Fuzzy based approach to the Recognition of Multi-Font Numerals

Madasu Hanmandlu
Department of Electrical Engg,
I.I.T. Delhi,
New Delhi – 110016, India.
mhmandlu@ee.iitd.ernet.in

Mohd. Hafizuddin Mohd. Yusof
Faculty of Information Technology,
Multimedia University,
Cyberjaya - 63100, Selangor, Malaysia.
hafizuddin.yusof@mmu.edu.my

Vamsi Krishna Madasu
School of ITEE,
University of Queensland,
QLD 4072, Australia.
madasu@itee.uq.edu.au

Abstract- *In this paper, we propose a new scheme for off-line recognition of multi-font numerals. In the proposed scheme, the binary image of a character is partitioned into a fixed number of sub-images called boxes. The features consist of normalized vector distances (γ) and angles (α) from each box. These features give rise to fuzzy sets. However, when we have small number of samples as in the case of multi-font numerals, the choice of a proper fuzzification function is crucial. Hence, we have devised a new fuzzification function involving parameters, which take account of the variations in the fuzzy sets. These parameters are obtained by optimization.*

Keywords- *Box approach, Fuzzy optimization, Multi-font Numerals, New Fuzzification function*

1. INTRODUCTION

From time immemorial, the magic of numbers has enchanted researchers and scientists across the globe. The applications of numbers are so varied that there is no point in discussing them further. However the styles of writing the numbers are also different and they come in various sizes, shapes and fonts. The ability to identify these machine printed numbers in an automated or semi automated manner led to the development of an entirely different field of research known as the Optical character recognition (OCR). This has been a topic of interest since possibly the late 1940's when Jacob Rabinow started his work in the field. The earliest OCR machines were primitive mechanical devices with fairly high failure rates. As the amount of new written material increased, so did the need to process it all in a fast and reliable manner, and these machines were clearly not up to the task. They quickly gave way to computer-based OCR devices that could outperform them both in terms of speed and reliability. The earlier computer based OCR systems [1] that had been developed were confined to recognize only the printed or handwritten characters of fixed size and fonts. But, the present study aims at producing a system, which could recognize characters, especially numerals of any arbitrary size, shape and fonts.

There are numerous approaches that address the problem of recognition of numerals depending on the type of features extracted and the different ways of extracting them. Although recognition of handwritten numerals is a well-researched topic and many techniques of recognition of both machine printed and hand printed characters are available in the literature [2,3,4] but not much work has been reported on the recognition of multi-font numerals, in recent times.

Recognition of printed Korean characters of multi-font, multi-size [5] was attempted using a two-stage classification method. Handwritten numeral character recognition using fuzzy logic was first attempted by Siy and Chen [6]. Here the handwritten numeral is decomposed into straight lines, portions of a circle and circles. A more flexible scheme is proposed in [7] where the decomposition is based on the detection of a set of feature points: terminal point, intersection point and bend point.

Handwritten numeral recognition using combined self-organizing maps (SOM) and fuzzy rules was presented in [8]. In this study, SOM algorithm is used in the learning phase to produce prototypes, which together with corresponding variances are used to determine fuzzy regions and membership functions. Learning from training patterns, then the algorithm generates fuzzy rules. In the recognition stage, a fuzzy rule based classifier is employed to classify an input pattern. An unsure pattern is re-classified by a SOM classifier.

A technique in [9] is used to produce fuzzy rules, based on the ID-3 approach optimizing the defuzziffication parameters with help of a two-layer perceptron. This technique overcomes the difficulties in a conventional

syntactic approach for handwritten character recognition including problems of choosing a starting or reference point scaling and learning by machines.

A new scheme for off-line recognition of totally unconstrained handwritten characters using a simple multi-layer cluster neural network trained with the back propagation algorithm is presented in [10] and it is shown that the use of genetic algorithms avoids the problem of finding local minima in training the multi-layer cluster neural network with gradient descent technique and improves the recognition rates. In that scheme, Kirsch masks are adopted for extracting feature vectors and a three-layer cluster neural network with five independent sub networks is developed for classifying similar characters efficiently.

