International Journal of Engineering Innovation & Research
Volume 2, Issue 6, ISSN: 2277 – 5668

Recognizing the Hand Written Characters

D. Kiran Babu, M. Lakshmi Prasanna, P. Suresh Babu, A. Balakrishna
Department of CSE, Kaushik College of Engineering, Visakhapatnam, Andhra Pradesh, India

***Abstract* – In zoning-based classification, a membership function defines the way a feature influences the different zones of the zoning method. This paper presents a new class of membership functions, which are called fuzzy-membership functions (FMFs), for zoning-based classification. These FMFs can be easily adapted to the specific characteristics of a classification problem in order to maximize classification performance. In this research, a real-coded genetic algorithm is presented to find, in a single optimization procedure, the optimal FMF, together with the optimal zoning described by Voronoi tessellation. The experimental results, which are carried out in the field of handwritten digit and character recognition, indicate that optimal FMF performs better than other membership functions based on abstract-level, ranked-level, and measurement-level weighting models, which can be found in the literature.**

***Keywords* – Fuzzy-Membership Function (FMF), Genetic Algorithm, Handwritten Character Recognition, Voronoi Tessellation, Zoning Method.**

I. INTRODUCTION

In the field of handwritten character recognition, zoning offers an easy way to obtain spatial information on feature distribution. There-fore, zoning has been largely considered to be a useful technique to handle the enormous variability of handwritten characters, due to dif-ferent handwriting styles and personal variations [1]. Given a pattern image B, a zoning method $Z_M = \{z_1 , z_2 , . . . , z_M \}$ is a partition of B into M subimages, which are called zones, that can each provide information related to a specific part of the pattern [2].

In the past, the static zoning methods used could be obtained by superimposing regular $m \times n$ grids on a pattern image. In this case, the most-profitable zoning for a given classification problem was selected on the basis of a designer's personal experience and on experimental tests. More recently, dynamic zoning methods have been introduced in which zoning design is considered as an optimization problem. In [3], a multiobjective evolutionary algorithm was used to define the zoning method based on two diverse optimality criteria: a minimal number of nonoverlapping zones and a minimal error rate. Gagne´ and Parizeau [4] used a hierarchical zoning and a genetic technique to opti-mize the feature-extraction step of a handwritten character recognizer. Genetic programming has also been used to simultaneously learn the best hierarchy and the best combination of fuzzy features. Impedovo *et al.* [5] identified the optimal zoning as the one which optimized the classification performance, estimated by the cost function (CF) that is associated with classification. In this case, a well-suited zoning rep-resentation was considered based on Voronoi tessellation [6]. Strictly speaking, the Voronoi tessellation is the partition of a plane into M zones $z_1 , z_2 , . . . , z_M$ that reflects proximity relationships among the given set of M distinct points $p_1 , p_2 , . . . , p_M$ in the plane. In other words, each point p_i determines a region z_i that is the locus of points that are closer to p_i than to any other point in the set. Since changing the position of the Voronoi points causes modification of the zoning methods, Voronoi tessellation allows the zoning to be easily adapted to the specific characteristics of the classification problem.

Although zoning methods have been extensively dealt with in the literature, so far little attention has been paid to the definition of feature-zone membership functions, i.e., functions that determine the way in which a feature influences the different zones of a zoning method [7]. Standard membership functions use static weighting models based on abstract-level, ranked-level, and measurement-level strategies [5], [7]. Unfortunately, they are unable to adapt themselves to the specific characteristics of feature distributions in order to achieve optimal performances.

This paper introduces a new class of fuzzy-membership functions (FMFs) and presents a real-coded genetic algorithm to detect, in a single optimization procedure, the optimal FMF, together with the op-timal zoning, for the specific classification problem. The experimental tests were carried out in the field of handwritten numeral and character recognition, using datasets from the databases of the Center of Ex-cellence for Document Analysis and Recognition (CEDAR), Buffalo, NY, and the Electrotechnical Laboratory (ETL), Tsukuba, Japan. The results demonstrate that FMFs can give better performance than other membership functions described in the literature. Additionally, they confirm the effectiveness of the genetic technique for the combined selection of the optimal zoning and optimal FMF.

The organization of this paper is as follows. Section II describes the problem of zoning-based classification, while Section III reports the standard membership functions and introduces the new class of FMFs. Section IV shows the genetic algorithm for the combined se-lection of the optimal zoning, together with the optimal FMF, and Section V reports the experimental results. The conclusion is reported in Section VI.

