demonstrated using identification simulation problem as shown in Figure 1. Adding impulse noise and/or Gaussian noise to the system output y_k forms the desired input signal d_k. The unknown system is an FIR filter with coefficients [0.1, 0.2, 0.3, 0.4, 0.5, 0.4, 0.3, 0.2, 0.1]. The length of the sliding window is 9. In all simulation studies in this paper, the adaptation gains are chosen such that algorithms are stable and the initial convergence rates of all algorithms are almost identical when no impulsive noise is present.

In order to evaluate the performance of the different algorithms, the logarithm of the normalized weight error vector norm (WEN) is used [1].

$$\text{WEN} = 10 \ \log \ [\frac{||W^o - W(k)||}{||W^o||}] \quad (33)$$

Graph of the WEN for algorithm A is shown in Figure 4 where the desired signal is corrupted only with Gaussian noise. Figure 5 indicates the convergence of filter A when the desired signal is corrupted with both Gaussian and impulsive noise. Similar graphs of filter algorithm B are shown in Figures 6 and 7.

As shown from Figures 4 to 7 the performance of the fuzzy based algorithms outperforms, in the steady state error, the individual SA, LMS, and LMF algorithms.

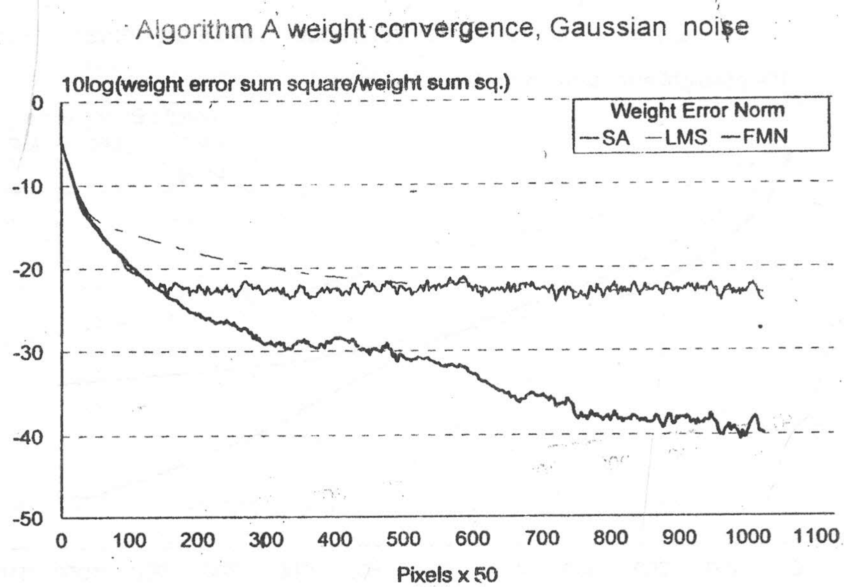


Figure 4 The weight error norm. Desired input corrupted by Gaussian noise of variance 9

