Middle-East Journal of Scientific Research 20 (11): 1615-1625, 2014
ISSN 1990-9233

DOI: 10.5829/idosi.mejsr.2014.20.11.114161

# Tailored Image District Processing Precision Andwritten Appreciation

*R. Udayakumar and T. Saravanan*

School of Computing,
Bharath University, India

**Abstract:** The project proposes the data hiding technique using adaptive pixel pair matching (PPM) and LSB Replacement for secret data communication. The basic idea in LSB is the direct replacement of LSBs of pixel bits of the cover image with the secret message bits. The basic idea of PPM is to use the values of pixel pair as a reference coordinate and search a coordinate in the neighborhood set of this pixel pair according to a given message digit. The proposed method always has lower distortion for various payloads Compared with the LSB and optimal pixel adjustment process (OPAP) method. Finally project prove that the proposed method provides better performance than those of LSB and OPAP.This method is more secure and it yields the better signal to noise ratio for preserving the image qualityI.

**Key words:** This method is more secure and it yields the better · Reference coordinate and search a coordinate.

## INTRODUCTION

IMAGE ZONING is a widespread feature extraction technique for handwritten character recognition. In fact, image zoning is rightly considered effective for coping with the changeability of handwritten patterns, due to different writing styles and personal variability of the writers. By letting *B* be a pattern image, an image zoning method *ZM* can be generally considered as a partition of *B* into *M* subimages (*M* is an integer which is >1), or named zones (i.e. $ZM = \{z1, z2,\ldots, zM\}$), each one providing local information on pattern images [1, 2]. In literature, the problem of zoning design has been mainly considered as related to the design of the topology to be used, which defines the way in which a pattern image mustbesegmentein order to extract as much discriminative information as possible [3]. The approaches proposed so far for Manuscript received October 25, 2011; revised March 13, 2012; Traditional approaches involve static topologies that are designed without using *a priori* information on feature distributions in pattern classes. In this case, zoning design is performed according to experimental evidences or on the basis of intuition and experience of the designer. In general, static topologies are designed considering $u \times v$ regular grids that are superimposed on the pattern image, determining uniform partitions of the pattern image into regions of equal shape. Blumenstein *et al.* [4] use a static topology obtained by a $3 \times 2$ regular grid for handwritten character recognition.

The same topology is considered by Morita *et al.* [5], who derive contour-based features for digit recognition; by Oliveira *et al.* [6], who adopt a $3 \times 2$ grid and extract contour-based features from each zone; and by Koerich [7] and Koerich and Kalva [8], who derive directional features. Suen *et al.* [9], [10] also use a $3\times2$ regular grid to define a model to evaluate the distinctive parts of handwritten characters and to compare human and machine capabilities in character recognition by parts. A $3 \times 3$ regular grid for zoning design is used by Baptista and Kulkarni [11] who extract geometrical feature distribution from each zone and by Singh and Hewitt [12] use a modified Hough transform method to extract features for handwritten digit and character recognition. Phokharatkul *et al.* [13] present a system for handwritten character recognition based on antminer algorithm. They use a $4 \times 3$ regular grid for zoning design in order to extract closed-loop and endpoint features from the pattern image. A $4 \times 4$ regular grid is used by Cha *et al.* [14] to extract gradient, structural and concavity information from the pattern image and by Negi *et al.* [15] to derive the density of pixels in the different zones. Kimura and Shridhar [15] use a zoning topology based on a $4 \times 4$ regular grid to detect information from the contour profiles

**Corresponding Author:** R. Udayakumar, School of Computing, Bharath University, India.

of the patterns. In each zone, the numberof segments on the contour of the pattern with the same orientation is counted. Four basic orientations are considered: 0°, 90°, +45° and -45°. The same grid is used by Liu *et al* to recognize Chinese characters by a directional decomposition approach.

Camastra and Vinciarelli use a 4 × 4 regular grid for recognizing isolated cursive characters extracted from word images. In this case, two sets of operators are applied to each zone. The operators of the first set measure the percentage of foreground pixels in the zone with respect to the total number of foreground pixels in the character image. The operators of the second set estimate to what extent the black pixels in the cell are aligned along some directions. Xiang *et al.* apply zoning to the recognition of car plates. They extract pixel density features dividing the character input image from car plates using a 4 × 4 regular grid. Impedovo *et al.* consider 3 × 2, 3 × 3 and 4 × 4 regular grids and use a genetic algorithm to determine the optimal weight vector to balance local decisions by using *M* zones. Sharma and Gupta use 4 × 4, 6 × 6 and 8 × 8 regular grids to extract pixel density from the pattern image. Rajashekararadhya and Ranjan use a 5 × 5 regular grid for zoning design. For each zone, the average distances from the character centroid to the pixels in each row/column are considered as features. A 5 × 5 regular grid is also used by Vamvakas *et al.* to compute local density in the character image. Kato and Suzuki use a 7 × 7 regular grid for Chinese and Japanese handwritten characters. A similar approach, which uses overlapped zones to reduce border effects, is also proposed by Kimura *et al*. Dynamic topologies are designed according to the result of optimization procedures. Aires *et al.* and Freitas *et al.* present a perception-oriented approach that uses nonregular grids for zoning design, resulting in a nonuniform splitting of the pattern image. They manually define the zoning grid by using the confusion matrices looking for the relation between the zones, in order to make the zoning design process less empirical. Other approaches, based on automatic optimization schemes, generally concern constrained zoning methods based on predetermined templates. Valveny and Lopez use a zoning method for digit recognition located on surgical sachets which pass through a computer vision system performing quality control. In this case, the authors divide the pattern image into five rows and three columns. The size of each row and column is determined in such a way to maximize the discriminating capabilities of the diverse zones of the pattern image. Dimauro *et al.* performed zoning design