II. ZONING-BASED CLASSIFICATION

Given the set of class labels $\Omega = \{C_1 , . . . , C_K \}$ and the rejection class C , the classification of a pattern x can be considered to be a mapping D, which is defined as [5]

$$D : S(x) \rightarrow \Omega \cup \{C \} \qquad (1)$$

with $S(x)$ being the description of x based on the features set $F = \{f_1 , . . . , f_T \}$. Moreover, when a zoning $Z_M = \{z_1, z_2$

$, \ldots, z_M \}$ is considered, $S(x)$ can be described as [7]

$$S(x) = \{(f_{i_q}, z_j, w_{i_q j}) | q = 1, 2, \ldots Q, j = 1, 2, \ldots, M\} \quad (2)$$

where $(f_{i_q}, z_j, w_{i_q j})$ means that an instance of feature f_{i_q} has been detected in pattern x, and the influence of this instance on zone z_j is equal to $w_{i_q j}$ (with $i_q \in \{1, 2, \ldots, T\}$, $j \in \{1, 2, \ldots, M\}$), and Q is the total number of features detected in x.

In this case, let $X^L = \{x^L_r \mid r = 1, 2, \ldots, N^L\}$ be the learning set, and let $S(x^L_r)$ the description set of x^L_r, for $r = 1, 2, \ldots, N^L$. The following statistical distributions for each class C_k are estimated in the learning phase.

1) The normalized total weight (NTW): This is defined as

$$NTW^k(f, z_i) = \frac{1}{n^k} \sum_{\substack{(f_i, z_j, \mu_{ij}) \in s(x^L_r) \\ i = 1, 2 \ldots, T,\ j = 1, 2 \ldots, M}} W_{ij} \quad (3)$$

with n^k being the number of learning patterns belonging to C_k
(i.e., $n^k = \text{card}\ \{x_r^L \in X^L | x_r^L \in C_k\}$).

2) The normalized relevance (NR): This is defined as

$$\mathrm{NR}^k(f_i, z_j) = \frac{NTWk(f_i, z_j)}{\sum_{k=1,2,\ldots k} NTW^k(f_i, z_j)}$$

Where,

$$\mathrm{NR}^k(f_i, z_j) = 0 \text{ if } \sum_{k=1,2,\ldots,K} NTW^k(f_i, z_j)$$

Unknown pattern x_t (with $i^t q \in \{1,2\ldots.T\}$). The confidence value that x_t belongs to the class C_k is given by

$$\mathrm{Score}_{ck}(x_t) = \sum_{(f^t_{i\,q}, z_j, w^t_{i\,q} z_j) \in S(x_t)} w_i^t q, j\ \ \mathrm{NR}^k(f^t_{i\,q}, z_j) \quad \ldots\ldots(5)$$

The pattern x_t is classified as belonging to the class C_k, if and only if

$$\frac{\mathrm{Score}_{Ck^*}(x_t) - \mathrm{Score}_{Ck^{**}}(x_t)}{\mathrm{Score}_{Ck^*}(x_t)} > \epsilon \quad \ldots(6)$$

With

$K^* = \arg\max_{k \in \{1,2\ldots.K\}} \mathrm{Score}_{Ck}(x_t)$ and $k^{**} = \arg\max_{k \in \{1,2\ldots.K\},\ k \neq k^*} \mathrm{Score}_{Ck}(x_t)$

and $\in \{1,2\ldots K\}$

III. New Class of Fuzzy-Membership Functions

For whatever zoning method considered, membership functions play a fundamental role in zoning-based classification since they determine the way in which a feature influences the different zones of the zoning method. In general, the degree of influence an instance of a feature f has on a zone z is defined by a weight that only depends on relative po-sitions between the instance of the feature and the zone, and therefore, a membership function is described by a set of weights $W = \{\mu_1, \mu_2, \ldots, \mu_M\}$ defined according to proximity criteria [7]. In particular, let $Z_M = \{z_1, z_2, \ldots, z_M\}$ be the zoning method defined by the set of Voronoi points $P = \{p_1, p_2, \ldots, p_M\}$, and let q be the position of feature f (it is worth noting that the position q of a feature f is assumed to be located at the center of gravity of f when structural features are considered, such as lines, loops, cavities, arcs, etc.). In order to intro-duce the membership functions proposed so far in the literature, let us define the *ranked index sequence* (*RIS*) as

$$\mathrm{RIS} = i_1, i_2, \ldots, i_m, i_{m+1}, \ldots, i_M \quad (7)$$

So that

$$i_m \in \{1, 2, \ldots, M\} \quad \forall m = 1, 2, \ldots, M$$
$$i_m = i_n \quad \forall m, n = 1, 2, \ldots, M, \quad m = n$$

Abstract-level membership functions: In this the degree of influence of each zone is defined by Boolean weights as follows

