

HANDWRITTEN CHARACTER RECOGNITION USING NONSYMMETRICAL PERCEPTUAL ZONING

CINTHIA O. A. FREITAS*, LUIZ S. OLIVEIRA† and FLÁVIO BORTOLOZZI‡

Pontifical Catholic University of Paraná (PUCPR)

Rua: Imaculada Conceição

1155 — 80215-901 — Curitiba (PR) — Brazil

**cinthia@ppgia.pucpr.br*

†soares@ppgia.pucpr.br

‡fborto@ppgia.pucpr.br

SIMONE B. K. AIRES

Universidade Tecnológica Federal do Paraná (UTFPR-PG)

Av. Monteiro Lobato

km 4 — 84016-210 — Ponta Grossa (PR) — Brazil

sbkaminski@pg.cefetpr.br

In this paper we present an alternative strategy to define zoning for handwriting recognition, which is based on nonsymmetrical perceptual zoning. The idea is to extract some knowledge from the confusion matrices in order to make the zoning process less empirical. The feature set considered in this work is based on concavities/convexities deficiencies, which are obtained by labeling the background pixels of the input image. To better assess the nonsymmetrical zoning we carried out experiments using four different zonings strategies. Experiments show that the nonsymmetrical zoning could be considered as a tool to build more reliable handwriting recognition systems.

Keywords: Visual perception; zoning mechanism; confusion matrix; handwriting recognition.

1. Introduction

The handwritten character recognition is a special subject and has become important as ICR systems (Intelligent Character Recognition) become more powerful and commercially available. However, there is a gap between human reading capabilities and the recognition systems. According to the literature,^{11,22,24} it is necessary to explore and capture information from human perception to design new systems.

The character tendency to be confused conveys important information to define the perceptual similarity of characters. The basic idea is that two characters that look a lot alike will often be confused with each other. Figure 1 presents this idea considering the following characters: “U” and “V”, “O” and “Q”. A good strategy of perceptual similarity should predict which pairs of characters are confused and

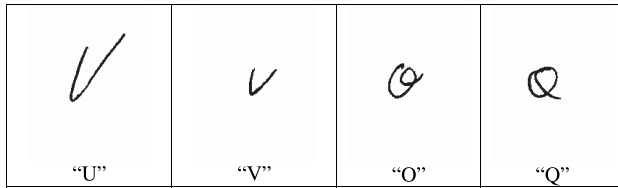


Fig. 1. Similarity between letters: (a) “U” and “V”; (b) “O” and “Q”.

which are not. The confusion matrix computed by the recognition process can provide us a solution for understanding and predicting the confusion. The idea is to analyze the confusion and use this information to build up recognition systems.

To understand the recognition process, let us consider two different aspects: features and spatial frequency.^{2,6,21} The former is based on features and it uses a checklist approach and it claims that characters are represented in the nervous system as a set of features, lines and contours of various orientations. This kind of approach states that if two characters have many features in common, they tend to be confused, otherwise, they do not. It can be noticed from Table 1 that the character “C” and “G” have features in common (convex segment and horizontal open). The only feature that discriminates these two characters is the bar-horizontal. This is not a reliable feature once it depends strongly on the writing style, as depicted in Fig. 2. The latter takes into account the spatial frequency and helps to understand that confusions tend to occur when the characters have similar spatial frequencies.

Table 1. List of features distinguishing characters of the alphabet.

	Features			
	A	C	G	T
<i>External</i>				
1. Horizontal				1
2. Vertical				1
3. Slant (/)	1			
4. Slant (\)	1			
5. Convex segment		3	3	
<i>Open</i>				
6. Horizontal		1	1	
7. Vertical				
8. Wedged, horizontal				
9. Wedged, vertical				
10. Internal protrusion				
11. Intersection, internal	2			
12. Bar-horizontal	1		1	
13. Bar-slant, crossing				
14. Symmetry, vertical	1			1
15. Symmetry, horizontal		1		

“C”				“G”			
c	c	C	c	G	G	G	G

Fig. 2. Characters “C” and “G”: difference based on bar-horizontal feature.