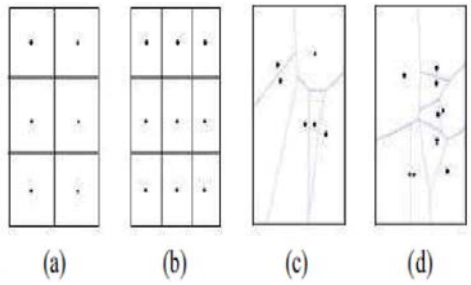



Fig. 1: Example of zoning method: static versus dynamic (Voronoi-based). (a) Static: $Z_{3\times2}$ (b Static: $Z_{3\times3}$ (c)Dynamic: $Z_6$* (d) Dynamic: Z*

according to the analysis of the discriminating capability of each zone, estimated by means of the statistical variance of feature distributions. Di Lecce *et al.* designed the zoning problem as an optimization problem in which the discrimination capability of each zone is stimated by the Shannon entropy. Lazzerini and Marcelloni applied a method to handwritten characters for the fuzzy classification and recognition of 2-D shapes. The character image is partitioned horizontally and vertically into stripes. For each dimension, a set of weights is determined that define the importance of each stripe in the classification process and a genetic algorithm is used to optimize stripe dimension with respect to the recognition rate (REC). Radtke *et al.* presented an automatic approach to define zoning based on fixed-position divisions of pattern images. Gagné and Parizeau used a tree-based hierarchical zoning or handwritten character classification and presented a genetic programming approach for optimizing the feature extraction step of a handwritten character recognizer. Converse to previous approaches, in which dynamic zoning methods were designed according to constrained topologies based on predetermined templates, Impedovo *et al.* proposed Voronoi tessellation for zoning description, since (a) (b) (c) (d).

Voronoi tessellation allows the design of dynamic zoning methods based on unconstrained topologies. In fact, given a set of points (named Voronoi points) in continuous space, Voronoi tessellation is a simple means of naturally partitioning the space into zones, according to proximity relationships among the set of points. More precisely, let *B* be a pattern image and $P = \{p1, p2,\ldots, pM\}$ a set of M distinct points in B. The Voronoi tessellation determined by *P* is the partition of *B* into *M* zones $\{z1, z2,\ldots, zM\}$, with the property that each region *zi* defined by *pi*, contains all the points *p* for which it results that

distance ($p$, $pi$ ) < distance ($p$, $p\,j$ ), for any $p\,j \_= pi$. In addition, concerning the boundaries of the zones, it is also assumed that the points of $B$ that are equidistant from two (or more) points of $P$ belong to the zone with minimum index. For instance, let $p$ be a point for which distance ($p$,$pi$ ) = distance ($p$, $p\,j$ ), for $i \_= j$. In this paper, it is assumed that $p \square zi$, if $i < j\ p \square z\,j$, if $i > j$. Of course, changing the position of the Voronoi points corresponds to the modification of the zoning method. Therefore, zoning description with Voronoi tessellation offers the possibility to easily adapt the zoning to the specific characteristics of the classification problem. In fact, a genetic algorithm for zoning design has also been proposed, in which each individual of the genetic population is a set of Voronoi points (corresponding to a zoning method) and the cost function (CF) associated to the classification is considered as a fitness function. Fig. 1 shows some examples of static and dynamic zoning topologies. Fig. 1(a) and (b) shows two uniform topologies obtained by 3 × 2 and 3 × 3 regular grids, respectively. Fig. 1(c) and (d) shows two examples (with six and nine zones, respectively) of optimized nonuniform topologies. In all cases, the Voronoi points of the zones are reported. As Fig. 1 demonstrates, Voronoi tessellations can be used for describing both static and dynamic zoning topologies. Of course, when uniform topologies are considered, as in Fig. 1(a) and (b), the Voronoi point of each zone corresponds to the center of that zone. Unfortunately, although zoning methods are largely adopted and advanced techniques for optimal topology design ave been proposed, aspects related to the choice of feature-zone membership functions have not yet been sufficiently addressed. Notwithstanding, membership function plays a crucial role in exploiting the potential of a zoning method since it should be able to model spatial distributions of features in the different zones. Thus, when zoning is used, the choice of a membership function needs specific attention. In literature, the membership values are assigned on the basis of the values of specific proximity-based functions. According to the type of values used to define membership weights,three classes of order-based membership functions can be defined [37]: abstract level, ranked level and measurement level. When abstract-level membership functions are considered, the membership values are given in the form of Boolean values. When ranked-level membership functions are used, the membership values are integers. When measurement-level membership functions are used, the membership values are real numbers. Fuzzy membership functions have also been proposed, in which the membership weights are assumed to be fuzzy values derived by an optimization procedure. It is worth noting that order-based functions are not able to cope with the specific characteristics of pattern distributions. Furthermore, both order-based membership functions and fuzzy membership functions follow a global strategy, that is, the membership functions are the same for all zones of a zoning method. The result being that they are unable to exploit the evidence that feature distributions in diverse zones of pattern images can be very different. Starting from this consideration, this paper introduces a new class of zone-based membership functions with adaptive capabilities and presents a real-coded genetic algorithm for determining-in a single process-both the optimal zoning method, based on Voronoi tessellation of the pattern image and the adaptive membership function most profitable for a given classification problem. Contrary to other approaches proposed in the literature so far, the new class of membership functions allows the membership function to adapt to the specific feature distribution of each zone of the zoning method. The experimental tests were carried out in the field of handwritten digit and character recognition using datasets from coupling, energetics and dynamics of atmospheric regions (CEDAR) and extract, transform, load (ETL) databases, respectively. As expected, the results show that the effectiveness of a zoning method strongly depends on the membership function considered. In addition, they demonstrate that adaptive membership functions are superior to traditional functions, whatever zoning topology is used. Of course, when adaptive zone-based membership functions were selected together with the optimal Voronoi-based zoning-according to the approach proposed in this paper-the recognition and reliability rates (RELs) achieved the best results for both the numeral and character recognition. This paper is organized as follows. The role of membership functions for feature extraction by zoning methods is focused on in Section II, which also illustrates the new class of adaptive membership functions proposed in this paper.

**Feature Extraction by Zoning Methods:** Let $ZM = \{z1, z2, \ldots, zM\}$ be a zoning method. A crucial aspect for zoning-based classification concerns the way in which each feature detected in a pattern $x$ has influence on each zone of $ZM$. In fact, let us consider the classification of a pattern $x$ into one class of $\_ = \{C1, \ldots, CK\}$ using the

feature set $F = \{ f1, \ldots, fT \}$. In this case, $x$ can be described by the feature matrix $Ax$ of $T$ rows (features) and $M$ columns (zones).

