2012 International Conference on Frontiers in Handwriting Recognition

Handwritten Digit Recognition by Multi-Objective Optimization of Zoning Methods

S. Impedovo, G. Pirlo, F. M. Mangini
Dipartimento di Informatica
Università degli Studi di Bari "A. Moro",
via Orabona 4, 70126, Bari, Italy
impedovo@di.uniba.it

Abstract — **This paper addresses the use of multi-objective optimization techniques for optimal zoning design in the context of handwritten digit recognition. More precisely, the Non-dominant Sorting Genetic Algorithm II (NSGA II) has been considered for the optimization of Voronoi-based zoning methods. In this case both the number of zones and the zone position and shape are optimized in a unique genetic procedure. The experimental results point out the usefulness of multi-objective genetic algorithms for achieving effective zoning topologies for handwritten digit recognition.**

Keywords: Handwritten Digit Recognition, Genetic Algorithms, NSGA II, Zoning Topology, Voronoi Diagrams.

I. Introduction

A zoning method can be defined as a partition of the pattern image that allows the extraction of local information useful for recognition. Formally speaking, given a pattern image B, a zoning $Z_M=\{z_1, z_2, ..., z_M\}$ of B can be considered as a partition of B into M sub-images, named zones, each one providing information related to a specific part of the pattern [1, 2].

Traditional zoning methods are based on static approaches in which zoning design is obtained by standard grids that are superimposed on pattern images. In this case, no a-priori information on feature distribution is considered for defining the zoning method. When dynamic zoning methods are used, zoning design is considered as an optimization problem and the optimal zoning method is generally found as the zoning which optimizes a well-suited cost function [3]. In literature both constrained and unconstrained dynamic zoning methods have been proposed so far. When constrained methods are considered, the optimal zoning topology is selected within a set of topologies having well-defined characteristics defined a–priori. For example, the system of Valveny and Lopez [4] divide the pattern image into five rows and three columns. The size of each row and column is then determined according to the discriminating capabilities of the diverse regions of the image. Radtke et al. [5] use a predefined 6x6 regular grid that can be optimized according to two diverse optimality criteria: a minimal number of non-overlapping zones and an error rate as low as possible. Gagné and Parizeau [6] use a hierarchical zoning for handwritten character classification. Their recognizer uses a multilayer perceptron as a classifier and operates on a hierarchical feature space of orientation, curvature, and center of mass primitives. The nodes of the hierarchy represent rectangular zones of their parent node whereas the tree root corresponds to the entire image pattern. Unconstrained zoning topologies are defined without imposing strict constraints on zoning topologies. For example, Dimauro et al. [7] perform zoning design according to the analysis of discriminating capability of each zone, estimated by statistical parameters. In this case a region-growing process is proposed for zoning design. Di Lecce et al. [8] designed the zoning problem as an optimization problem in which the Shannon entropy is used to evaluate the discrimination capability of each zone. Impedovo et al. [9, 10, 11] define the optimal zoning as the zoning for which the cost function associated to the classification is minimum. In this case, Voronoi Tessellation is proposed for zoning description, since it provides a means of naturally partitioning the space into zones. Strictly speaking, given a set of a finite number of M distinct points $p_1,p_2,\ldots,p_M$ in the Euclidean plane, the Voronoi Tessellation is the partition of the plane into M zones $z_1,z_2,\ldots,z_M$ that reflects proximity relationships among the set of points. In other words, each point p_i determines a region z_i that is the locus of points which are closer to p_i than to any other point of the set, according to the Euclidean distance [9, 10]. In addition, the role of membership functions for zoning based classification is also analyzed [11, 12, 13, 14, 15].

Although several approach to topology design have been proposed so far, little attention has been devoted until now to the automatic definition of the optimal number of zones for a given classification task [16]. This paper discusses the effectiveness of multi-objective genetic algorithms applied to Voronoi-based zoning design in order to combine, in a unique optimization process, the selection of the optimal number of zones along with the optimal Voronoi zones for a given classification problem. In particular, in this paper the effectiveness of the Non-dominated Sorting Genetic Algorithm (NSGA II) is considered for finding the optimal zoning method [17]. The experimental tests, carried out in the field of handwritten digit recognition, demonstrate that the optimal zoning methods derived from multi-objective optimization technique outperform traditional zoning methods based on single-objective optimization techniques. In addition, some issues related to the convergence of the genetic approach are highlighted and discussed.