Since we consider features to discriminate classes of character, it is necessary first to compile a checklist of features. In other words, it is necessary to decide which features should make up the list. Table 1 presents a list of the features distinguishing some characters of the alphabet, such as “A”, “C”, “G” and “T”. A complete list is presented in Ref. 21. The feature approach considers that each character in the alphabet has a set of features to be distinguished.

When designing recognition systems we take into account that perception depends on cooperative interaction between the processing of global (low-frequency) and local (high-frequency) information. The character or word recognition is an example of stimulus that contains both kinds of information.

We are experts in recognition of characters from early childhood onwards. But, when we observe only a part of the character, its identification is not that obvious. In the first observation, we process global information, while in the second, we process local information. We go through the characters stored in the brain, choose a possible candidate which contains the same part, and then try to add other parts to it to form this possible character.²⁴ Another possibility is to decompose a possible character in the same way as the given partition does. If the first one does not fit, try another one, and so on until a suitable part is found.²⁴ Based on this concept, methods for local information analysis on partitions of the character, also known as zoning, have been proposed to evaluate the recognition rates of the distinct parts of characters. Most of the works define empirically symmetrical zoning while others use complex and expensive search mechanisms to find the best zoning.^{9,20,27}

This paper discusses the zoning mechanism taking into consideration the methodologies used to define this kind of approach. In a general way, the researchers define zoning as an empirical process or as a result of searching algorithms. In this paper, we present an alternative methodology to define zoning, which is based on the concept of nonsymmetry. The features considered in this work are based on concavities/convexities deficiencies, which are obtained by labeling the background pixels of the input image. Four different perceptual zoning (symmetrical and nonsymmetrical) are then discussed. Experiments show that the nonsymmetrical zoning could be considered as a part of the solution in handwritten character recognition. Differently from Refs. 10 and 24 we have observed that more cells in the zoning do not bring more confusing parts, when those cells are nonsymmetrical. Our experimental results presented in Sec. 4 demonstrate that this strategy is reliable and very useful to help design zoning.

The paper is divided into five sections. Section 2 presents the handwritten character recognition problem summarizing the most important visual perception concepts to our study and our approach to solve the preprocessing and feature extraction stages. Section 3 summarizes the zoning mechanism concepts and presents our approach based on perceptual zoning considering a nonsymmetrical strategy. In Sec. 4, the database used in the experiments is presented and the experimental results are discussed. The experimental results are supported by the confusion matrices analysis and how to use this information as a zoning mechanism. Finally, in Sec. 5, our conclusion and a plan for future works are presented.

2. Handwriting Recognition

Since the late 1960s, research on recognition of unconstrained handwritten characters has made impressive progress and many systems have been developed, particularly in machine printed and online character recognition. However, there is still a significant performance gap between humans and machines in the recognition of offline totally unconstrained handwritten character recognition.

Character recognition techniques have potential application in any domain where a large mass of document image-bearing texts must be interpreted or analyzed. Conventionally, such images are processed by human operators who act according to what has been written or simply key in what they read onto a computer system that carries out further processing, say of postal address. However, automation of the entire process requires high recognition rates, as well as maximum reliability.

Generally speaking, an offline handwriting recognition system includes four stages: image preprocessing, segmentation, feature extraction and classification. Preprocessing is primarily used to reduce noise or variations of handwritten characters. Segmentation consists in locating and extracting the handwritten information from the image. Feature extraction is essential for data representation and extracting meaningful features for later processing. Classification assigns the characters to one of the several classes.

Considering the influence on recognition performance, the feature extraction plays a very important role in handwriting recognition. This has led to the development of a variety of features for handwriting recognition and their recognition performances have been reported by several authors.^{4,25,26} Therefore, the idea in this section is to present the handwriting recognition problem summarizing the most important visual perception concepts to this work and to present our approach to solve the preprocessing, feature extraction and classification stages.

2.1. *Visual perception concepts*

Here we introduce a summary about the visual perception concepts related to handwriting recognition. The Gestalt Theory describes the **principles of organization**,

which tend to encourage the emergence of perceptual forms and promote the **grouping** of those forms, segregated from their surroundings. This theory is beyond the scope of this paper. However, an excellent introduction to this subject can be found in Ref. 21.