$Ax = Ax(1, 1)\ Ax(1, 2).\ Ax(1, j).\ Ax(1, M)$
$Ax(2, 1)\ Ax(2, 2)...\ Ax(2, j)...\ Ax(2, M)$
$Ax(i, 1)\ Ax(i, 2)...\ Ax(i, j)...\ Ax(i, M)$
$Ax(T, 1)\ Ax(T, 2)..\ Ax(T, j)...\ Ax(T, M)$

With $Ax(i, j) = \_wij$ instances of $fi$ in $x$ (2) where $wij$ is the weight that defines the degree of influence of an instance of feature $fi$ (detected in $x$) on zone $zj$. Now, the influence weight $wij$ of an instance of $fi$ on zone $zj$ is determined on the basis of the proximity condition between the position of the instance of $fi$ and $zj$ (it is worth noting that the position of a feature $fi$ is assumed to be located at the center of gravity of $fi$ when structural features are considered, such as lines, loops, cavities, arcs, etc. More precisely, let $ZM = \{z1, z2, \ldots, zM\}$ be a zoning method corresponding to the Voronoi points $P = \{p1\ p2, \ldots, pM\}$, where $zj$ is the Voronoi region corresponding to the Voronoi point $pj, j = 1, 2, \ldots, M$; let $qi$ be the point in which the instance of feature $fi$ is found; let $dij = \text{dist}(qi, pj)$ be the Euclidean distance between $qi$ and $pj$ ; the ranked index sequence ($RISi$) associated with the instance of feature $fi$, which denotes the sequence of the zones ranked according to their proximity to $qi$, is defined as

$RISi = < i1, i2,... ik, ik+1,... iK >$ (3) with:1) $ik \in \{1, 2, \ldots, M\}\ \forall k = 1, 2, \ldots, K$;
2) $ik1 \neq ik2$
$\forall k1, k2 = 1, 2,... K, k1 \neq k2$
and for which it results
$dik1 < dik2,\ k1 < k2\ \forall k1, k2 = 1, 2,, K$ (4)
we also assume that in the case $dik1 = dik2$, $ik1$ precedes $ik2$,
if $k1 < k2$.

Furthermore, let $counti(j)$ be the function providing the position of the index $j$ (i.e. concerning zone $zj$) in the sequence $RISi$ [i.e. $counti(j) = k$ for $j = ik$, according to (3)], the following feature-zone membership functions can be considered. 1) *Abstract-Level Membership Functions:* Membership functions at abstract level assign Boolean influence weights on the basis of the first $k$ zones in $RISi$. a) *The winner-takes-all (WTA) membership function:* This is the standard membership function used in traditional zoning-based classification. In this case, the results are $wij = 1$, if $counti(j) = 1$ (5a) $wij = 0$, otherwise. (5b) b) *The k-nearest zone (k-NZ) membership function:*

This is a generalization of the WTA function. In this case, the results are $wij = 1$, if $counti(j) \in \{1, 2,... k\}$ (6a) $wij = 0$, otherwise. (6b) c) *Ranked-level membership functions:* Membership functions at ranked level assign integer influence weights on all zones, the basis of their position in the $RISi$. d) *The ranked-based (R) membership function.* In this case, the result is $wij = M - k$, if $counti(j) = k$. (7) 2) *Measurement-Level Membership Functions:* Membership functions at measurement level assign real influence weights according to the distance between the zones and the instance of the feature $fi$. In this paper, three measurement-level membership functions are considered.

- *Linear weighting model (LWM)* $wij = 1\ dij$. (8a)
- *Quadratic weighting model (QWM)* $wij = 1\ d2ij$ (8b)
- *Exponential weighting model (EWM)* $wij = 1\ edi$. For example, for the instance of $fi$ (an endpoint) shown in Fig. 2, the numerical values of the membership functions are reported in Table I (for the sake of clarity, the values of the measurement-based membership functions have been normalized). In this case, we consider the zoning method $Z9 = \{z1, z2, \ldots, z9\}$ ($M = 9$) corresponding to the set of Voronoi points $P = \{p1, p2, \ldots, pM\} = \{(9, 60), (27, 60), (45, 60), (9, 36), (27, 36), (45, 36), (9, 12), (27, 12), (45, 12)\}$, whereas the position of $fi$ is $qi = (qxi, qyi) = (44, 57.8)$. Starting from the set of Euclidean distances $dij = \text{dist}(qi, pj) = [(pxi - qxi)2 + (pyi - qyi)2]1/2$, for $j = 1, 2, \ldots, 9$, the values of the membership functions are computed. It is worth noting that here $RISi = <3, 2, 6, 5, 1, 4, 9, 8, 7>$. In this paper, starting from the basic idea that pattern features are spatially distributed according to local characteristics, a new adaptive technique to membership function design is considered. In fact, there are regions of the pattern image in which features are arranged into a small area (stable regions), whereas there are regions in which features are spread over a very large area (variable regions).

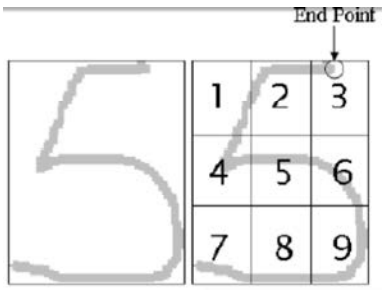



Fig. 2: Zoning methods: $RIS_i = <3, 2, 6, 5, 1, 4, 9, 8, 7>$.

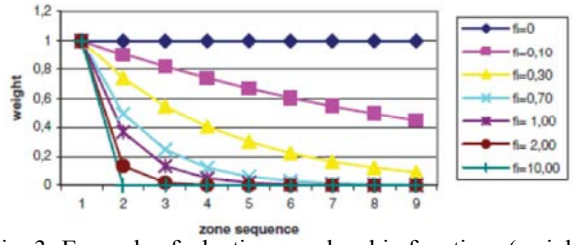



Fig. 3: Example of adaptive membership functions (weight versus zone).