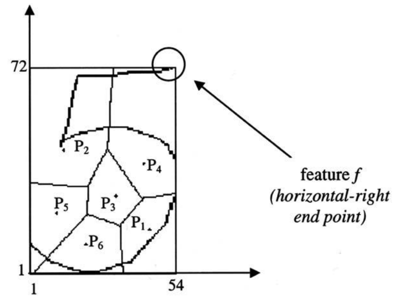


Fig.1. Voronoi-based zoning Z_6 (RIS = _4, 2, 3, 1, 5, 6_).

a) Winner-takes-all (WTA), i.e.,

$$\mu_m = 1, \text{ if } m = i1$$
$$\mu_m = 0, \text{ otherwise.} \quad \ldots..(8)$$

b) k-Nearest zone (k-NZ), i.e.,

$$\mu_m = 1, \text{ if } m \in \{i_1, i_2, \ldots, i_k\}$$
$$\mu_m = 0, \text{ otherwise.} \quad \ldots..(9)$$

Note that for $k = 1$, the 1-NZ function is equal to the WTA membership function.

Ranked-level membership functions:

In this, the degree of influ-ence of each zone is defined by integer weights as follows.

a) Ranked-based (R), i.e.,

$$\mu_m = M - m \quad \ldots..(10)$$

Measurement-level membership functions:

In this, the degree of influence of each zone is defined by real weights as follows.

a) *Linear (L), i.e.,*

$$\mu_m = \frac{1}{d_{im}} \quad (11)$$

b) Quadratic (Q), i.e..,

$$\mu_m = \frac{1}{d^2_{im}} \quad (12)$$

c) Exponential (E). i.e..,

$$\mu_m = \frac{1}{T^{p.\dim}} \quad (13)$$

In addition, a new class of membership functions, which is named FMF, is introduced in this paper for zoning-based classification. An FMF is weighting function for which

International Journal of Engineering Innovation & Research
Volume 2, Issue 6, ISSN: 2277 – 5668

$\mu m \geq 0,\ m=1,2,\ldots M$ ……(14a)

$\mu m \geq 0,\ m=1,2,\ldots M-1$ ……(14b)

$\mu m \geq 0,\ m=1,2,\ldots M-1$ ……(14c)

IV. Zoning Design Using Fuzzy-Membership Functions

In the literature, the problem of optimal zoning design for handwritten character recognition was recently defined as follows [5],[7].

(1) Find the set of optimal voronoi points $P^*=\{p_1^*, p_2^*, \ldots p_j^* \ldots p^*_M\}$ so that

$$CF(Z^*) = \min_{\{2\}} CF(Z)$$

With

(a) $z=\{z_1,z_2\ldots,z_j,\ldots z_m\}$, where z_j is the voroni region rlated to p_j $\forall$ J =1,2…..,M;

(b) $Z^* =\{z^*1,z^*2,\ldots,z^*j,,,,z^*m\}$, where z_j^* is the optimal voronoi region related to p_j^* $\forall$ j=1,2…..M, and the CF was defined as [8]

CF = n error rate + rejection rate (15)

With [9]

(a) Error Rate = card$\{x_r \in X^L \mid x_r \in C_k$ and $D(S(x_r)) = C_{k^*},\ k=k^*\}$/card$(X^L)$, i.e., the percentage of patterns misclassified by the zoning-based classifiers;

(a) Rejection Rate = card $\{x_r \in X^L \setminus D(S(x_r)) = C\}$/ card$(X^L)) = C\}$ /card (X^L), i.e., the percentage of patterns rejected by the zoning-based classifier.

(c) η being the cost value associated with the treatment of an error with respect to a rejection.

It is worth noting that the error rate and the rejection rate of the zoning-based classifier are estimated using the patterns of the learning set, according to the decision rule of(6) In this paper, the design of the optimal zoning is considered along with the design of the optimal FMF. In this case, the optimization problem concerning the combined selection of the optimal FMF, together with the optimal zoning method, as represented by voronoi tessellation, is defined as follows.

(2) Find the set of optimal weitths $w^*=\{\mu^*1,\mu^*2\ldots,\mu^*m\ldots\mu^*_M\}$ together with the set of optimal voroni points $p^*=\{p^*1,p^*2,\ldots p^*_{j\ldots} p_M\}$ so that

$$CF(Z^*,FMF^*) = \min_{\{z,FMF\}} CF(Z,FMF)$$

With

a) $FMF = \{\mu_1,\mu_2\ldots\ldots,\mu_m,\ldots\mu M\}$, where μ_m is the weight for the zone in the mth position of RIS $\forall$ m=1,2….M;

b) $FMF^*= \{\mu_1,\mu_2\ldots\ldots,\mu_m,\ldots\mu M\}$, where μ^*_m is the optimal weight for the zone in the mth position of RIS $\forall$ m=1,2,….M;

c) $Z=\{z_1,z_2,\ldots\ldots z^*_j\ldots z^*_m\}$, where z_j is the voronoi region related to p_j $\forall$ j=1,2…..,M;
$Z^*=\{z_1,z_2,\ldots\ldots z^*_j\ldots z^*_m\}$, where z^*j is the optimal vornoni region related to p^*_j $\forall$ = 1,2….M;

and where the CF was defined according to …(16)