Generally speaking, people organize what they see. Because only particular properties promote grouping, these properties may constitute the basic elements of perception. We call these visual elements **primitives**. And these primitives need to be extracted from the pattern or shape to be recognized and to build up a feature vector or to generate graphs, string of codes or sequence of symbols.

On the assumption that knowledge of these primitives might reveal how grouping processes work, many researchers^{2,6} try to identify vision's primitives. The Gestalt theory postulated that the human being has a tendency to interpret a visual stimulus as a complete scene. This tendency is known in Gestalt theory as closure concept. In handwriting recognition, this approach is called *Global Approach*.

Another approach considered in the literature is the structuralist.¹⁹ It treated form perception as an analytical process, whereby complex forms were decomposed into small, simple elements. For handwriting recognition, it is the same approach that segments words into letters or pseudo-letters, called by the researchers as *Local* or *Analytical Approach*. In summary, this work takes into account the Gestalt theory applying first a *Global* feature extraction based on concavities/convexities deficiencies, and a *Local* perceptual zoning mechanism. The zoning mechanism allows scrutinizing the elements (shapes) individually.

2.2. Preprocessing and feature extraction stages

The system uses as input a 256 grey-level image. Then, a preprocessing step, which is composed of binarization¹⁵ and a bounding box definition is applied. After the preprocessing step we apply a feature extraction stage. Selection of a feature extraction method is probably the most important factor to achieve high recognition performance. For this reason, such a subject has gained considerable attention of the scientific community. A good survey about feature extraction can be found in Refs. 11 and 26. The literature of handwriting recognition shows us basically three classes of features: (a) grey-level or binary values of all the pixels in an image usually represented by a N -dimensional vector, where N is the number of pixels in the image, (b) structural features of the image, which are typically perceptual entities of the character such as bends, end points, intersections, loops, measures of concavity, distance information and directional features, (c) statistical features which are the result of global mathematical transformations such as moments,^{1,5} Fourier descriptors,²³ and wavelet transforms.¹⁷ To build more reliable systems, many researchers have turned towards the combination of statistical and structural features in a same structure of classification (e.g. in a one-shot classifier).^{4,26}

The feature set considered in this work is based on Concavities/Convexities deficiencies. This feature set puts on evidence the topological and geometrical properties

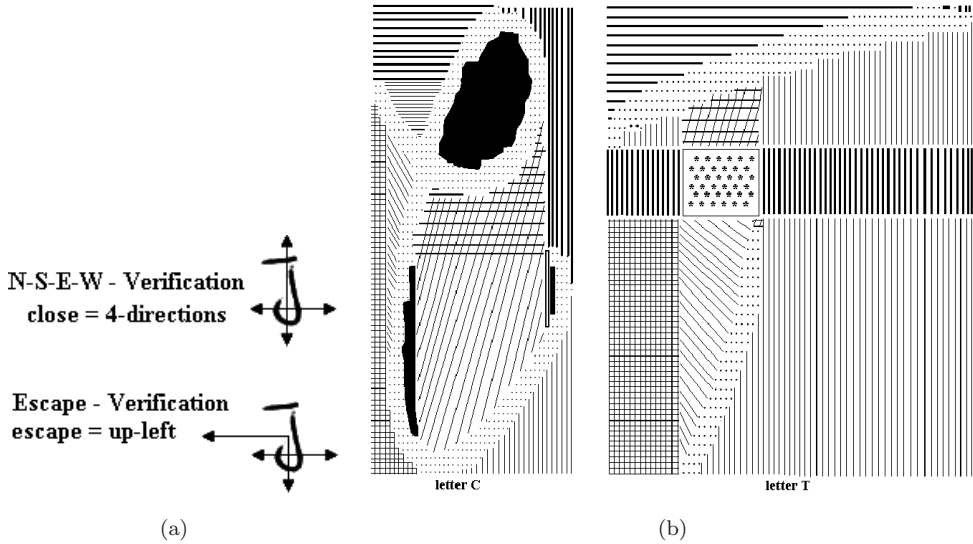


Fig. 3. Feature extraction: (a) verification process, and (b) labeling process.