The perturbations due to writing habits and instruments are taken into account in the recognition of off-line handwritten numerals in [11]. A model based segmentation algorithm is proposed in [12] for handwritten character strings. This method is based on the boundary analysis of the character strings where a set of new features is extracted. These features can be used for global interpretation of the boundary structures. Various models are then constructed for the segmentation and for reducing the segmentation time.

The back-propagation neural network is used in [13] for the recognition of handwritten characters. In that, feature extraction is done using three different approaches, namely, ring, sector and hybrid. The features consist of normalized vector distances and angles. The hybrid approach, which combines the ring and sector approaches, was found to yield the best results. The same features are adopted in [14]. We follow the same theme of feature extraction as employed in [14] for our present study. However, this study is meant for the case when we have a small set of samples. For this reason, we have devised a new fuzzification function, which is a modified version of the function used in [14]. This new function contains structural parameters, which are derived so as to maximize the recognition rate and minimize the entropy.

The organization of the paper is as follows: Section 2 briefly discusses the preprocessing steps used. Section 3 deals with the feature extraction. Recognition system is presented in Section 4 and results and conclusions are presented in Sections 5 and 6 respectively.

2. PRE-PROCESSING

The normalization of the numerals is essential because of the different types of fonts, which result in several variations in the shapes and sizes. Therefore, to bring about uniformity among the input numerals, all of them should be made of the same size. For this reason, the numerals are fit into a standard size window of 42 x 32. Every measure has to be taken to preserve the exact aspect ratio of the numerals. This size of window is selected due to the fact that usually height of the numeral is almost 1.5 times more than the width. Experiments show that the normalization of thinned numeral causes the skeleton to be discontinuous. Hence, the un-thinned numeral is normalized and later thinned. The result of thinning constitutes a skeleton of the original pattern, one-pixel wide, consisting of a subset of the original black pixels. This thinned image is then considered for extracting features.

3. FEATURE EXTRACTION

Feature extraction is the crucial phase in numeral identification as each numeral is unique in its own way, thus distinguishing itself from other numerals. Hence, it is very important to extract features in such a way that the recognition of different numerals becomes easier on the basis of the individual features of each numeral.

For extracting the features, we have used the Box-Method approach proposed in [14]. The major advantage of this approach stems from its robustness to small variations, easy to implement and relatively high recognition rate. In this approach, the binary array of 42 x 32 is fit into horizontal and vertical grid lines of 6 x 4. Thus as shown in Fig. 1, 24 boxes of equal size are derived in which the complete binary image is enclosed. Each box will have the size 7 x 8, so that, the portions of a numeral will lie in some of these boxes. There could be boxes that are empty. However, all boxes are considered for analysis in a sequential order. The choice of number of boxes is arrived at by experimentation. By considering the bottom left corner as the absolute origin (0,0), the coordinate distances (Vector Distances) of all the pixels of a box are calculated. By averaging the total of all distances with the number of pixels in a box, a Normalized Vector Distance (γ) for each box is obtained. Therefore, 24 γ's corresponding to 24 boxes will constitute a feature set based on the vector distances. However, for empty boxes, the value will be zero. The Vector Distance is calculated for each pixel at coordinates (i, j) with respect to the origin, using the formula:

$$d_{kb} = (i^2 + j^2)^{1/2} \quad (1)$$

Similarly, for all pixels in a box, the vector distances are computed and normalized with the total number of pixels available in that box using the formula:

$$\gamma_b = \frac{1}{n}\sum_{k=1}^{n} d_{kb} \quad (2)$$

where, n is number of pixels in a box and b is the number of boxes

The angle features are not considered here for recognition, as they are not as much effective as the distance features.