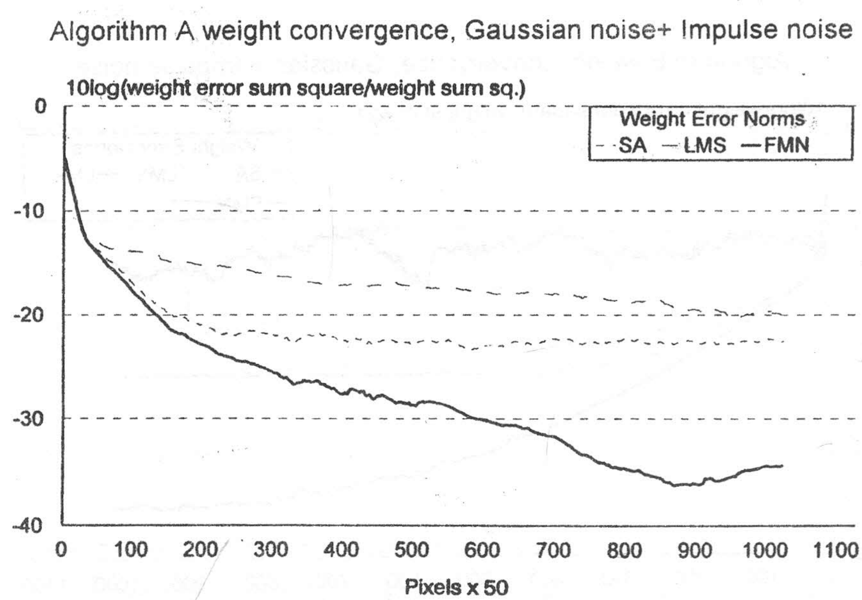


Figure 5 The weight error norm. Desired input corrupted by Gaussian noise of variance 9 + impulsive Noise

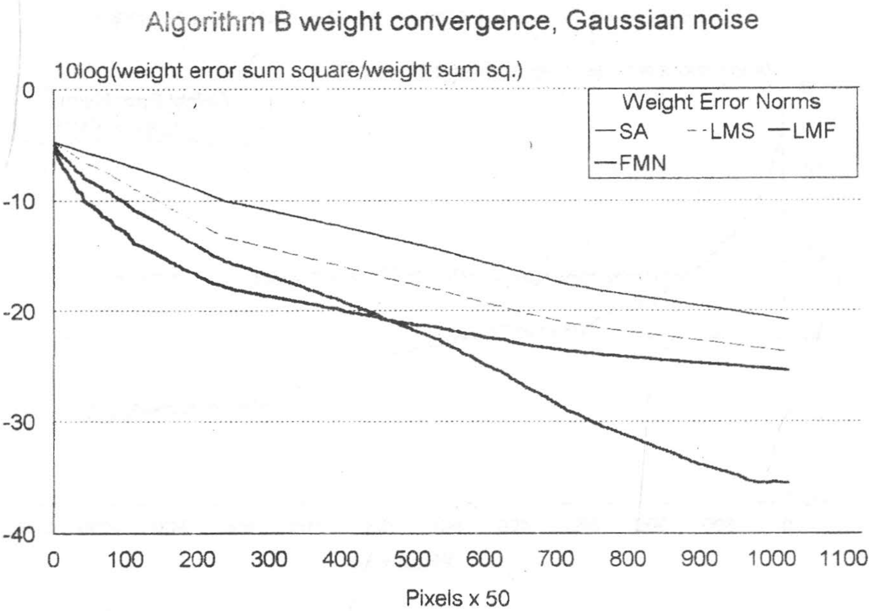


Figure 6 The weight error norm. Desired input corrupted by Gaussian noise of variance 9