Therefore, a membership function could be able to adapt itself to the local distributions of patterns. In addition, this paper takes advantage from the evidence that the membership function based on the exponential weighted model generally leads to superior performance than other membership functions, as already discussed in the literature. For this purpose, the new zone-based membership functions, for each zone *z j*, are here defined according to an adaptive weighted model (AWM) *wi j* = *e-λj di j* (9) where *λj* is a positive parameter named falling rate undergoing exponential decay of the adaptive membership function. Larger falling rates make the value of the membership function vanish much more rapidly, as the distance between the position of the feature and the zone increases. The membership functions for different values of *λj* are shown in Fig. 3. It should be noted that these work as a traditional WTA strategy, for *λj* = 10. In fact, in this case, feature *fi* has influence (with *wi j* = 1) only on the zone *z j* in which *fi* is positioned. Conversely, when *λj* = 0, *fi* has an equal influence (with *wi j* = 1) on all zones, no matter where *fi* is positioned.

**Classification in Voronoi Tessellation by Adaptive Membership Functions:** As discussed before, the zoning design process concernsboth the definition of the optimal topology along with the definition of the optimal membership functions. According to previous studies in the literature, the CF of a zoning-based classifier, which depends on both zoning method (*ZM)* and membership function (*FM)*, is here defined as follows: CF *(ZM,FM)* = *ç* • Err*(ZM,FM)* + Rej *(ZM,FM)* (10)

where:

- Err(Z*M,*F*M)* and Rej(Z*M,*F*M)* are the misrecognition rate and the rejection rate of the zoning-based classifierrespectively;2) coefficient *ç* is the cost value associated with the treatment of an error with respect to a rejection. More precisely, in the proposed approach, Err(Z*M,* F*M)* andRej(Z*M,* F*M)* are estimated using the patterns of the learning set *XL* as follows: 1) Err(*ZM*, *FM)* = card{*xr* □ *XL* |*xr* is misclassified by the zoning-based classifier}/card(*XL)*;
- Rej(Z*M*, *FM)* = card {*xr* □ *XL* |*xr* is rejected by the zoning-based classifier}/card(*XL)*. Moreover, since Voronoi tessellation is used for zoning description and the adaptive membership functions are considered, the following formulation of the problem of optimal zoning design is given. Find the sets {*p**1, *p**2,…, *p***M*} (Voronoi points) and {*λ**1, *λ**2,…, *λ***M*} (falling values) so that CF*(Z*M, F*M)* = min {*ZM,FM*} CF*(ZM, FM)* (11) with:

TABLE I

MEMBERSHIP FUNCTIONS: A NUMERICAL EXAMPLE

| Zone | $q_i = (qx_i, qy_i)$ | $p_i = (px_j, py_j)$ | $d_{ij} = \text{dist}(q_i, p_j)$ | count$_j$(*j*) | Abstract | | | Ranked | Measurement | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | WTA | 2-NZ | 3-NZ | R | LWM | QWM | EWM |
| | | | | | (5) | (6-*k* = 2) | (6-*k* = 3) | (7) | (8a) | (8b) | (8c) |
| $z_1$ | (44, 57.8) | (9, 60) | 35.06 | 5 | 0 | 0 | 0 | 4 | 0.046 | 0.006 | 0 |
| $z_2$ | (44,57.8) | (27,60) | 17.14 | 2 | 0 | 1 | 1 | 7 | 0.096 | 0.025 | 0.015 |
| $z_3$ | (44,57.8) | (45,60) | 2.41 | 1 | 1 | 1 | 1 | 8 | 0.586 | 0.932 | 0.981 |
| $z_4$ | (44, 57.8) | (9,36) | 41.23 | 6 | 0 | 0 | 0 | 3 | 0.040 | 0.004 | 0 |
| $z_5$ | (44, 57.8) | (27, 36) | 27.64 | 4 | 0 | 0 | 0 | 5 | 0.060 | 0.010 | 0.001 |
| $z_6$ | (44, 57.8) | (45, 36) | 21.82 | 3 | 0 | 0 | 1 | 6 | 0.075 | 0.015 | 0.004 |
| $z_7$ | (44, 57.8) | (9, 12) | 57.64 | 9 | 0 | 0 | 0 | 0 | 0.029 | 0.002 | 0 |
| $z_8$ | (44, 57.8) | (27, 12) | 48.85 | 8 | 0 | 0 | 0 | 1 | 0.033 | 0.003 | 0 |
| $z_9$ | (44, 57.8) | (45, 12) | 45.81 | 7 | 0 | 0 | 0 | 2 | 0.035 | 0.003 | 0 |

- $Z^*M$ = {$z^*1$, $z^*2$,…, $z^*M$}, $z^*$ $j$ being the Voronoi region corresponding to $p^*$ $j$, □$j$ = 1, 2,…, M;
- $ZM$ = {$z1$, $z2$,…, $zM$}, $z$ $j$ being the Voronoi region corresponding to $p$ $j$, □$j$ = 1, 2,…, $M$;
- $F^*M$ = {$\lambda^*1$, $\lambda^*2$,…, $\lambda^*M$}, $\lambda^*$ $j$ being the falling value of the adaptive membership functions associated to the zone $z^*$ $j$, □$j$ = 1, 2,…, $M$;
- $FM$ = {$\lambda1$, $\lambda2$,…, $\lambda M$}, $\lambda j$ being the falling value of the adaptive membership functions associated to the zone $z$ $j$, □ $j$ = 1, 2,…, $M$. In order to solve the optimization problem (11), a real-coded genetic approach is used since it has potential for solving nonlinear optimization problems in which the analytical expression of the object function is not known. In the following, the genetic algorithm is described for the design of the adaptive membership functions together with the optimal zoning. The initial population Pop = _i,_2,_i..., _l,...,_Nð0ð for the genetic algorithm is created by generating $N$pop random individuals ($N$pop even). Each individual is a vector $i$ =$p\lambda1$, $p2\lambda2$,$p3$ $\lambda j$,$Pm$,$\lambda M$--- (12) where each element $p$ $j$ $\lambda j$ consists of: 1) $p$ $j$ : a point defined as $p$ $j$ = ($x$ $j$, $y$ $j$ ), that corresponds to the Voronoi point of the zone $z$ $j$ of $ZM$ = {$z1$, $z2$, …, $zM$}; 2) $\lambda j$ : a falling value that defines the adaptive model for the membership function of the zone $z$ $j$. Consequently, the fitness value of the individual is taken as the classification cost $CF$($ZM$, $FM$), obtained by (10), where: 1) $ZM$ = {$z1$, $z2$,…, $zM$} is the Voronoi tessellation and $z$ $j$ is the Voronoi region corresponding to $p$ $j$, □$j$ = 1, 2,…, $M$;
- $FM$ = {$\lambda1$, $\lambda2$,…, $\lambda M$} is the set of adaptive membership functions and $\lambda j$ is the falling value of the adaptive weighing model associated to the zone $z$ $j$,□$j$ =1, 2,... $M$.