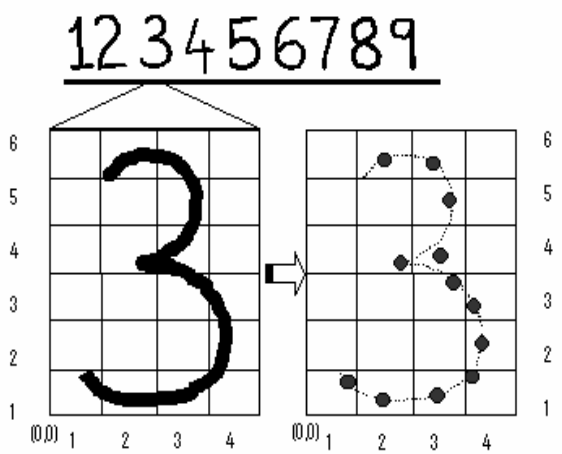


Fig. 1. Numeral '3' enclosed in 6x4 grid

4. RECOGNITION SYSTEM

The extracted features are used to identify multi-font numerals. Therefore, a database of numerals of different fonts is required to extract the features, which form the knowledge base. We have considered here a standard database of multi-font numerals consisting of thirty fonts. Only ten samples of each font are taken into account, as there is not much difference between the individual samples. Thus we have 300 samples in all.

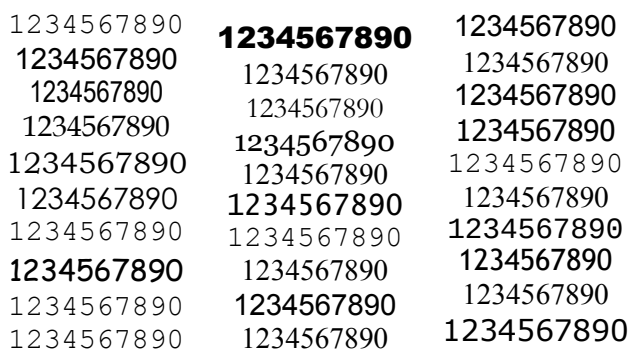


Fig.2. Database of multi-font numerals

The features extracted from the size normalized and thinned binary array, form the input for recognition process. The recognition system is designed so as to optimize the structural parameters in the fuzzification function, in order to enhance the recognition rate.

4.1. Fuzzy Sets and choice of membership function

In order to recognize the unknown numeral set using fuzzy logic, an exponential variant of fuzzy membership function is selected. The fuzzy membership function is proposed in [14] based on the normalized vector distance. The application of fuzzy logic to the recognition of numerals was proposed in [13] and was extensively used in [14]. The concept of fuzzy sets is as follows.

If there are 'n' possible features for each numeral and if there are 'm' such samples, then a particular feature from each of the samples forms a fuzzy cluster. Thus, for each particular numeral, the resultant matrix thus formed is m x n. The reference numeral data set is obtained from training samples. The mean and variances are computed for each of the 24 clusters and taken as Knowledge Base (KB). The procedure is repeated for all reference numerals. Therefore, there are 24 mean and variances corresponding to 24 features for each of the 10 numerals. Given a very large number of samples, by choosing the fuzzification function, the membership function of each feature value in the cluster can be determined. However, we need to compute the membership functions for the features of unknown numerals and not the reference numerals. The unknown numeral features are matched with all reference numeral features available in the form of KB. It is possible to compute the membership functions by associating the features of the input numeral with the clusters. The KB consists of means m_i and variance σ_i^2 for each of the 24 clusters using the formulae:

$$m_i = \frac{1}{N_i}\sum_{j=1}^{N_i} f_{ij} \tag{3}$$

$$\sigma_i^2 = \frac{1}{N_i}\sum_{j=1}^{N_i} (f_{ij} - m_i)^2 \tag{4}$$

where, N_i is the number of samples in the ith cluster and f_{ij} stands for the jth feature value of reference numeral in the ith cluster where (i=1,2…24).

For an unknown input numeral x, the 24 features are extracted using the box method. The membership function is chosen as,

$$\mu_{xi} = \mathrm{e}^{\frac{-|x_i - m_i|}{\sigma_i^2}} \tag{5}$$

where, x_i is the ith feature of the unknown numeral.

If all x_i's are close to m_i's which represent the known statistics of a reference character, then unknown numeral is identified with this known numeral because all the membership functions are close 1 and hence the average membership function is almost 1 as explained below.