of the shape to be recognized. The labeling process can be done in two ways: (a) labeling the background pixels of the input images¹⁶ and (b) labeling line segment information.³

The basic idea of concavity/convexity deficiencies is the following: for each white pixel in the image of the character we search in four-directions [North, South, East, West — Fig. 3(a)], the number of black pixels that it can reach as well as which directions the black pixel is not reached. When black pixels are reached in all directions, we branch out in four auxiliary directions in order to confirm if the current white pixel is really inside a closed contour. Figure 3(b) shows the obtained result after the labeling process. The entire and definitive alphabet has 24 different symbols. The next section presents the classification stage.

2.3. Classification stage

The baseline system uses a Class-Modular architecture feedforward MLP (Multiple Layer Perceptron) in the classification stage. Oh and Suen¹³ have demonstrated that class-modular NN can produce better results than just one single NN. Based on this and other works we have done,^{7,14} we have chosen this architecture for our experiments. Moreover, experimental results (Sec. 4.2) support this architecture.

In class-modular architecture a single task is decomposed into multiple subtasks and each subtask is allocated to an expert network. In this paper, as well as in Ref. 13, in the class-modular classification, the K -classes classification problem is decomposed into K two-class subproblems, each for one of the K classes. A two-class subproblem is solved by the two-class classifier specifically designed for the

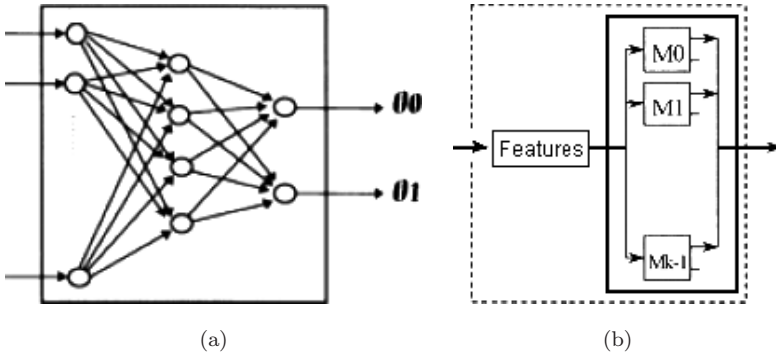


Fig. 4. Class-modular architecture: (a) subnetwork, and (b) whole network with K modules.

corresponding class. The two-class classifier is only responsible for one specific class and discriminates that class from the other $K - 1$ classes. In the class-modular framework, K two-class classifiers solve the original K -class classification problem cooperatively and the class decision module integrates the outputs from K two-class classifiers.

In Fig. 4(a), we can see the MLP architecture for two-class classifier. The modular MLP classifier consists of K subnetworks, M_i for $0 \leq i \leq K - 1$, each one responsible for one of the K classes. The architecture for the entire network composed of K subnetworks is shown in Fig. 4(b).

The next section presents the zoning mechanism concepts and discusses about the symmetrical and nonsymmetrical strategies.

3. Zoning Mechanism

Several authors have presented zoning mechanisms or regional decomposition methods to investigate the recognition rates and discover potential candidates when confusion occurs at a given part. Let us analyze the human brain during the character reading process. Humans often concentrate on the significant parts of the characters for effective and efficient reading. But, do we really know which are the significant parts? Where are the significant parts of the characters located? This section summarizes the zoning mechanism concepts and presents our approach based on perceptual zoning considering a nonsymmetrical strategy.

3.1. Zoning mechanism concepts

Zoning is a simple way to obtain local information and it is used to extract topological information from patterns.⁸ The goal of the zoning is to obtain local characteristics instead of global characteristics. A zoning is a partition of the control box of the pattern (i.e. the smallest rectangle containing the pattern); the elements of such partitions are used to identify the position in which features of the pattern

are detected. The zoning design, that is the way in which the partition into the bounding box is defined, can be considered in two different ways:

- *Fixed* or *symmetrical*: the bounding box is divided into zones of equal size zoning,^{3,8,10,12,24,30}
- *Variable* or *nonsymmetrical*: the bounding box is nonuniformly divided according to pattern density.^{9,20,27}

Depending on the domain of application or the experience of the researcher the zoning can be carried out exclusively on the basis of intuitive motivations⁹ or based on the easier manner, i.e. fixed or symmetrical zoning.