From the initial population, the following genetic operatorsare used to generate new populations of individuals: individual selection, crossover, mutation and elitism. In the following, a brief explanation of the adopted operators is reported (the complete description of these genetic operators can be found in). where $i$ is a random value generated in the range, according to a uniform distribution; $\delta$_displ the maximum displacement allowed; $b$ a parameter determining the degree

**Individual Selection:** Npop/2 random pairs of individuals are selected for crossover, according to a roulette wheel strategy. of nonuniformity; iter the count of the

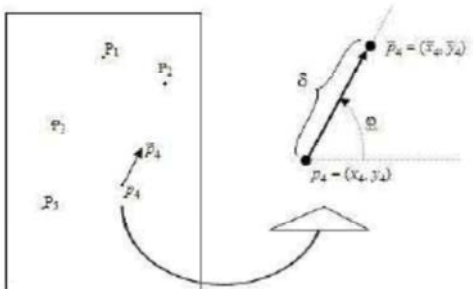


Fig. 4 : Mutation operator

generations performed; $N$iter the maximum number of generations. It is worth noting that (16) causes the operator to search the space almost uniformly initially when iter is small and locally in later stages. b) Similarly, concerning ˜$\lambda$ $j$, we have ˜$\lambda j$ = $\lambda j$+(-1)$s$ •$\lambda$_displ•1 - ç _ 1-iter $N$iter _$c$_ (17)

where $s$ is a random Boolean value generated according to a equally distributed probability function; ç is a random value generated in the range, according to a uniform distribution; $\lambda$_displ is the maximum displacement allowed; $c$ is a parameter determining the degree of nonuniformity; and iter denotes the count of the generations performed while Niter denotes the maximum number of generations.

**Elitist Strategy:** from the N$pop$ individuals generated by the above operations, one individual is randomly removed and the individual with the minimum cost in the previous population is added to the current population. Steps from 1) to 4) are repeated until $N$iter successive populations of individuals are generated. When the process stops, the optimal zoning is obtained by the best individual of the last-generated population. horizontal-left endpoints. Please note that the description of this feature set is behind the aims of this paper and the interested reader can find a detailed description in the literature. For pattern classification, a $k$-NN classifier ($k$ = 1) was considered. Pattern rejection occurred when the two training vectors closest to the test vector were related to two diverse classes and the difference of the distances between each one of the two training vectors and the test vector was smaller than a suitable threshold $\hat{\imath}$ ($\hat{\imath}$ = 0.7 in the tests). In addition, the following parameter values for the genetic algorithm were selected by $k$-fold cross validation ($k$ = 10) on the training sets: $N$Pop = 10; $N$iter = 300; Mut_prob = 0.35; $\delta$_displ = 5; $b$ = 1.0; $\lambda$_displ = 0.5 and $c$ = 3.0.

TABLE II(a)
PERFORMANCE ANALYSIS ON $\Omega_1$: REC

| M | Zoning | Performance | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Abstract | | | Ranked | Measurement | | | Adaptive |
| | | WTA | 2-NZ | 3-NZ | R | LWM | QWM | EWM | AWM |
| 6 | $Z_{3\times2}$ | 0.74 | 0.70 | 0.62 | 0.50 | 0.44 | 0.50 | 0.73 | 0.81 |
| | $Z_6^*$ | 0.87 | 0.84 | 0.75 | 0.62 | 0.52 | 0.57 | 0.87 | 0.89 |
| 9 | $Z_{3\times3}$ | 0.80 | 0.73 | 0.69 | 0.46 | 0.47 | 0.51 | 0.81 | 0.88 |
| | $Z_9^*$ | 0.89 | 0.84 | 0.79 | 0.53 | 0.52 | 0.58 | 0.93 | 0.97 |
| 16 | $Z_{4\times4}$ | 0.82 | 0.79 | 0.77 | 0.44 | 0.49 | 0.47 | 0.80 | 0.85 |
| | $Z_{16}^*$ | 0.88 | 0.85 | 0.84 | 0.54 | 0.54 | 0.63 | 0.91 | 0.93 |
| 25 | $Z_{5x5}$ | 0.81 | 0.79 | 0.76 | 0.47 | 0.46 | 0.49 | 0.78 | 0.86 |
| | $Z_{25}^*$ | 0.89 | 0.87 | 0.85 | 0.55 | 0.57 | 0.65 | 0.91 | 0.93 |
| Average | | 0.83 | 0.80 | 0.75 | 0.51 | 0.50 | 0.55 | 0.84 | 0.89 |

TABLE II(b)
PERFORMANCE ANALYSIS ON $\Omega_1$: REL

| M | Zoning | Performance | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Abstract | | | Ranked | Measurement | | | Adaptive |
| | | WTA | 2-NZ | 3-NZ | R | LWM | QWM | EWM | AWM |
| 6 | $Z_{3\times2}$ | 0.77 | 0.73 | 0.65 | 0.54 | 0.51 | 0.61 | 0.76 | 0.82 |
| | $Z_6^*$ | 0.91 | 0.88 | 0.79 | 0.73 | 0.61 | 0.64 | 0.91 | 0.93 |
| 9 | $Z_{3\times3}$ | 0.86 | 0.77 | 0.73 | 0.51 | 0.54 | 0.63 | 0.84 | 0.91 |
| | $Z_9^*$ | 0.92 | 0.89 | 0.83 | 0.63 | 0.58 | 0.66 | 0.95 | 0.99 |
| 16 | $Z_{4\times4}$ | 0.86 | 0.84 | 0.81 | 0.49 | 0.54 | 0.53 | 0.83 | 0.88 |
| | $Z_{16}^*$ | 0.91 | 0.88 | 0.87 | 0.62 | 0.61 | 0.69 | 0.94 | 0.95 |
| 25 | $Z_{5\times5}$ | 0.85 | 0.85 | 0.79 | 0.55 | 0.53 | 0.55 | 0.81 | 0.90 |
| | $Z_{25}^*$ | 0.93 | 0.92 | 0.89 | 0.61 | 0.63 | 0.70 | 0.93 | 0.97 |
| Average | | 0.87 | 0.84 | 0.79 | 0.58 | 0.56 | 0.62 | 0.87 | 0.91 |