Let, $m_j(r), \sigma_j^2(r)$ belong to the rth reference numeral with r = 0,1…9. We then calculate the average membership as,

$$\mu_{av}(r) = \frac{1}{c}\sum_{j=1}^{c} e^{\frac{-|x_j - m_j(r)|}{\sigma_j^2(r)}} \tag{6}$$

Then $x \in r$ if $\mu_{av}(r)$ is the maximum for r = 0,1…9

4.2. Modified Fuzzification functions

The fuzzification function (5) does not perform well when we have a limited database such as the case of multi-font numerals, shown in Fig.2. This is because some of the fuzzy sets have a very small variance and others have a large variance. So, we have devised a scheme to separate the fuzzification functions on the basis of the value of variance. In doing so, our objective is to get larger values for the membership functions.

Accordingly, we split up the equation (5) into the following two exponential functions,

$$\mu_{xi} = e^{\frac{-|x_i - m_i|}{\sigma_i^2}} \qquad \text{For } \sigma_i^2 \succ 1 \qquad (7)$$

$$\mu_{xi} = e^{-\left\{(x_i - m_i)^2 \cdot \sigma_i^2\right\}} \qquad \text{For } \sigma_i^2 \prec 1 \qquad (8)$$

The bifurcation of the membership function into two different functions depending upon the variance is compelled by the fact that some variances are too small and some are too large. This has spurred the choice of a new membership function involving structural parameters s and t are given by,

$$\mu_{xi} = e^{-\frac{\Delta x_i'}{\sigma_i^{2\prime}}} \qquad (9)$$

where,

$$\sigma_i^{2\prime} = (1+t) + t^2\sigma_i^2$$
$$\Delta x_i' = (1-s) + s^2 \Delta x_i$$
$$\Delta x_i = |x_i - m_i|$$

Thus, the new variance and the new mean are the functions of the variance and mean of original clusters. Thus, the structural parameters s, t model the variations in the mean and variance over all 24 clusters (boxes). The choice of these parameters has reasoning. That is, if s=1, $\Delta x_i' = \Delta x_i$. Thus, s would be perturbed around 1 to reflect changes in the means. Similarly, if t=-1, then $\sigma_i^{2\prime} = \sigma_i^2$. Thus, t would reflect the changes in the variances. The relations in (9) are intuitive.

We now present the estimations of these parameters using an optimization approach. The average membership function of any unknown numeral must be close to 1, when its features are fit into the statistics of a known numeral. This will serve as an identification criterion, defined by the equations

$$J = \frac{1}{c}\sum_{j=1}^{c} \mu_{xj} \qquad (10)$$

$$J_1 = [1 - J]^2 \qquad (11)$$

where, c denotes the number of clusters.

In order to reduce the uncertainty in s and t, we define the entropy E as,

$$E = \frac{1}{c \ln 2} \sum -[\mu_{xj} \ln \mu_{xj} + (1 - \mu_{xj}) \ln(1 - \mu_{xj})]$$

So we reduce E subject to the satisfaction of J_1. The objective function to be minimized is therefore chosen as,

$$G = E + \lambda J_1$$

Now, in order to find the parameters s and t, we use the gradient descent technique,

$$s^{new} = s^{old} - \varepsilon_1 \frac{\partial G}{\partial s} \tag{12}$$

$$t^{new} = t^{old} - \varepsilon_2 \frac{\partial G}{\partial t} \tag{13}$$

$$\lambda^{new} = \lambda^{old} - \varepsilon_3 \frac{\partial G}{\partial s} \tag{14}$$

The derivatives in the above equations are given in appendix-A. The learning factors $\varepsilon_1, \varepsilon_2, \varepsilon_3$ are chosen as 0.01. The value of λ is to taken to be 1. The initial values of s and t are taken as 1 and 2 respectively. With the choice of learning factors, the converged values of s, t are obtained as given in Table I.

5. RESULTS

The various types of fonts introduce large variations in the numerals. In this work, we have used 30 fonts. The structural parameters are given in Table I. We have obtained a recognition rate of 96 % with the functions (7) and (8). But with the proposed modification of the membership function (9), the recognition rate has risen to 98.67 as shown in Table II.