The paper is organized as follows. Section II presents the problem of zoning-based classification. Section III presents the problem of optimal Voronoi-based zoning design by multi-objective genetic algorithm,. The experimental results, carried out in the field of handwritten digit recognition, are reported in Section IV. The conclusion of the paper is reported in Section V.

978-0-7695-4774-9/12 $26.00 © 2012 IEEE
DOI 10.1109/ICFHR.2012.209

II. ZONING-BASED CLASSIFICATION OF HANDWRITTEN DIGITS

Let $Z_M=\{z_1, z_2, ..., z_M\}$ be a zoning method, and let us consider the classification of a pattern x into one class of the set $\Omega=\{C_1 ,..., C_K\}$ using the feature set $F=\{f_1,...,f_T\}$. In this case x can be described by the feature matrix A_x of T rows (features) and M columns (zones) [16]:

$$A_x = \begin{bmatrix} A_x(1,1) & A_x(1,2) & ... & A_x(1,j) & ... & A_x(1,M) \\ A_x(2,1) & A_x(2,2) & ... & A_x(2,j) & ... & A_x(2,M) \\ ... & ... & ... & ... & ... & ... \\ A_x(i,1) & A_x(i,2) & ... & A_x(i,j) & ... & A_x(i,M) \\ ... & ... & ... & ... & ... & ... \\ A_x(T,1) & A_x(T,2) & ... & A_x(T,j) & ... & A_x(T,M) \end{bmatrix}$$

with

$$A_x(i,j) = \sum_{instances\ of\ f_i\ in\ x} w_{ij} \qquad (1)$$

being w_{ij} the weight that defines the degree of influence of an instance of feature f_i (detected in x) on zone z_j.

Now, if the Winner-takes-All strategy is used, the weights are defined according to the following simple rule:

- $w_{ij} = 1$ iff the instance of f_i has been detected in z_j;
- $w_{ij} = 0$ otherwise.

In this case the classification phase can be performed considering the matrix A_x , by a simple k-nn classifier (k=1).

III. MULTI-OBJECTIVE OPTIMIZATION FOR ZONING DESIGN

In this paper the problem of optimal zoning design is considered as the result of a multi-objective optimization problem. More precisely, it is formulated as the problem to define the optimal zoning for which the cost function associated to the classification is minimum and in which the number of zones is minimum. Therefore, the two cost functions to be minimized are the following [16]:

1) $CF_1(Z_M)= \mu\cdot Err(Z_M)+ Rej(Z_M)$ (2a)
   where $Err(Z_M)$ is the error rate (estimated on the learning set); $Rej(Z_M)$ is the rejection rate (estimated on the learning set); the coefficient μ is the cost value associated to the treatment of an error with respect to a rejection.
2) $CF_2(Z_M)= M$ (2b)
   where M is the number of zones of the zoning method Z_M.

Of course, since Voronoi Diagram is used for zoning description, the problem of optimal zoning design becomes:

Find the set of Voronoi points $\{p^*_1, p^*_2, ..., p^*_M\}$ with minimum cardinality (M minimum) and for which it results :

$$CF(Z^*_M) = \min_{\{ZM\}} CF(Z_M) \qquad (3)$$

with:

- $Z^*_M = \{z^*_1, z^*_2,\dots, z^*_M\}$, z^*_j being the Voronoi region corresponding to p^*_j , $\forall j=1,2,\dots,M$;
- $Z_M = \{z_1, z_2,\dots, z_M\}$, z_j being the Voronoi region corresponding to p_j , $\forall j=1,2,\dots,M$.