What follows is a description of the real-coded genetic technique used for the optimization problem, which concerns the combined selection of the optimal FMF, together with the optimal zoning method ….. (17)

Step 1: the initial population pop = $\{\Phi1, \Phi2, \ldots \Phi i\ldots \Phi_{Npop}\}$ was created by generating N_{pop} random individuals (N_{pop} even), with

$$\Phi i = (P,W)$$

Where

1) $P = (p_1,p_2\ldots p_j\ldots pM)$ is the set of M voronoi points, with $p_j = (x_j, y_j)$ being the voronoi point of the zone z_j of $z_M = \{z_1, z_2,\ldots z_j\ldots z_M\}$;

2) $W = (\mu1, \mu2,\ldots \mu_m,\ldots, \mu_M)$ is the set of weights of the FMF and μ_m, is the weight value associated with the zone in the m mt position of the RIS.

Step 2: the fitness value of an individual $\Phi_i=(P,W)=\{(p1,p2\ldots p_j\ldots,p_m)\}$ was taken as the classification cost CF(Z,FMF), which was obtained by (16), where z is the voronoi tessellation corresponding to the se p, and FMF is the FMF of weights w.

Step 3: From the initial population, the new population of individuals were generated by the genetic operators of(A) selection , (B) crossover, (C), mutation and (D) elitism[10] as following:

$((p_1^a,p_2^a,\ldots p^a_j\ldots p^a_M), (\mu_1^a,\mu_2^a\ldots u_m^a\ldots u^a_M))$

and

$((p_1^b,p_2^b,\ldots pb_j\ldots pb_M), (\mu_1^b,\mu_2^b\ldots u_m^b\ldots u^b_M))$

be two individuals selected for crossover, then the two offspring individuals

$((p_1^{\sim a},p_{21}^{\sim a},\ldots p_1^{\sim a}{}_{j\ldots}p_1^{\sim a}{}_M), (\mu_1^{\sim a},\mu_2^{\sim a}\ldots u_m^{\sim a}\ldots u_1^{\sim a}{}_M))$

and

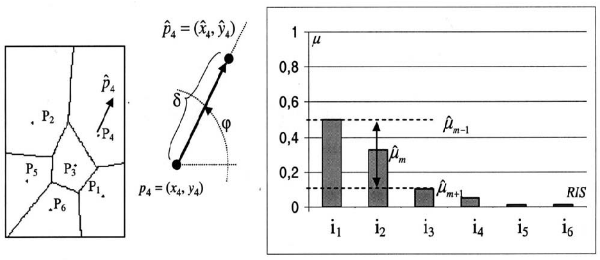
$((p_1^{\sim b},p_2^{\sim b},\ldots p^{\sim b}{}_{j\ldots}p^{\sim b}{}_M), (\mu_1^{\sim b},\mu_2^{\sim b}\ldots u_m^{\sim b}\ldots u^{\sim b}{}_M))$

Of the next generation were obtained as a linear combination of the parent individuals, according to the random values $\alpha,\beta \in [0,1]$:

$p_j^{\sim a} = \alpha.p_j^{\sim a} + (1-\alpha).p_j^{\sim b}{}_j,\ p_j^{\sim b} = \alpha.p_j^{\sim b} + (1-\alpha).p_j^a$

and

$\mu_m^{\sim b} = \beta.\mu_m^b + (1-\beta).\mu_m^b,\ \mu_m^{\sim b} = \beta.\mu_m^{\sim b} + 1-\beta).\mu_m^a$



(a) (b)

Fig.2. Mutation operator

(C): Given an individual $\Phi_i = \{P,W\} = \{(p_1,p_2,\ldots p_j\ldots,p_m), (\mu_1, \mu 2_{\ldots}\ \mu_{m,\ldots}\ \mu_M)\}$, the mutation operator used tow distinct procedures for $P=(p_1,p_2,\ldots p_j\ldots,p_m)$ and $W=(\mu_1, \mu 2_{\ldots}\ \mu_{m,\ldots}\ \mu_M)$.