Table I: Estimation of Structural parameters

NUMERAL	s	t
0	1.1671	2.1818
1	1.2446	2.1619
2	1.2467	2.1636
3	1.2011	2.1731
4	1.2466	2.1635
5	1.2177	2.1716
6	1.2232	2.1689
7	1.2012	2.1730
8	1.2093	2.1729
9	1.2308	2.1676

Table II: Recognition Rates with function (9)

NUMERAL	RECOGNITION RATE (%)
0	100.0
1	100.0
2	100.0
3	100.0
4	100.0
5	96.67
6	96.67
7	100.0
8	96.67
9	96.67
Overall	**98.67**

The recognition process is illustrated in Appendix-B. The unknown numeral, which is to be recognized, is taken to be 4. The second curve in all the figures shows the membership function of 4 given the statistics of 0,1,2…9. The first curve shows the membership of the known numeral, with given statistics. We observe from the membership function for 4 that the two curves are closer. In this figure, the first curve is the membership function of any known numeral 4 and second curve is the membership function of the unknown numeral. This indicates the identity of the unknown numeral.

6. CONCLUSIONS

In this paper, we have devised a new fuzzification function for the recognition of unconstrained, isolated, multi-font machine printed characters, especially, numerals. The box method has been used for obtaining the distance features. The mean and variances of these features are adapted using two structural parameters, which model their variations across the different fonts and styles of numerals. A recognition rate of almost 99% was achieved with the new fuzzification function. Suggested future work may include testing of the algorithm with more numeral data, consisting of many other fonts, which have not been considered in the present study. The relations between the modified and the actual means and variances involving structural parameters are intuitive. Further work is required to devise a sound framework.

REFERENCES

[1] S. Mori, C.Y. Suen and K. Kamamoto, "Historical review of OCR research and development", *Proc. of IEEE*, vol. 80, pp. 1029-1058, July 1992.

[2] V.K. Govindan and A.P. Shivaprasad, "Character Recognition - A review", *Pattern Recognition, v*ol. 23, no. 7, pp. 671-683, 1990.

[3] C.Y. Suen, C. Nadal, R. Legault, T.A. Mai and L. Lam, "Computer Recognition of Unconstrained handwritten numerals", *Proc. of IEEE*, vol. 80, no. 7, pp. 1162-1180, 1992.

[4] R. Plamondon, N. Srihari, "On-Line and Off-line Hand Writing Recognition: A Comprehensive Survey", *IEEE Trans. on PAMI*, vol. 22, no. 1, pp. 63-84, January 2000.

[5] J.S.Lee, O.J.Kwon and S.Y. Bang, "Highly accurate recognition of printed Korean characters through an improved two stage classification method", *Pattern Recognition*, vol. 32, no. 12, pp. 1935-1945, 1999.

[6] P. Siy and C.S. Chen, "Fuzzy Logic for handwritten numeral character recognition", *IEEE Trans. Systems, Man and Cybernetics*, vol. 4, pp. 570-575, 1974.

[7] G. Baptista and K.M. Kulkarni, "A high accuracy algorithm for the recognition of hand written numerals", *Pattern Recognition*, vol. 21, pp. 287-291, 1998.

[8] Z. Chi, J. Wu and H. Yan, "Handwritten Numeral Recognition Using Self Organizing Maps and Fuzzy Rules", *Pattern Recognition*, vol. 28, no. 1, pp.59-66, 1995.

[9] Z. Chi and H. Yan, "ID-3 Derived Fuzzy Rules and Optimized defuzziffication For Handwritten Numeral Recognition", *IEEE Trans. on Fuzzy Systems*, vol. 4, no. 1, pp.24-31, February 1996.

[10] S.W. Lee, "Off Line Recognition of Totally Unconstrained Handwritten Numerals Using Multilayer Cluster Neural Network", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, no. 6, pp. 48-652, June 1996.

[11] T.M. Ha and H. Bunke, "Off-Line handwritten numeral recognition by perturbation method", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, pp. 535-539, May 1997.