Suen *et al.*²⁴ and Li *et al.*¹⁰ applied a zoning mechanism in their experiments using hand printed characters. They analyzed four different configurations where the characters were divided into Z parts, say $Z = 2, 4,$ and 6 as presented in Fig. 5. They observed that character “D” always lies on the top (100%), characters “A”, “K” and “G” give higher recognition rates (100%) than “P”, “I” and “T” (54%) and, the recognition rates considering $Z = 2LR, 2UD, 4$ and 6 were: 86.12%, 85.88%, 61.73%, and 42.91%, respectively. The authors comment about the case 2LR for “Y” and explain that this zoning is perfect for recognition; but it brings a difficulty to “B” because the left half is confusing with “E”. Therefore, it should be noticed that different partitions may produce big differences in the recognition rates. In addition, more partitions bring more confusing parts. For instance, for $Z = 6$, a character is confused with six characters, e.g. letter “B” is confused with: “C”, “G”, “J”, “O”, “S”, “U”.

Blumenstein *et al.*³ presented a study where the character is zoned into six windows of equal size. Morita *et al.*¹² used the same strategy for handwritten digits. Xiang *et al.*³⁰ extracted features dividing the character input image from car plates into $n \times m$ ($n = 4, m = 4$) zones. Koerich and Kalva⁸ examined the input image dividing the handwritten character into 3×2 zones. These authors applied a symmetrical zoning mechanism as local feature extractor.

Other authors used complex and expensive search mechanisms to find the best zoning. Lecce *et al.*⁹ designed the zoning problem as an optimization problem in which the Shannon entropy is used to evaluate the discrimination capability of each zone when a specific feature set is considered.

Radtke *et al.*²⁰ presented an automatic approach to define the zoning for offline handwritten digit recognition, using Multi-Objective Evolutionary Algorithms (MOEAs). The idea was to provide a self adaptive methodology to define

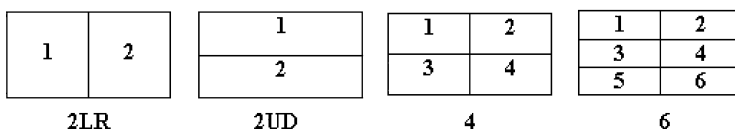


Fig. 5. Zoning mechanism: $Z = 2$ (Left Right and Up Down), 4, and, 6 parts.

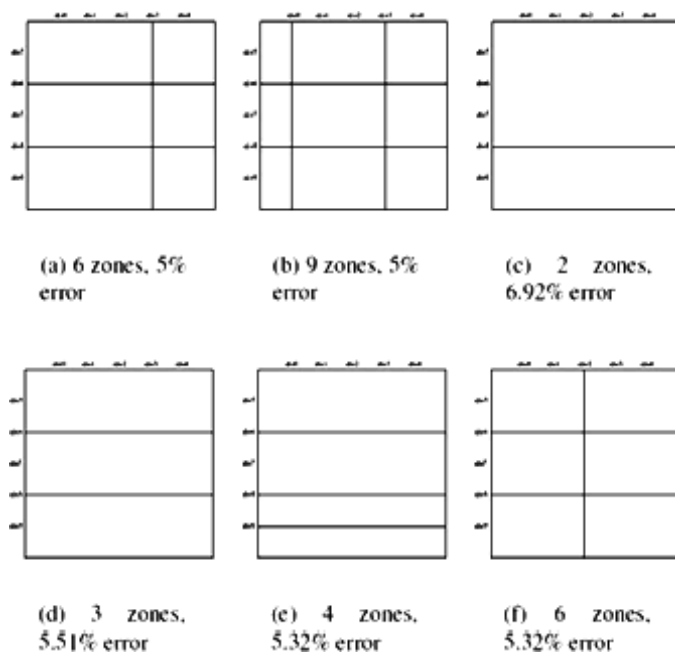


Fig. 6. Zoning strategies comparison.

the zoning strategy with m nonoverlapping zones and an acceptable error rate, with no need of human intervention during the search stage. The best result was obtained using six zones composed by three symmetrical rows (horizontal: 2/6, 2/6, 2/6) and three nonsymmetrical columns (vertical: 1/6, 3/6, 2/6), as presented in Fig. 6.