In order to solve this optimization problem the Non-dominated Sorting Genetic Algorithm (NSGA II) has been considered [17]. In this case, individuals of the genetic population are evaluated by non-dominance and by spatial distribution criteria in order to derive a set of non-dominated solutions evenly spaced (such set is known as the *Pareto-front*), which represents the best configurations for the two objectives being optimized. In this paper, order to solve the optimization problem (3), the non-dominated sorting genetic algorithm (NSGA II) reported in Figure 1 is adopted [17]:

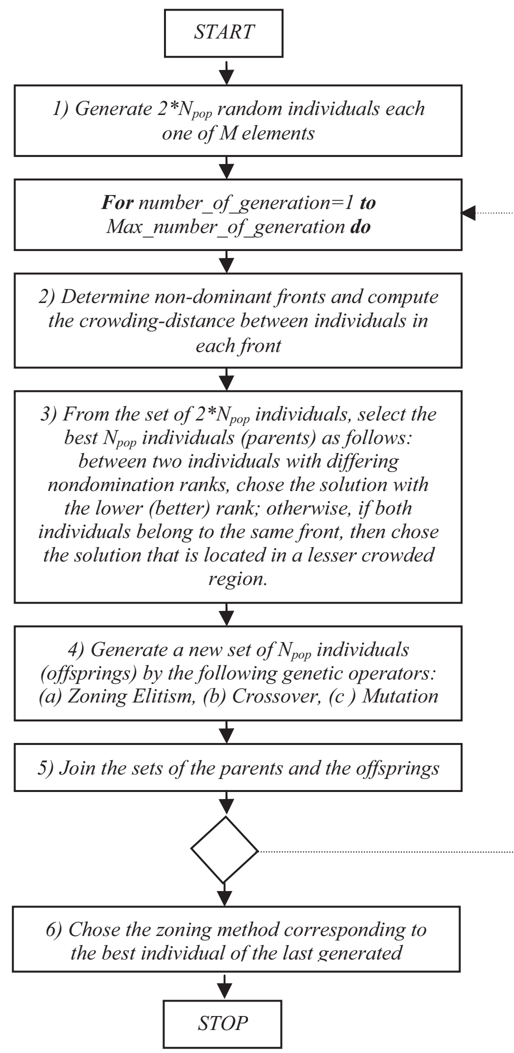


Figure 1. NSGA II for Zoning Design

In the following, a detailed description of each phase of the algorithm is reported.

1) In this phase the initial population

$$\text{Pop}=\{\Phi_1, \Phi_2, ...,\Phi_\iota, ... ,\Phi_{2*N\text{Pop}}\} \tag{4}$$

for the genetic algorithm is created. Each individual is a vector

$$\Phi_i = \left\langle p_1, p_2,.., p_j,.., p_M \right\rangle \tag{5}$$

where each element $p_j=(x_j,y_j)$ is a Voronoi point corresponding to the zone z_j of $Z_M=\{z_1, z_2, ..., z_M\}$.

2) In this phase the non-dominant fronts are determined and the crowding-distance between individuals in each front is computed. These measures are useful to characterize each solution and to select the best ones. More precisely, non-dominated fronts are determined by the approach of ref. [18], that is based on two entities that are computer for each solution Φ: n_Φ , that counts the number of solutions which dominate the solution p; S_Φ, a set of solutions that the solution Φ dominates. Figure 2 shows an example of non-dominant fronts of solutions.

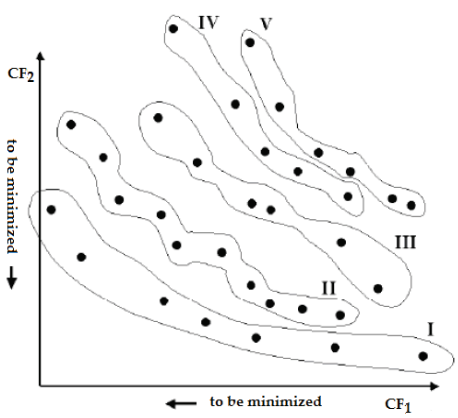


Figure 2. Example of non-dominant fronts

The crowding distance has been introduced as an estimator of the density of solutions surrounding a particular solution in the population [18]. The computation of the crowding distance requires sorting the population according to each objective function value in ascending order of magnitude. Thereafter, for each objective function, the boundary solutions (solutions with smallest and largest function values) are assigned an infinite distance value. All other intermediate solutions are assigned a distance value equal to the absolute normalized difference in the function values of two adjacent solutions. This calculation is continued with other objective functions. The overall crowding-distance value is calculated as the sum of individual distance values corresponding to each objective. Of course, each objective function is normalized before calculating the crowding distance.