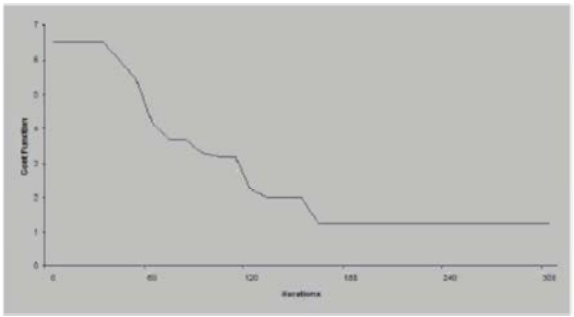
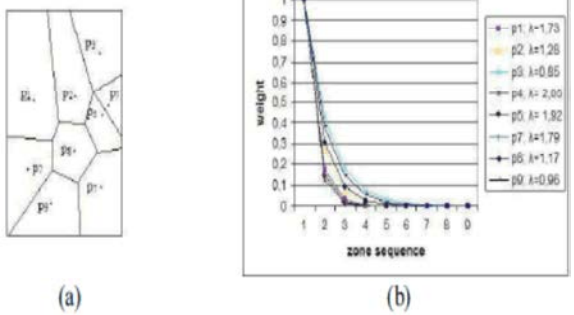


Fig. 5: Genetic algorithm: CF versus iteration (M=9)




(a) (b)

Fig. 6: Optimal $Z_9$* zoning method : an example. (a) Optimal topolgy, (b) Optimal adapive membership function.

An example of the convergence of the genetic algorithm is shown in Fig. 5, for $M$ = 9. Fig. 6 shows the results obtained for the specific case of $M$ = 9, when handwritten digits are considered. Fig. 6(a) shows the optimal zoning $Z$* 9 and Fig. 6(b) reports the set of optimal adaptive membership functions $F$* 9. Concerning the time complexity of the new technique, it can be measured by the number of fitness function evaluations done during the course of a run. Hence, for fixed population sizes, the number of fitness function evaluations is given (in the worst case) by the product of population size ($N$pop*)* per number of generations ($N$iter*)*. Of course, faster convergence of the genetic algorithm can be expected if the initial population is not random, but it is defined starting from the analysis of local properties of statistical feature distributions. Anyway, it is worth noting that time complexity is not a limitation of the new technique since it concerns the optimization of the zoning topology and membership functions which occurs during the tuning

TABLE III(a)
PERFORMANCE ANALYSIS ON $\Omega_2$: REC

| M | Zoning | Performance | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Abstract | | | Ranked | Measurement | | | Adaptive |
| | | WTA | 2-NZ | 3-NZ | R | LWM | QWM | EWM | AWM |
| 6 | $Z_{3\times2}$ | 0.71 | 0.63 | 0.6 | 0.51 | 0.50 | 0.58 | 0.72 | 0.74 |
| | $Z_6^*$ | 0.77 | 0.71 | 0.68 | 0.54 | 0.52 | 0.63 | 0.78 | 0.80 |
| 9 | $Z_{3\times3}$ | 0.80 | 0.64 | 0.62 | 0.55 | 0.57 | 0.62 | 0.82 | 0.88 |
| | $Z_9^*$ | 0.85 | 0.74 | 0.69 | 0.60 | 0.62 | 0.71 | 0.87 | 0.92 |
| 16 | $Z_{4\times4}$ | 0.84 | 0.7 | 0.67 | 0.54 | 0.55 | 0.64 | 0.85 | 0.89 |
| | $Z_{16}^*$ | 0.87 | 0.79 | 0.72 | 0.59 | 0.60 | 0.73 | 0.88 | 0.93 |
| 25 | $Z_{5\times5}$ | 0.87 | 0.71 | 0.67 | 0.54 | 0.56 | 0.62 | 0.87 | 0.91 |
| | $Z_{25}^*$ | 0.91 | 0.8 | 0.75 | 0.60 | 0.63 | 0.70 | 0.92 | 0.95 |
| Average | | 0.82 | 0.71 | 0.67 | 0.55 | 0.56 | 0.65 | 0.83 | 0.87 |

TABLE III(b)
PERFORMANCE ANALYSIS ON $\Omega_2$: REL

| M | Zoning | Performance | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Abstract | | | Ranked | Measurement | | | Adaptive |
| | | WTA | 2-NZ | 3-NZ | R | LWM | QWM | EWM | AWM |
| 6 | $Z_{3\times2}$ | 0.73 | 0.66 | 0.68 | 0.57 | 0.55 | 0.62 | 0.78 | 0.82 |
| | $Z_6^*$ | 0.80 | 0.76 | 0.75 | 0.61 | 0.60 | 0.69 | 0.86 | 0.85 |
| 9 | $Z_{3\times3}$ | 0.81 | 0.68 | 0.65 | 0.60 | 0.62 | 0.67 | 0.87 | 0.91 |
| | $Z_9^*$ | 0.86 | 0.79 | 0.73 | 0.68 | 0.67 | 0.76 | 0.93 | 0.94 |
| 16 | $Z_{4\times4}$ | 0.85 | 0.75 | 0.71 | 0.62 | 0.59 | 0.67 | 0.88 | 0.92 |
| | $Z_{16}^*$ | 0.89 | 0.84 | 0.77 | 0.67 | 0.66 | 0.73 | 0.91 | 0.95 |
| 25 | $Z_{5\times5}$ | 0.88 | 0.76 | 0.73 | 0.71 | 0.66 | 0.70 | 0.90 | 0.93 |
| | $Z_{25}^*$ | 0.93 | 0.85 | 0.84 | 0.77 | 0.74 | 0.78 | 0.94 | 0.97 |
| Average | | 0.84 | 0.76 | 0.73 | 0.65 | 0.63 | 0.70 | 0.88 | 0.91 |