[12] J. Hu and H. Yan, "A model based segmentation method for handwritten character strings", *Computer Vision and Image Understanding*, vol. 70, no. 3, pp. 383-403, 1998.

[13] M. Hanmandlu, K.R.M. Mohan, and H. Kumar, "Neural-based Handwritten character recognition ", in *Proceedings of Fifth IEEE International Conference on Document Analysis and Recognition*, *ICDAR'99*, pp. 241-244, Bangalore, India, 1999.

[14] M. Hanmandlu, K.R.M. Mohan, S. Chakraborty, S. Goyal and D. Roy Choudhury, "Unconstrained handwritten character recognition based on fuzzy logic", *Pattern Recognition*, vol. 36, no. 3, pp. 603-623, March 2003.

APPENDIX-A

The entropy, which indicates the uncertainty in the parameters in s and t is

$$E = \frac{1}{c \ln 2} \sum -[\mu_{xj} \ln \mu_{xj} + (1-\mu_{xj}) \ln(1-\mu_{xj})] \qquad \text{(A.1)}$$

The derivatives E with respect to s and are:

$$\frac{\partial E}{\partial s} = \frac{\partial E}{\partial \mu_j} \cdot \frac{\partial \mu_{xj}}{\partial s} \qquad \text{(A.2)}$$

$$\frac{\partial E}{\partial t} = \frac{\partial E}{\partial \mu_j} \cdot \frac{\partial \mu_{xj}}{\partial t} \qquad \text{(A.3)}$$

$$\frac{\partial E}{\partial \mu_{xj}} = \frac{1}{c \ln 2} \sum ln \left[\frac{1-\mu_{xj}}{\mu_{xj}} \right] \qquad \text{(A.4)}$$

But, our aim is to not only reduce the uncertainty but to pave the way for the identification of the unknown numeral. So, we have

$$G = E + \lambda J_1 \tag{A.5}$$

Its derivatives with respect to s, t and λ are

$$\frac{\partial G}{\partial s} = \frac{\partial E}{\partial s} + \lambda \frac{\partial J_1}{\partial s} \tag{A.6}$$

$$\frac{\partial G}{\partial t} = \frac{\partial E}{\partial t} + \lambda \frac{\partial J_1}{\partial t} \tag{A.7}$$

$$\frac{\partial G}{\partial \lambda} = J_1 \tag{A.8}$$

The derivative of J_1 with respect to J is

$$\frac{\partial J_1}{\partial J} = 2[1 - J] \tag{A.9}$$

The derivatives of J with respect to s and t are as follows:

$$\frac{\partial J}{\partial s} = \frac{1}{c} \sum \frac{\partial \mu_{xj}}{\partial s} \tag{A.10}$$

$$\frac{\partial J}{\partial t} = \frac{1}{c} \sum \frac{\partial \mu_{xj}}{\partial t} \tag{A.11}$$

Therefore,

$$\frac{\partial J_1}{\partial s} = \frac{\partial J_1}{\partial J} \cdot \frac{\partial J}{\partial s} = 2[1 - J] \cdot \frac{1}{c} \sum \frac{\partial \mu_{xj}}{\partial s} = \frac{2}{c}[1 - J] \sum \frac{\mu_{xj}[1 - 2s\left|x_j - m_j\right|]}{(1+t) + t^2 \sigma_j^2} = \frac{2}{c}[1 - J] \sum_{j=1}^{c} F_{sj} \tag{A.12}$$

$$\frac{\partial J_1}{\partial t} = \frac{\partial J_1}{\partial J} \cdot \frac{\partial J}{\partial t} = 2[1 - J] \cdot \frac{1}{c} \sum \frac{\partial \mu_{xj}}{\partial t} = \frac{2}{c}[1 - J] \sum \frac{\mu_{xj}[(1-s) + s^2\left|x_j - m_j\right|][1 + 2t\sigma_j^2]}{[(1+t) + t^2 \sigma_j^2]^2}$$