Let $p_j = (x_j,y_j)$ be an element of P selected for mutation, according to mutation probability p_Mut_Prob. The nouniform mutation operator changed $p_j = (x_j,y_j)$ into the new element $p_j = (x'_j, y'_j)$ that could then be defined as follows [see Fing.2(a)]:

$$\left.\begin{aligned} x'_j &= x_j+\delta.\cos(\mu) \\ y'_j &= y_j+\delta.\cos(\mu) \end{aligned}\right\} \qquad \ldots.(18)$$

Where μ is a random value generated according to a uniform distribution and $\mu \in [0,2\prod]$ and δ is a displacement determined according to the following equation:

$$\delta=\delta\_displ.(1-v^{[1-(iter/N\ iter)\}b}) \qquad \ldots.(19)$$

Being *V a* random value generated in the rage [0, 1] according to uniform distribution, δ_displ is the maximum displacement allowed, b is a parameter determining the degree of non-uniformity, iter is the counter of generations performed, and N^{iter} is the maximum number of generation. It is worth noting that, in (19), when the inter was small, it initially caused the operator to search the space almost uniformly and then locally during the later stages[5],[10].

Simlilarly, by letting be an element of W selected of mutation, according to a mutation probability μ_Mut_Prob. The nonuniform mutation operator changed μ_m, with

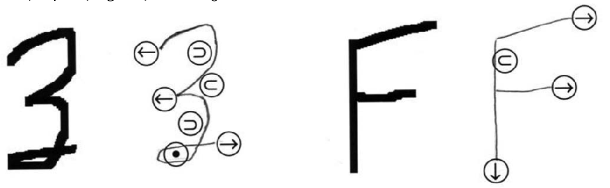
$\mu'_m=\mu 1+9\text{-}1)^e.\eta.[1-\mu 2]$,if m=1 …(20a)
$\mu'_m=\mu 1+9\text{-}1)^e.\eta.[\mu_m\text{-}1 - \mu_m+1]$,if m=2,3…M-1 …(20b)
$\mu'_m=\mu 1+9\text{-}1)^e.\eta.[\mu_m\text{-}1 - \mu_m+1]$,if m=2,3…M-1 …(20b)
$\mu'_m=\mu_M+(-1)^e.\eta.[\mu_M\text{-}1]$, if m=M …(20c)

Where e is a random Boolean value generated according to a equally distributed probability function, and n is a random value generated in the range [0,1], according to uniform distribution. In addition, the following rules were used to guarantee that the membership values μ'_m are positive, monotonically decreasing, and normalized [according to (14)].

1) Positivity:
   If $\mu'_M < 0$, then $\mu'_M = 0$
2) Monotonicity:
   If $\mu'_m < \mu'_{m+1}$, then $\mu'_{m=} \mu'_{m+1}$,
   For m=M-1 , M-2,…1… (21b)
3) Total energy (with $T_e=\Sigma^M_{m=1}\mu'_m$):

$$\mu'_m = \frac{\mu'_m}{T_e}\text{, fro m=1,2,….M} \qquad \ldots(21c)$$

Fig.2(b) shows an example of FMF values after mutation. Since RIS ={4,2,3,1,5,6} we have $i_{1=}$ 4, i_2=2, i_3=3, i_4=1, i_5=5, and i_6=6.



(a) (b) (a) (b)

Fig.3. Examples of feature extraction

(**D**): The elitism strategy substituted one random individual of the current population with the best individual of the previous population[10].

In the genetic process. Steps 2 and 3 were repeated until at least one of the following stop conditions occurred.

1) N^{iter} successive populations were generated.
2) No improvement was registered in L successive populations.

When a stop condition occurred, the process stopped, and the optimal FMF along with the optimal zoning was obtained form the best individual of the last-generated population.

V. Proposed System

1. More recently, dynamic zoning methods have been introduced in which zoning design is considered as an optimization problem.
2. This paper introduces a new class of fuzzy-membership functions (FMFs) and presents a real- coded genetic algorithm to detect, in a single optimization procedure, the optimal FMF, together with the optimal zoning, for the specific classification problem.
3. The experimental tests were carried out in the field of handwritten numeral and character recognition, using datasets.
4. The proposed system confirms the effectiveness of the genetic technique for the combined selection of the optimal zoning and optimal FMF

VI. Modules

1) Training Phase
2) Testing Phase
3) Symbol image detection
4) Symbol image matrix mapping

VII. Modules Description

Training Phase has various functionalities such as:

1. Analyze image for characters
2. Convert symbols to pixel matrices
3. Retrieve corresponding desired output character and con vert to Unicode
4. Lineraize matrix and feed to network
5. Compute output
6. Compare output with desired output Unicode value and compute error
7. Adjust weights accordingly and repeat process until preset number of iterations.