phase of the system, before the running phase. Therefore, the time complexity of the technique has no effect on the classification speed of the zoning-based classifier, which remains the same as that of traditional zoning methods. The main results of the experimental tests are summarized in Tables II and III, which report the performance obtained on _1 and _2, respectively. The effectiveness of different membership functions is compared on both static zonings and dynamic, Voronoi-based zoning methods. The performance is reported in terms of REC and REL, for $ç$ = 10 [see (10)]. The result shows that dynamic zoning methods based on optimal Voronoi tessellation always outperform static methods, whatever number of zones ($M$) and membership function ($F$) are considered. In particular, when handwritten numerals are considered, Table II(a) shows that the improvement in REC ranges from 7% when $M$ = 16 and $F$ = WTA [REC($Z$* 16, WTA) = 0.88 versus REC($Z$4×4, WTA) = 0.82] up to 34% when $M$ = 16 and $F$ = QWM (REC($Z$* 16, QWM) = 0.63 versus REC($Z$4×4, QWM) = 0.47, whereas improvement in REL [Table II(b)] ranges from 4% when $M$ = 16 and 2-NZ (REL($Z$* 16, 2-NZ) = 0.88 versus REL($Z$4×4, 2-$NZ$) = 0.84 up to 35% when $M$ = 6 and $F$ = $R$ (REL(Z6, $R$) = 0.73 versus REL ($Z$3×2, $R$) = 0.54. Conversely, when handwritten characters are used, improvement in recognition rate [Table III(a)] ranges from 3% when $M$ = 16 and $F$ = WTA (REC($Z$* 16, WTA) = 0.87 versus REC(Z4×4, WTA) = 0.84 up to 14% when $M$ = 9 and $F$ = QWM (REC($Z$* 9, QWM) = 0.71 versus REC($Z$3×3, QWM) = 0.62, whereas improvement in REL [Table III(b)] ranges from 3% when $M$ = 6 and $F$ = AWM (REL($Z$* 16, AWM) = 0.95 versus REL($Z$4×4, AWM) = 0.92 up to 16% when $M$ = 9 and $F$ = 2-NZ [REL($Z$* 9, 2-NZ) = 0.79 versus REL($Z$3×3, 2- NZ) = 0.68]. Furthermore, Tables II and III show that AWM is superior to other membership functions, whatever zoning methods are used. In particular, concerning handwritten digit recognition, Table II (a) shows that, when abstract membership functions are used, the average REC is 0.83, 0.80 and 0.75 for WTA, 2-NZ and 3-NZ, respectively.When ranked membership func tion is considered (R), the REC is 0.51 on average. When measurement membership functions are used, the average REC is 0.50, 0.55 and 0.84 for LWM, QWM and EWM, respectively. The average recognition rate is 0.89 when AWM is used. Moreover, Table II(a) shows that the best result occurs for $Z$* 9. In this case, REC($Z$* 9, AWM) =

0.97 results and the improvement is equal to 8% with respect to WTA, 15% with respect to 2-NZ, 25% with respect to 3-NZ, 83% with respect to R, 98% with respect to LWM, 67% with respect to QWM and 4% with respect to EWM. Concerning the REL, Table II(b) shows that AWM always outperforms other membership functions. In particular, when abstract membership functions are used, the average REL is 0.87, 0.84 and 0.79 for WTA, 2-NZ and 3-NZ, respectively. When ranked membership function is considered (R), the REL is 0.58 on average. When measurement membership functions are used, the average REL is 0.56, 0.62 and 0.87 for LWM, QWM and EWM, respectively. Finally, when the new AWM is used, the average REL is 0.91. Furthermore, Table II(b) shows that the best reliability occurs for $Z^*$ 9. In this case, REL($Z^*$ 9, AWM) = 0.99 results and the improvement is equal to 7% with respect to WTA, 8% with respect to 2-NZ, 19% with respect to 3-NZ, 57% with respect to R, 70% with respect to LWM, 50.0% with respect to QWM and 4% with respect to EWM. In order to evaluate the statistical significance of the classification results on handwritten digits, the two way analysis of variance (ANOVA) has been performed on the data of Table II(a) and (b). The ANOVA test (with the significance level $\alpha$ 0.05) demonstrated that differences between traditional zoning methods and optimized methods based on Voronoi tessellation and adaptive membership functions are significant in terms of both RECs and REL. Concerning handwritten character recognition, Table III(a) shows that AWM provides the best results. Precisely, when abstract membership functions are used, the average RECs is 0.82, 0.71 and 0.67 for WTA, 2-NZ and 3-NZ, respectively. When ranked membership function is considered (R), the REC is 0.55 on average. When measurement membership functions are used, the average REC is 0.56, 0.65 and 0.83 for LWM, QWM and EWM, respectively. Finally, when the AWM is used, the average REC is 0.87. More precisely, Table III(a) shows that the best result occurs for $Z^*$ 25. In this case, REC($Z^*$ 25, AWM) = 0.95 results and the improvement is equal to 4% with respect to WTA, 18% with respect to 2-NZ, 26% with respect to 3-NZ, 58% with respect to R, 50% with respect to LWM, 35% with respect to QWM and 3% with respect to EWM. Table III(b) shows that the AWM is always superior to other membership functions, when the REL is considered. More precisely, when abstract membership-functions are used, the average REL is 0.84, 0.76 and 0.73 for WTA, 2-NZ and 3-NZ, respectively. When ranked membership function is considered (R), the REL is 0.65 on average. When measurement membership functions are used, the average REL is0.63, 0.70 and 0.88 for LWM, QWM and EWM, respectively. Finally, when the AWM is used, the average REL is 0.91. Furthermore, Table III(b) shows that the best reliability occurs for $Z^*$ 25. In this case, REL($Z^*$ 25, AWM) = 0.97 results and the improvement is equal to 4% with respect to WTA, 14% with respect to 2-NZ, 15% with respect to 3-NZ, 25% with respect to R, 31% with respect to LWM, 24% with respect to QWM and 3% with respect to EWM. Also, in the case of the classification results of Table III(a) and (b), the ANOVA test (with the significance level $\alpha$ 0.05) demonstrated that traditional zoning methods and optimized methods based on Voronoi tessellations and adaptive membership functions provide statistically different classification performances on handwritten characters, in terms of both RECs and REL.