3) In this phase the best N_{Pop} individuals are selected from the set of $2*N_{pop}$ individuals. Precisely, since every individual in the population has two attributes: (I) nondomination rank (i_{rank}) , (II) crowding distance ($i_{distance}$); a partial order relation is defined as follows: between two solutions with differing nondomination ranks, the solution with the lower (better) rank should be preferred. Otherwise, if both solutions belong to the same front, then the best solution is that located in a lesser crowded region.

4) In this phase the new set of N_{Pop} individuals (offsprings) is generated. This is performed according to the following genetic operations [9, 16]:

<u>a) Zoning Elitism.</u> The zoning elitist technique selects randomly some individuals of the population and removes the element corresponding to the less significant zone from the individual. The significance of an element (i.e. a zone) is here defined according to the number of instances a feature in the learning patterns lies in that zone. The lower the number of instances in a zone the lower the significance of that zone. This operation allows the production of zoning methods with a reduced number of zones. It is worth noting that this strategy does not apply to two-zone zoning methods.

<u>b) Crossover.</u> One-point crossover is used to combine information from diverse individuals. Let

$$\Phi_i = \left\langle p_1{}^a, p_2{}^a,.., p_j{}^a,.., p_{M_1}{}^a \right\rangle \tag{6a}$$

and

$$\Phi_i = \left\langle p_1{}^b, p_2{}^b,.., p_j{}^b,.., p_{M_2}{}^b \right\rangle \tag{6b}$$

be two individuals selected for crossover, the two offspring individuals

$$\Phi_i = \left\langle \cancel{p}_1{}^a, \cancel{p}_2{}^a,.., \cancel{p}_j{}^a,.., \cancel{p}_{M_2}{}^a \right\rangle \tag{7a}$$

and

$$\Phi_i = \left\langle \cancel{p}_1{}^b, \cancel{p}_2{}^b,.., \cancel{p}_j{}^b,.., \cancel{p}_{M_1}{}^b \right\rangle \tag{7b}$$

of the next generation are obtained as follows:

- $\cancel{p}^a{}_s=p^a{}_s$, for s=1,…j ; $\cancel{p}^a{}_s=p^b{}_s$, for s=j+1,…M2
- $\cancel{p}^b{}_s=p^b{}_s$, for s=1,…j ; $\cancel{p}^b{}_s=p^a{}_s$, for s=j+1,…M1

being s a random integer in the range [1, min(M_1,M_2)].

<u>c) Mutation.</u> A non-uniform mutation operator has been used. Let us consider the individual Φ_ι and an element selected for mutation, according to a mutation probability Mut_prob. The non-uniform mutation changes p_j in the new element $\tilde{p}_j = (\tilde{x}_j, \tilde{y}_j)$ that is defined as follows:

$$\begin{cases} \tilde{x}_j = x_j + \delta \cdot \cos(\varphi) \\ \tilde{y}_j = y_j + \delta \cdot \sin(\varphi) \end{cases} \tag{8}$$

where:

- φ is a random value generated according to a uniform distribution, $\varphi \in [0,2\pi[$;

- δ is a displacement determined according to the a non-linear model that allows the operator to search the space almost uniformly initially, when *iter* is small, and locally in later stages [9].

5) In this phase the two sets of parents and offsprings are joint together.

Steps from (2) to (5) are repeated until *Max_number_of_generation* successive populations of individuals are generated.

6) In this phase the optimal zoning is obtained by the best individual of the last-generated population.