Testing Phase

Testing Phase has various functionalities such as:

1. Analyze image for characters
2. Convert symbols to pixel matrices.
3. Compute output
4. Display character representation of the Unicode output.

VIII. Symbol Image Detection

The process of image analysis to detect character symbols by examining pixels is the core part of input set preparation in both the training and testing phase.

Symbolic extents are recognized out of an input image file based on the color value of individual pixels, which for the limits of this project is assumed to be either black RGB (255,0,0,0) or white RGB (255, 255, 255, 255). The input images are assumed to be in bitmap form of any resolution which can be mapped to an internal bitmap object in the Microsoft Visual Studio environment. The procedure also assumes the input image is composed of only characters and any other type of bounding object like a boarder line is not taken into consideration.

IX. Algorithm

The training routine implemented the following basic algorithm

1. Form network according to the specified topology parameters
1) Initialize weights with random values within the specified
2) load trainer set files (both input image and desired output text)
3) analyze input image and map all detected symbols into linear arrays
4) Form network according to the specified topology parameters
5) 2.Initialize weights with random values within the specified
6) load trainer set files (both input image and desired output text)
7) analyze input image and map all detected symbols into linear arrays
8) read desired output text from file and convert each character to a binary Unicode value to store
9) separately
10) for each character :
(a) calculate the output of the feed forward network
(b) compare with the desired output corresponding to the symbol and compute error
(c) back propagate error across each link to ad just the weights
11) move to the next character and repeat step 6 until all characters are visited
12) compute the average error of all characters
13) repeat steps 6 and 8 until the specified number of epochs
(a) Is error threshold reached? If so abort iteration.
(b) If not continue iteration

X. Experimental Results

Two groups of experiments were carried out using the set $\Omega_1 = \{0, 1, 2, \ldots, 9\}$ of handwritten numerals and the set $\Omega_2 = \{A, B, C, \ldots, Z\}$ of handwritten characters. Experiments on handwritten numerals have been performed using 18 468 patterns of the CEDAR database. Experiments on handwritten characters have been performed using 29 770 patterns of the ETL database. The patterns were thinned and normalized to 72×54 pixels [11]. Successively, the feature set $F = \{f_1, \ldots, f_9\}$ was considered for pattern description [5] as follows:

f_1 — holes (·);
f_2 — vertical-up cavities);
f_3 —vertical-down cavities (∩);
f_4 —horizontal-right cavities( );
f_5 —horizontal-left cavities ( );
f_6 —vertical-up endpoints (↑);
f_7 —vertical-down endpoints(↓);
f_8 —horizontal-right endpoints (→);
f_9 —horizontal-left endpoints (←).

Fig. 3(a) shows two handwritten patterns. Fig. 3(b) shows the geo-metrical features extracted from the thinned images.

In the latter, the structural features (that were localized at their center of gravity) are f_1 , f_2 , f_3 , f_4 , and f_5 .

Tables II and III show the experimental results [for $\varepsilon = 0.05$ of (6) and $\eta = 2$ of (16)] achieved, respectively, on handwritten numerals and characters. Each row of Tables II and III reports the performances of the optimal zonings Z_M (with M = 4, 6, 9, 16, and 25 being the number of zones of the zoning method) defined by considering dif-ferent membership functions. Accuracy estimation was performed by K -fold cross-validation technique, with K = 10 [12]. Several com-binations of working parameters have been tested, and the following

Table I: Handwritten Numerals: Performance Versus Membership Function

Zoning	Recognition Rate							
	Membership Function							
	Abstract			Ran.	Measurement			Fuzzy
	WTA [5, 7]	2-NZ [7]	3-NZ [7]	R [7]	LWM [7]	QWM [7]	EWM [7]	FMF
Z^*_4	0.81	0.78	0.62	0.68	0.51	0.58	0.82	**0.86**
Z^*_6	0.87	0.83	0.75	0.62	0.53	0.61	0.88	**0.92**
Z^*_9	0.92	0.86	0.83	0.55	0.56	0.66	0.93	**0.95**
Z^*_{16}	0.88	0.85	0.81	0.53	0.55	0.63	0.91	**0.94**
Z^*_{25}	0.87	0.84	0.79	0.52	0.53	0.62	0.90	**0.93**
Zoning	Reliability Rate							
	Membership Function							
	Abstract			Ran.	Measurement			Fuzzy
	WTA [5, 7]	2-NZ [7]	3-NZ [7]	R [7]	LWM [7]	QWM [7]	EWM [7]	FMF
Z^*_4	0.82	0.80	0.63	0.70	0.54	0.61	0.83	**0.89**
Z^*_6	0.89	0.85	0.76	0.62	0.56	0.63	0.90	**0.94**
Z^*_9	0.93	0.88	0.84	0.57	0.58	0.69	0.96	**0.97**
Z^*_{16}	0.89	0.86	0.83	0.54	0.57	0.66	0.95	**0.96**