$$= \frac{2}{c}[1 - J] \sum_{j=1}^{c} F_{tj} \tag{A.13}$$

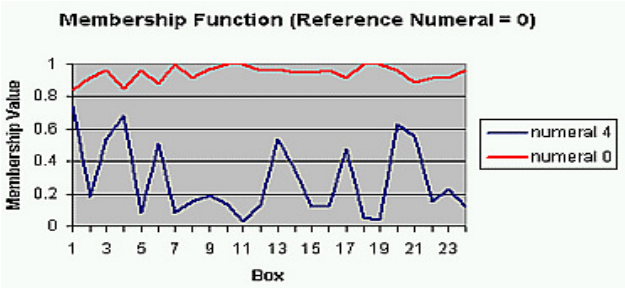
The equations (A.6) and (A.7) are simplified as,

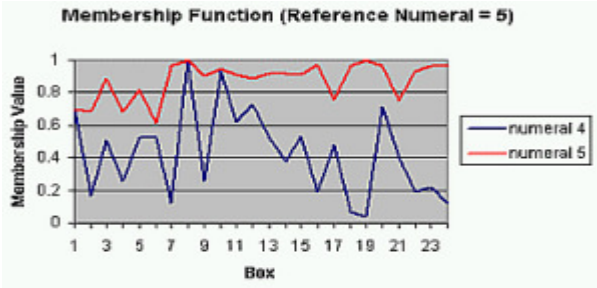
$$\frac{\partial G}{\partial s} = \frac{1}{c \ln 2} \sum_{j=1}^{c} E_t F_{sj} + \frac{2\lambda[1 - J]}{c} \sum_{j=1}^{c} F_{sj} \tag{A.14}$$

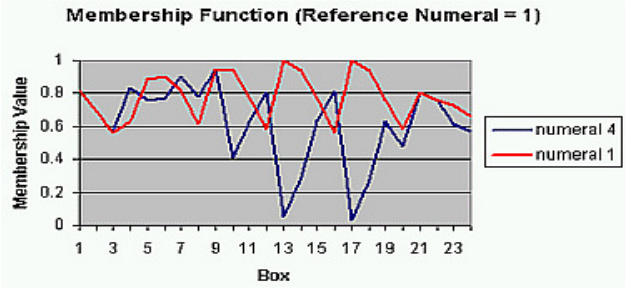
$$\frac{\partial G}{\partial t} = \frac{1}{c \ln 2} \sum_{j=1}^{c} E_t F_{tj} + \frac{2\lambda[1 - J]}{c} \sum_{j=1}^{c} F_{tj} \tag{A.15}$$

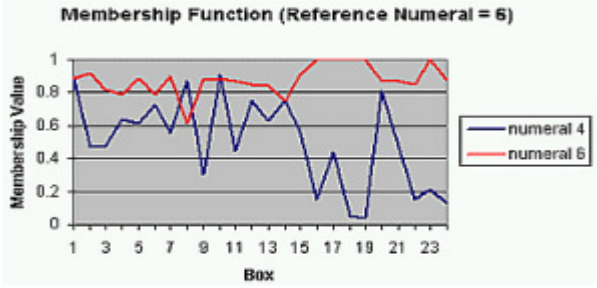
where, $E_t = \dfrac{(1-s) + s^2 \Delta x_j}{c \ln 2 (1 + t + t^2 \sigma_j^2)}$

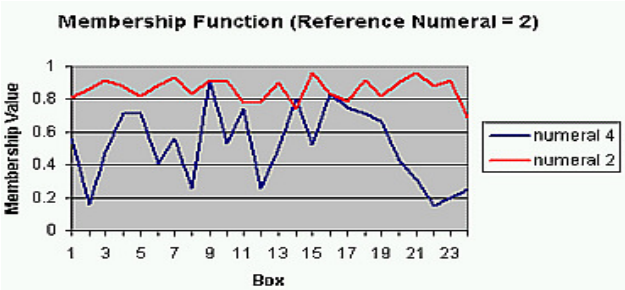
APPENDIX – B

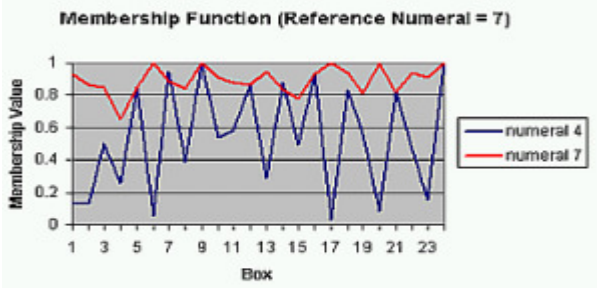

Membership Function (Reference Numeral = 0)
Membership Value
Box
numeral 4
numeral 0