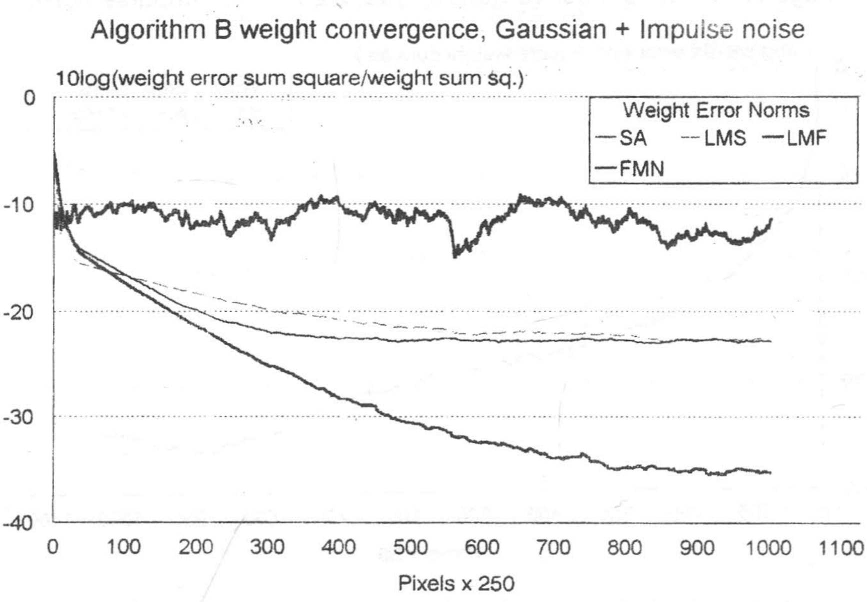


Figure 7 The weight error norm. Desired input corrupted by Gaussian noise of variance 9 + impulsive Noise

CONCLUSIONS

A systematic method to the design of fuzzy-logic based adaptive algorithm, which minimizes a generalized mixed norm cost function, has been introduced. The potential of the proposed method is illustrated by designing two fuzzy mixed-norm algorithms. The proposed fuzzy-logic approach can actually be used for a variety of applications. For example, soft transitions among parallel algorithms, variable structure algorithms or algorithms with different window sizes.

The performance has been demonstrated in system identification simulation problem. Results clearly indicate that the FMN algorithm converges to true parameters with very small steady state weight error in the presence of Gaussian and impulsive noise.

REFERENCES

1 Jonathon Chambers, and Apostolos Avlonitis, "A robust Mixed-Norm Adaptive Filter Algorithm", IEEE Signal Processing Letters, Vol. 4, No. 2, pp. 46, (1997)

2 S. Alliney and R. A. Ruzinsky, "An Algorithm for the Minimization of Mixed l_1 and l_2 Norms with Application to Bayesian Estimation", IEEE trans. on Signal Processing, Vol. 42, No. 3, pp. 618, (1994)

3 T. Takagi and M. Sugeno, "Fuzzy Identification of Systems and its Application to Modelling and Control", IEEE Trans. Systems Man Cybern., Vol. 15, pp. 1116, (1985).

4 J. A. Chambers, O. Tanrikulu, and A. G. Constantantinides, "Least Mean Mixed-Norm adaptive Filtering", Electronics Letters, Vol 30, pp. 1574-1575, (1994).

5 Simon Haykin", Adaptive filter theory, Second Ed., Prentice-Hall, (1991).

Received July 10, 1999
Accepted October 19, 1999

طريقه لتصميم خوارزمات المرشحات المتكيفة مختلطة المقياس باستخدام المنطق الشامل

محيى محمد هدهود* و عبد العظيم صبيح إبراهيم**

* قسم هندسة الاتصالات الكهربية - جامعة المنوفية

*قسم هندسة التحكم و القياسات - جامعة المنوفيه

ملخص البحث:

يقدم هذا البحث طريقه لتصميم الخوارزمات المتكيفة مختلطة المقياس (**NORMS**) باستخدام المنطق الشامل (**FUZZY LOGIC**) حيث يتم الاستفادة من هذه الطريقة لتجميع مميزات اى عدد من الخوارزمات المتكيفة في خوار يزم واحد. ويتم ذلك عن طريق الانتقال الناعم بين المرشحات المختلفة بناء علي الخواص المحلية للدخل.
تم تصميم دالة للتعبير عن الدخل في فراغ المطابقة. و تم تصميم مقياس للخواص المحلية للإشارة باستخدام دالة العضوية المعرفة في فراغ المطابقة. و لزيادة كفاءة الخوار يزم تم عمل تنعيم لدالة العضوية أثناء التشغيل.
الخوار يزم المقترح تم استخدامه لتصميم خوار يزم رقمي متكيف من مجموعة من الخوارزمات المتكيفة مختلفة المقياس و تجميع مميزات عدد من الخوارزمات المتكيفة في خوار يزم واحد له اقل خطأ في حالة الاستقرار.