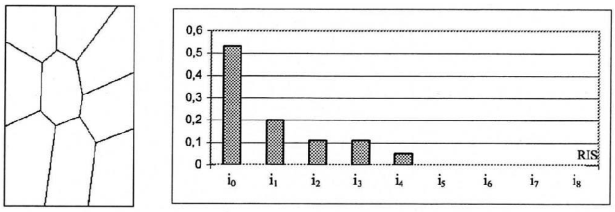
Table II shows that, whatever the number of zones considered, FMF performed better than other membership functions that are previously considered in the literature. In particular, the best result occurs for the optimal zoning Z_9 and the optimal FMF (see Fig. 4). In this case, the recognition rate and reliability rate are, respectively, 95% and 97%, which makes results better than those obtained by the technique given in [5] (i.e., 92% and 93%) and in [7] (i.e., 93% and 96%, at the best).

In addition, Table III shows that FMF outperformed abstract-level, ranked-level, and measurement-level membership functions [5], [7]

Table II: Handwritten Characters: Performance Versus Membership Function

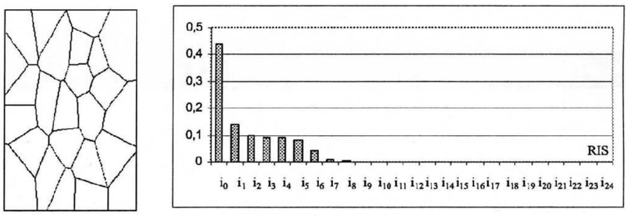
Zoning	Recognition Rate							
	Membership Function							
	Abstract			Ran.	Measurement			Fuzzy
	WTA [5, 7]	2-NZ [7]	3-NZ [7]	R [7]	LWM [7]	QWM [7]	EWM [7]	FMF
Z^*_4	0.73	0.73	0.59	0.65	0.50	0.53	0.79	**0.81**
Z^*_6	0.82	0.79	0.76	0.64	0.53	0.55	0.82	**0.84**
Z^*_9	0.84	0.82	0.83	0.62	0.53	0.57	0.84	**0.86**
Z^*_{16}	0.86	0.83	0.85	0.62	0.54	0.60	0.86	**0.89**
Z^*_{25}	0.90	0.86	0.84	0.60	0.57	0.60	0.88	**0.93**
Zoning	Reliability Rate							
	Membership Function							
	Abstract			Ran.	Measurement			Fuzzy
	WTA [5, 7]	2-NZ [7]	3-NZ [7]	R [7]	LWM [7]	QWM [7]	EWM [7]	FMF
Z^*_4	0.75	0.75	0.61	0.66	0.52	0.54	0.81	**0.84**
Z^*_6	0.83	0.80	0.77	0.65	0.55	0.57	0.83	**0.86**
Z^*_9	0.86	0.83	0.85	0.63	0.54	0.59	0.86	**0.89**
Z^*_{16}	0.88	0.85	0.86	0.64	0.55	0.62	0.89	**0.92**
Z^*_{25}	0.91	0.88	0.85	0.62	0.56	0.61	0.90	**0.95**

sets were selected for the GAs [5], [7], [10]: $N_{pop} = 2M$, with M being the number of zones, $N_{iter} = 100$, $L = 10$, p_mut_prob = 0.35, μ_mut_prob = 0.35, δ_displ = 5, b = 1, λ_displ = 0.5, and c = 3.



(a) (b)

Fig.4. Example of: (a) optimal zoning Z_9 (firstfold).



(a) (b)

Fig.5. Example of: (a) optimal zoning Z_{25} (firstfold). and (b) optimal FMF values

whatever the number of zones considered. In particular, when hand-written characters are considered, the best result occurs for the zoning Z_{25} and the fuzzy membership functions (see Fig. 5). In this case, the recognition rate and the reliability rate are 93% and 95%, respectively. This result is superior to the performance achieved by the optimal zoning-design technique, as discussed in [5], in which the recognition rate is equal to 90%, and the reliability rate is 91%, and as discussed in [7], in which the best values of recognition rate and reliability rate are equal to 88% and 90%, respectively.

XI. Conclusion

This paper has introduced a single optimization process to address the problem of zoning-based classification. Unlike previous approaches in the literature, here, the problem of optimal zoning design was con-sidered, together with the design of the optimal membership function. This paper has introduced a new class of fuzzy membership functions and presented an integrated approach that uses a real-coded genetic algorithm while exploiting the potential of Voronoi tessellation for zoning representation. This approach was employed to find, in a single optimization process, both the optimal fuzzy membership function and the optimal zoning method, for which the cost function associated with the classification problem was at its lowest.