Valveny and López²⁷ applied a zoning mechanism to digit recognition located in sachets containing surgical materials which pass through a computer vision system performing several quality controls. In this case, the size of each region is not constant. Each image is divided into five rows and three columns. However, the size of all rows and all columns is not always the same. Each row and column has its own size in order to locate them in the most discriminated areas of the image.

3.2. Nonsymmetrical perceptual zoning

In this paper we analyze the significant parts of the characters using the confusion matrix obtained in the recognition process. The idea consists in looking for the relationship between the regions and the confusions, thus allowing us to understand which parts are making up the confusions. Our approach does not use any complex and extensive search algorithm to design the zoning. It is a perceptual zoning based on confusions among specific regions in the handwritten character images, which are provided by the confusion matrices. We take into account that the confusion

matrix is a consistent analysis of classifier behavior and it can provide quantitative representation for each classifier in terms of recognition. Beyond, the confusion matrix allows us to design a nonsymmetrical zoning.

Our methodology started with experiments based on the following studies applying $Z = 4$, then we analyzed the confusion matrix looking for the relationship between the regions and the confusions. Analyzing the confusion matrix shown in Fig. 7, we can observe that the main confusions in this kind of zoning are: "B", "D" and "O"; "C" and "E"; "D" and "O"; "H" and "M"; "G" and "Q"; "T", "E" and "J"; "J" and "D"; "K" and "M"; "N" and "W"; "R" and "A"; "S" and "D"; "W", "U" and "V"; "X" and "K"; "Y" and "X". Some examples of those confusions are also depicted in Fig. 7. This analysis emphasizes that the confusions do not occur in the same region, such as "G" and "Q" in lower region or "B", "D" and "O" in middle region.