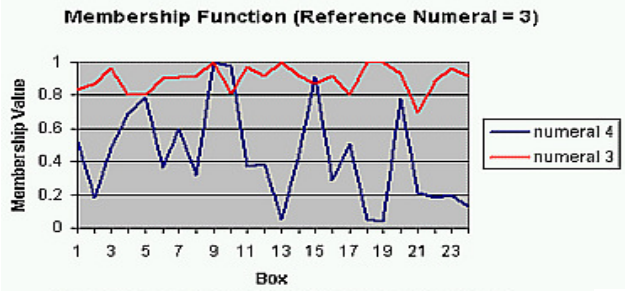

Membership Function (Reference Numeral = 5)
Membership Value
Box
numeral 4
numeral 5

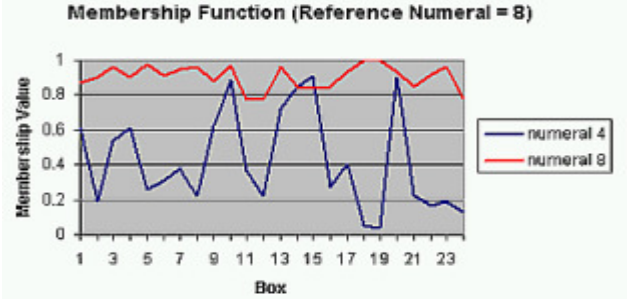

Membership Function (Reference Numeral = 1)
Membership Value
Box
numeral 4
numeral 1

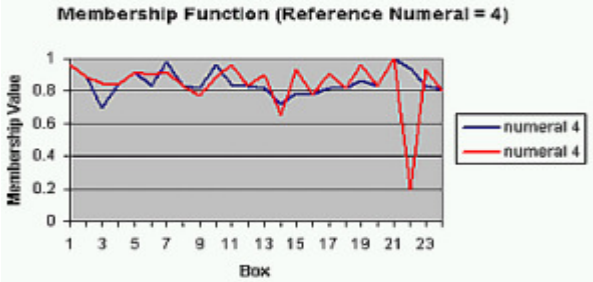

Membership Function (Reference Numeral = 6)
Membership Value
Box
numeral 4
numeral 6

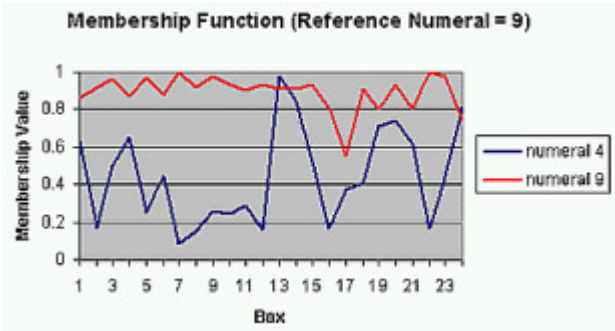

Membership Function (Reference Numeral = 2)
Membership Value
Box
numeral 4
numeral 2


Membership Function (Reference Numeral = 7)
Membership Value
Box
numeral 4
numeral 7


Membership Function (Reference Numeral = 3)
Membership Value
Box
numeral 4
numeral 3


Membership Function (Reference Numeral = 8)
Membership Value
Box
numeral 4
numeral 8


Membership Function (Reference Numeral = 4)
Membership Value
Box
numeral 4
numeral 4


Membership Function (Reference Numeral = 9)
Membership Value
Box
numeral 4
numeral 9