The experimental results, which are obtained in the field of handwrit-ten digit and character recognition, confirm the soundness of consider-ing the problem of optimal zoning design, together with that of optimal membership-function design. Furthermore, they show that fuzzy mem-bership functions can provide a significant improvement in the clas-sification performance with respect to standard membership functions based on abstract-level, ranked-level, and measurement-level weight-ing models, as previously considered in the literature.

References

[1] G.Pirlo and D. Impedovo, “Fuzzy-Zoning-based Classification for Handwritten Characters”, IEEETransactions on Fuzzy Systems, Vol. 19, No.4, August 2011.

[2] O. D. Trier, A. K. Jain, and T. Taxt, “Feature extraction methods for charac-ter recognition—A survey,” *Pattern Recogn.*, vol. 29, no. 4, pp. 641–662, 1996.

[3] Y. Suen, J. Guo, and Z. C. Li, “Analysis and recognition of alphanumeric handprints by parts,” *IEEE Trans. Syst., Man, Cybern.*, vol. 24, no. 4, pp. 614–630, Apr. 1994.

[4] P. V. W. Radtke, L. S. Oliveira, R. Sabourin, and T. Wong, “Intelligent zoning design using multi-objective evolutionary algorithms,” in *Proc 7th ICDAR*, 2003, pp. 824–828.

[5] C. Gagne´ and M. Parizeau, “Genetic engineering of hierarchical fuzzy regional representations for handwritten character recognition,” *IJDAR*, vol. 8, no. 4, pp. 223–231, 2006.

[6] S. Impedovo, M. G. Lucchese, and G. Pirlo, “Optimal zoning design by genetic algorithms,” *IEEE Trans. Syst., Man, Cybern A, Syst., Humans*, vol. 36, no. 5, pp. 833–846, Sep. 2006.

[7] F. Aurenhammer, “Voronoi diagrams: A survey of a fundamental geomet-ric data structure,” *ACM Comput. Surveys*, vol. 3, no. 3, pp. 345–405, 1991. A. Ferrante, S. Impedovo, R. Modugno, and G. Pirlo, “Feature

[8] mem-bership functions in Voronoi-based zoning,” in *Emergent Perspectives in Artificial Intelligence* (LNAI, vol. 5883), R. Serra and R. Cucchiara, Eds. Berlin, Germany: Springer, 2009, pp. 202–211. ISSN: 0302-9743.

[9] T. Matsui, I. Yamashita, and T. Wakahara, “Results of second IPTP char-acter recognition competition and studies on multi-expert hand-written numeral recognition,” in *Proc. IWFHR-4*, 1994, pp. 338–346.

[10] L. Lam, Y.-S. Huang, and C.-Y. Suen, “Combination of multiple classifier decisions for optical character recognition,” in *Handbook of Character Recognition and Document Image Analysis*, H. Bunke and P. S. P. Wang, Eds. Singapore: World Scientific, 1997, pp. 79–101.

[11] T. Baeck, *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolution Programming, Genetic Algorithms*. New York: Oxford Univ. Press, 1996.

[12] G. Dimauro, S. Impedovo, and G. Pirlo, “A new thinning algorithm based on controlled deletion of edge regions,” *IJPRAI*, vol. 7, no. 5, pp. 969–986, 1993.
[13] R. Kohavi, “A study of cross-validation and bootstrap for accuracy esti-mation and model selection,” in *Proc. 14th Int. Joint Conf. Artif. Intell.*, Montreal, QC, Canada, 1995, pp. 1137–1145.

AUTHOR’S PROFILE

Dannana Kiran Babu
BCA, B.Ed, M.A., M.C.A.
Student of Kaushik College of Engineering,
Visakhapatnam, Andhra Pradesh.

M. Lakshmi Prasanna,
M.Tech. (CSE), (Ph.D)
Asst. Professor
Department of Computer Science & Engineering
Kaushik College of Engineering,
Visakhapatnam, Andhra Pradesh.
Teaching Experience: 6Years

Sri. P.Suresh Babu
B.Tech,M.E.,CSI, Associate Professor
Department of Computer Science & Engineering
Kaushik College of Engineering,
Visakhapatnam, Andhra Pradesh.
Teaching Experience: 14 Years
Industrial Experience: 4 Years

A. Bala Krishna
Student of Kaushik College of Engineering,
M.Tech. (CSE) Second Year,
Visakhapatnam, Andhra Pradesh.