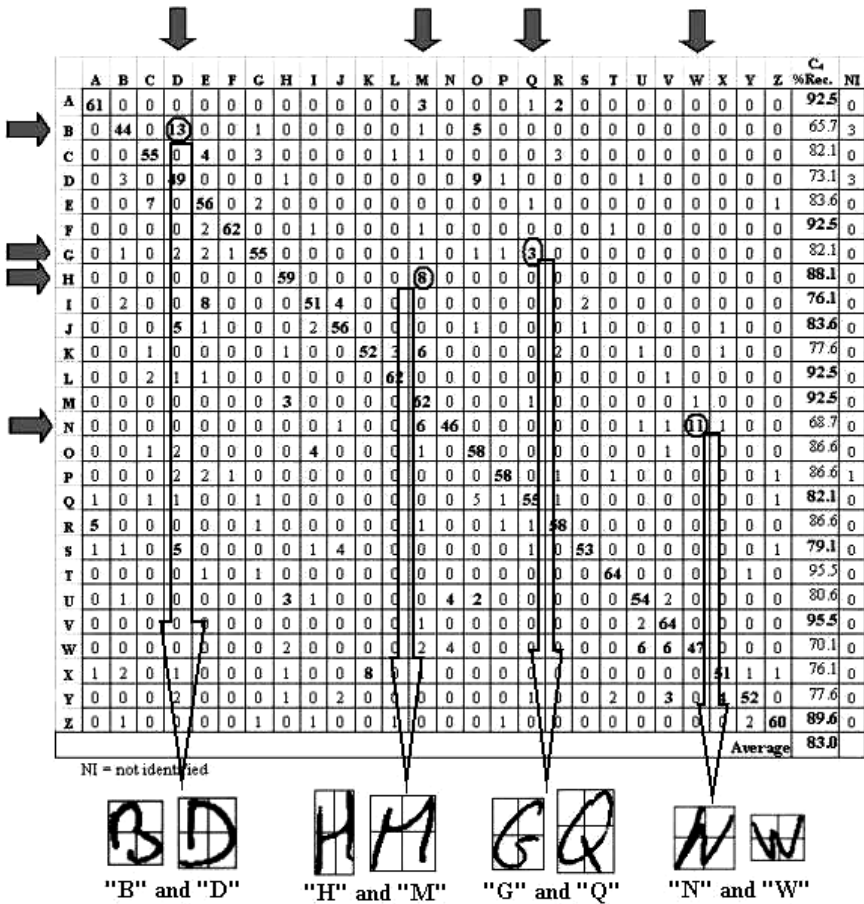


Fig. 7. Confusion matrix and parts of the letters: $Z = 4$.

Considering the problems pointed out in Fig. 7, it is possible to identify the regions where the confusions occur, hence, defining new strategies of zoning:

- lower region as “G” and “Q” [Fig. 8(a)];
- right half region as “E” and “I”; “N” and “W”; “X” and “K” [Fig. 8(b)];
- middle region as “B”, “D” and “O”; “U”, “V” and “W” [Fig. 8(c)].

The idea in observing the lower and right half regions is to give more emphasis to similar parts by increasing the number of zones. This proposal was evaluated with zoning mechanism based on $Z = 5V$ (Vertical – right half) and $Z = 5H$ (Horizontal – lower), as depicted in Figs. 9(b) and 9(c).

Following the same concept, we have investigated seven-part zoning ($Z = 7$), as shown in Fig. 9(d). In this case, the idea is to solve confusions among nonsymmetrical shapes but representing differentially the middle region in the character, such as “B” and “D”; “N” and “W”; “Y” and “X”. In a general way, this zoning mechanism contributes to solve the confusions among characters that need best representation in middle zone, such as “B”, “C”, “D”, “E”, and “F” as demonstrated by experimental results presented in Sec. 4.

Since the confusion matrix provides the necessary information for the analysis process, this kind of methodology allows us to extract some knowledge from the matrices in order to make the zoning process less empirical.

Differently from Refs. 10 and 24, we have observed that more cells in the zoning do not bring more confusing parts, when those cells are nonsymmetrical. Our

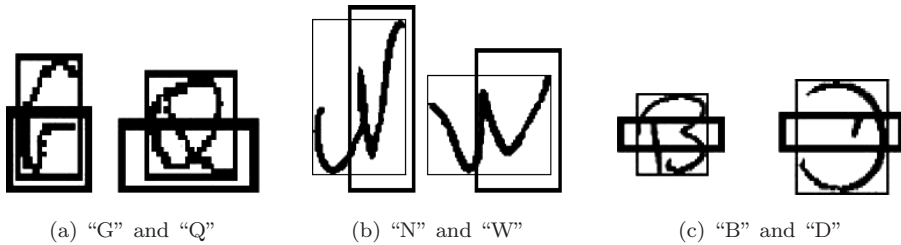


Fig. 8. Confusion analysis based on parts in the characters: (a) lower, (b) right half, and (c) middle parts.

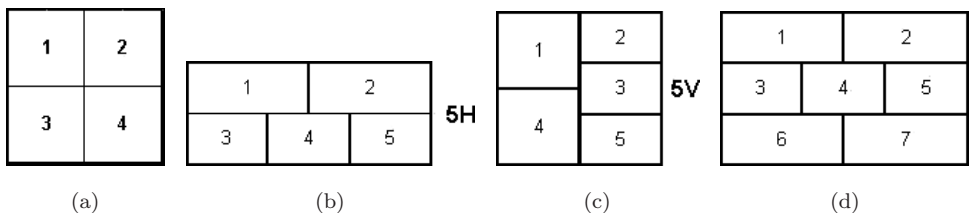


Fig. 9. Zoning mechanism: (a) $Z = 4$, (b) $Z = 5H$, (c) $Z = 5V$, and (d) $Z = 7$ parts.

experimental results presented in Sec. 4 demonstrate that this strategy is reliable and very useful to help design zoning.