IV. Experimental Results

The experiments have been carried out using the set of handwritten numeral digits $\Omega_1=\{0,1,2,3,4,5,6,7,8,9\}$ extracted from the CEDAR database [19]. Precisely, 18467 learning patterns (BR directory) and 2189 testing patterns (BS directory) were considered for the test. The feature set $F=\{f_1,\ldots,f_9\}$ is considered for pattern description, where [9]:

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 end-points;*
f_7 - *vertical-down end-points;*
f_8 - *horizontal-right end-points;*
f_9 - *horizontal-left end-points*.

The following parameter values have been considered for the genetic algorithm: $N_{Pop}=10$; *Max_number_of_generation* = 100; *Mut_prob* = 0.35; $\delta\_displ=5$; $b=1.0$; $\lambda\_displ=0.5$, $c=3.0$ and $\mu=1$ [16].

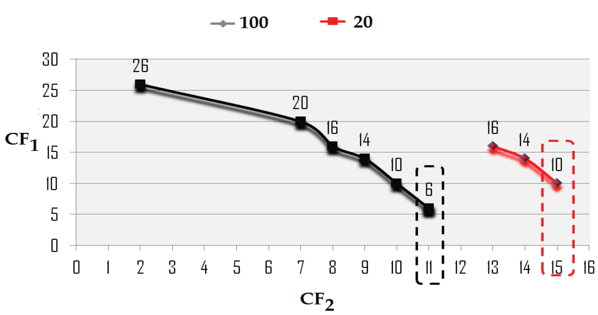
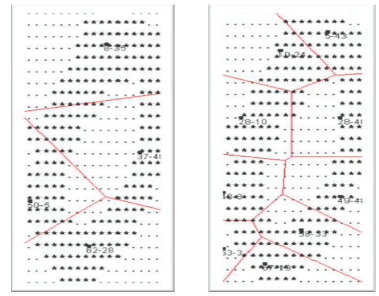


Figure 3. The Pareto Front: 20 vs 100 iterations

Figure 3 reports the Pareto fronts of the genetic algorithm, as the number of iterations increases. In particular the cases of 20 and 100 iterations are displayed. It is easy to observe the improvement of both CF_1 and CF_2 as the number of iterations increases. Furthermore, after 100 iterations the algorithm finds seven solutions on the Pareto front (Rank 1). Examples of generated zoning topologies are reported in Figure 4 , for the case of (a) M= 4, (b) M=9.



(a) (b)

Figure 4. Voronoi-based Zoning Topologies

Figure 5 reports the Pareto Front –Rank 1 and Rank 2 of non-dominated solutions after 100 iterations.

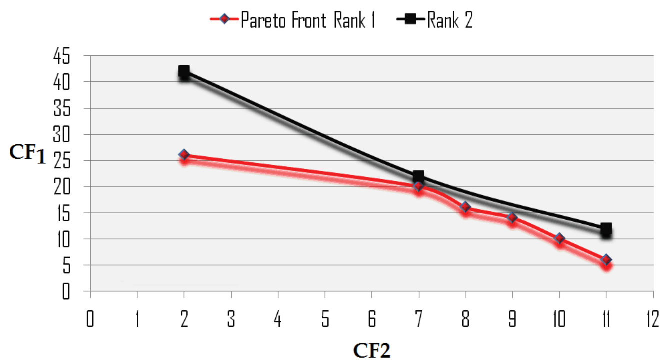


Figure 5. Performance

The results show that using single-objective optimization techniques (M=2,4,6,9 and 16 considered), the best zoning method is for M=9. In this case the error rate is equal to 14%. When the multi-objective optimization technique is considered the optimal number of zones is M=11. In this case the error rate is equal to 6%.

V. Conclusion

This paper addresses the problem of optimal zoning design, for handwritten digit recognition by using multi-objective genetic algorithms. The new strategy allows to define, in a unique optimization process, the zoning with optimal (minimum) number of zones and best performances. The strategy, that is based on the non-dominant sorting genetic algorithm (NSGA II), has shown to be effective with respect to traditional optimization approaches.

REFERENCES

[1] O.D.Trier, A.K.Jain , T.Taxt, "Feature Extraction Methods For Character Recognition – A Survey", *Pattern Recognition*, Vol. 29, n.4, pp. 641-662, 1996.