## CONCLUSION

This paper addressed the problem of membership function selection for zoning-based classification in the context of handwritten numeral and character recognition. For this purpose, zoning techniques were first introduced and static and dynamic zoning methods already presented in the literature were discussed, with specific consideration to the use of Voronoi tessellation for the design of optimal zoning topologies. After that, traditional membership functions, based on nonadaptive global strategies, were presented and a new class of adaptive zone-based membership functions was introduced. terms of minimal CF [which was defined by (10). Of course, this characteristic, which differentiated the new technique with respect to other approaches in the literature, led to different performance levels even though it made the new technique easily adaptable to different application requirements. The main idea was to have, for each zone of the zoning method, a membership function well suited for exploiting the specific characteristics of feature distribution in that zone. Successively, in order to take advantage of the potential of both adaptive zone-based membership functions and dynamic Voronoi-based zoning topologies, a new formulation of the problem of zoning design was given and a real-coded genetic approach was proposed for determining-in a unique optimization process-the adaptive membership functions and the optimal Voronoi-based topology most profitable for a given classification problem. The experimental results, carried out on standard benchmark databases of handwritten numerals and characters, demonstrated that

the new class of adaptive membership functions along with the optimal Voronoi- based zonings leads to better classification results than traditional approaches. More precisely, when handwritten numerals were considered, the best classification results were achieved for $M = 9$ zones. In this case, the REC and the REL were equal to 97% and 99%, respectively. When handwritten characters were considered, the best classification results were achieved for $M = 25$ zones. In this case, the REC and the REL were equal to 95% and 97%, respectively. These results were very satisfying compared to results concerning other approaches in the literature that employ the same datasets. For instance, when handwritten digits were considered, a REC higher than 98% was achieved only when using high-performance features and classification methods, whereas a recognition rate higher than 99% was obtained with multiclassifier systems, without assuming any rejection. Furthermore, it is worth noting that the new technique was not optimized in terms of REC, but in terms of minimal CF [which was defined by (10)]. Of course, this characteristic, which differentiated the new technique with respect to other approaches in the literature, led to different performance levels even though it made the new technique easily adaptable to different application requirements. For example, in some applications (for instance, those based on mobile handheld devices), it may be desirable to carry out the classification not considering the risk of a high error rate since the classification results are manually checked afterward; in other cases (for instance, those concerning administrative form recognition and bank check processing systems), the treatment of a substituted pattern has a high cost, hence the error rate must be kept as low as possible. A further important consideration is that the proposed technique for zoning design can be applied to any zoning-based classifier, without limitations in terms of feature type and classification technique. In fact, depending on the requirements of the specific application-that are formalized through the CF associated to the classification performance-the new technique is able to determine the optimal zoning topology and to select the best adaptive membership functions depending on the feature set and classification technique considered. Of course, a main weakness of the proposed approach is that the number of zones must be defined *a priori*. Therefore, an important advancement in the technique can be certainly achieved by optimizing also the number of zones of the zoning method. In addition, it should be pointed out that the proposed technique is general and can be applied to other image-processing tasks, since it performs-in an automatic and efficient way-optimal image segmentation by Voronoi tessellation and membership function selection, according to a given optimality criterion.

## REFERENCES


1. Plamondon, R. and S.N. Srihari. 2000. On-line and off-line handwriting recognition: A comprehensive survey, IEEE Trans. Pattern Anal. Mach. Intell., 22(1): 63-84.
2. Suen, C.Y., C. Nadal, R. Legault, T.A. Mai and L. Lam, 1992. Computer recognition of unconstrained handwritten numerals, Proc. IEEE, 80(7): 1162-1180.
3. Ferrante, A., S. Impedovo, R. Modugno and G. Pirlo. 2010. Zoning methods for hand-written character recognition: An overview, in Proc. Int. Conf. Frontiers Handwritten Recognit. Kolkata, India, pp: 329- 334.

4/. Blumenstein, M., B. Verma and H. Basli, 2003. A novel feature extraction technique for the recognition of segmented handwritten characters, in Proc. 7th Int. Conf. Document Anal. Recognit. pp: 137-141.

5. Morita, M.E., R. Sabourin, F. Bortolozzi and C.Y. Suen. 2003. Segmentation and recognition of handwritten dates: An HMM-MLP hybrid approach, Int. J. Document Anal. Recognit. 6(4): 248-262.
6. Oliveira, L.S., R. Sobourin, F. Bartolozzi and C.Y. Suen. 2002. Automatic recognition of handwritten numeral strings: A recognition and verification strategy, IEEE Trans. Pattern Anal. Mach. Intell., 24(11): 1438-1454.
7. Koerich, A., 2003. Unconstrained handwritten character recognition using different classification strategies, in Proc. Int. Workshop Artif. Neural Netw. Pattern Recognit. Florence, Italy, pp: 1-5.
8. Koerich, A.L. and P.R. Kalva, 2005. Unconstrained handwritten character recognition using metaclasses of characters, in Proc. IEEE Int. Conf. Image Process. pp: 542-545.
9. Suen, C.Y., J. Guo and Z.C. Li, 1992. Computer and human recognition of hand-printed characters by parts, in From Pixel to Features III: Frontiers in Handwriting Recognition, S. Impedovo and J. C. Simon, Eds. Amsterdam, The Netherlands: Elsevier, pp: 223-236.

10 Suen, C.Y., J. Guo and Z.C. Li. 1994. Analysis and recognition of alphanumeric handprints by parts, IEEE Trans. Syst. Man, Cybern. 24(4): 614-630.

11. Baptista, G. and K.M. Kulkarni, 1988. A high accuracy algorithm for recognition of hand-written numerals, Pattern Recognit. 21(4): 287-291.
12. Singh, S. and M. Hewitt, 2000 Cursive digit and character recognition on cedar database, in Proc. 15th Int. Conf. Pattern Recognit. Barcelona, Spain, pp: 569-572.
13. Phokharatkul, P., K. Sankhuangaw, S. Phaiboon, S. Somkuarnpanit and C. Kimpan. 2005. Off-line hand written thai character recognition using antminer algorithm, Trans. Enformatika Syst. Sci. Eng., 8: 276- 281.
14. Cha, S.H., C.C. Tappert and S.N. Srihari, 2003. Optimizing binary feature vector similarity using genetic algorithm and handwritten character recognition, in Proc. 7th Int. Conf. Document Anal. Recognit. Edinburgh, U.K. pp: 662-665.
15. Kimura, F. and M. Shridhar, 1991. Handwritten numerical recognition based on multiple algorithms, J. Pattern Recognit. 24(10): 969- 983.