[2] C.Y.Suen, J.Guo, Z.C.Li , "Analysis and Recognition of Alphanumeric Handprints by Parts", *IEEE T-SMC*, Vol. 24, n. 4, pp. 614-630, 1994.

[3] S. Impedovo, G. Pirlo, R. Modugno, A. Ferrante, "Zoning Methods for Handwritten Character Recognition: An Overview", Proc. 12th ICFHR, Kolkata (India), 16-18 Nov. 2010, pp. 329-334.

[4] E. Valveny and A. Lopez, "Numeral recognition for quality control of surgical sachets", 7th ICDAR, 2003, pp. 379-383.

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

[6] C. Gagné and M. Parizeau, "Genetic engineering of hierarchical fuzzy regional representations for handwritten character recognition", *IJDAR*, 2006, Vol. 8, n. 4, pp. 223-231.

[7] G. Dimauro, S.Impedovo, G. Pirlo, A.Salzo, "Zoning Design for Handwritten Numeral Recognition", in *Lecture Notes in Computer Sciences*, Vol. 1311, Ed. by A. Del Bimbo, Springer Verlag, Berlin, 1997, pp. 592-599.

[8] V. Di Lecce, G. Dimauro , A. Guerriero , S. Impedovo , G. Pirlo, A.Salzo, "Zoning Design for Handwritten Numeral Recognition", Proc. of Seventh International Workshop on Frontiers in Handwirting Recognition (IWFHR-7), pp. 583-588, 2000.

[9] S, Impedovo, M.G. Lucchese, G. Pirlo, "Optimal Zoning Design by Genetic Algorithms",IEEE T-SMC-A, Vol.36, n.5, pp.833-846, 2006.

[10] A. Ferrante, S. Impedovo, G. Pirlo, C. Trullo "On the Design of Optimal Zoning for Pattern Classification", *Proc. of the 11th IAPR Conference on Frontiers in Handwriting Recognition*, August 19-21, 2008, Concordia University, Montreal, Quebec, Canada, CENPARMI Press, pp. 130-134.

[11] S. Impedovo, A. Ferrante, R. Modugno, G. Pirlo, "Feature Membership Functions in Voronoi-based Zoning", Emergent Perspectives in Artificial Intelligence, Eds. R. Serra and R. Cucchiara, Lecture Notes in Artificial Intelligence, Vol. 5883, 2009, ISSN: 0302-9743, Springer Publ., pp. 202-211.

[12] S. Impedovo, R. Modugno, G. Pirlo, "Membership Functions for Zoning-based Recognition of Handwritten Digits", *Proc. International Conference on Pattern Recognition*, Istanbul, Turkey, August 2010, pp. 1876-1879.

[13] S. Impedovo, R. Modugno, G. Pirlo, "Analysis of membership Functions for Voronoi-based Classification", *Proc. International Conference on Frontiers in Handwriting Recognition 2010*, Kolkata, India, 2010, pp. 220-225.

[14] S. Impedovo and G. Pirlo, "Tuning between Exponential Functions and Zones for Membership Functions Selection in Voronoi-based Zoning for Handwritten Character Recognition", Proc. 2011 International Conference on Document Analysis and Recognition, Bejing (China), Sept. 18-20, 2011, pp. 997-1001.

[15] G. Pirlo and D. Impedovo, "Adaptive Membership Functions for Hand-Written Character Recognition by Voronoi-based Image Zoning", IEEE Transactions on Image Processing, 2012.

[16] G. Pirlo and D. Impedovo, "Voronoi-based Zoning Design by Multi-Objective Genetic Optimization", Proc. 10th DAS 2012, March 27-29, 2012, Gold Coast, Australia, pp. 220-224.

[17] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II", IEEE T-EC, Vol. 6, No. 2, April 2002, pp. 182-197.

[18] K. Deb, *Multiobjective Optimization Using Evolutionary Algorithms*.

[19] Chichester, U.K.: Wiley, 2001.J. Hull, "A database for handwritten text recognition research", *IEEE T-PAMI*, Vol. 16, n. 5, pp. 550–554, 1994.