As described so far, our idea is to use the perceptual information contained in the confusion matrices to propose a perceptual nonsymmetrical strategy. An important contribution of this work is to provide to the researchers an alternative strategy to define zoning for handwriting recognition.

4. Experimental Results and Discussions

This section presents the IRONOFF database used during the experiments, the obtained results based on the nonsymmetrical perceptual zoning mechanism, and the discussions about the relationship between the zones and the confusion, using the confusion matrix to this analysis. We are interested in designing a method that does not use just the first-order information (classifier’s score output) to evaluate the system performance. The idea is to use information from the confusion matrix for each individual classifier and to understand which parts of the character are making up the confusions.

4.1. IRONOFF database

The experiments were carried out using the handwritten character database from IRESTE/University of Nantes (France), called IRONOFF (IREste ON/OFF Dual Database), which is composed of 26 classes of uppercase characters from Form B: B27 ... B52 fields.^{28,29} The IRONOFF database was selected because it is fully cursive. It was collected from about 700 writers, mainly of French nationality. The experiments were carried out using three subsets, which we called the training, validation and testing sets. Their compositions are as follows: 60%, 20% and 20% for training, validation, and testing, respectively. The database sums up 10,510 images of handwritten characters. Figure 10 illustrates the variability of writing style in IRONOFF database.

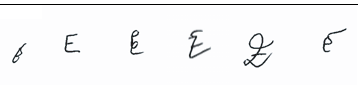
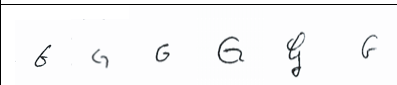
 <p>“A”</p>	 <p>“B”</p>
 <p>“C”</p>	 <p>“D”</p>
 <p>“E”</p>	 <p>“G”</p>

Fig. 10. Samples from IRONOFF database.

4.2. Experimental results

In order to get some basis for comparison, the first zoning we have used was $Z = 4$ [Fig. 9(a)] and the recognition rates for each class are presented in Table 2. This table allows us to understand that there is no best classifier based on a specific zoning. These results are a good indicator for studies on ensemble of classifiers as a future work.

The average recognition rate for $Z = 4$ is 83.0%. This zoning achieves best results for the following characters when this approach is compared with other zoning mechanisms (see Table 2): “A”, “F”, “H”, “I”, “J”, “L”, “M”, “Q”, “S”, “V” and “Z”. It is important to observe that the majority of these characters are symmetrical in terms of vertical (“A”), horizontal (“I”) or both (“H”) directions.

The confusions and the analysis allow us to experiment the five-part zoning (5V-vertical and 5H-horizontal — Figs. 9(b) and 9(c) trying to provide more information to solve confusions among shapes that are not symmetric on both sides (horizontal and vertical), such as “G” and “Q”; “D” and “O”; “Y” and “X”. The recognition rates for the classifiers $Z = 5H$ and $Z = 5V$ are 81.7%, 80.9%, respectively.

Table 2. Recognition rate (%).

Character	$Z = 4$	$Z = 5H$	$Z = 5V$	$Z = 7$
A	92.5	86.6	89.6	91.0
B	65.7	64.2	74.6	79.1
C	82.1	79.1	68.7	88.1
D	73.1	65.7	68.7	82.1
E	83.6	85.1	89.6	95.5
F	92.5	91.0	89.6	92.5
G	82.1	86.6	80.6	80.6
H	88.1	85.1	70.1	76.1
I	76.1	71.6	76.1	71.6
J	83.6	79.1	79.1	82.1
K	77.6	76.1	77.6	80.6
L	92.5	89.6	86.6	91.0
M	92.5	82.1	85.1	88.1
N	68.7	77.6	70.1	86.6
O	86.6	89.6	88.1	83.6
P	86.6	92.5	91.0	94.0
Q	82.1	64.2	76.1	80.6
R	86.6	89.6	88.1	91.0
S	79.1	79.1	79.1	76.1
T	95.5	97.0	97.0	97.0
U	80.6	85.1	82.1	86.6
V	95.5	82.1	88.1	82.1
W	70.1	74.6	65.7	79.1
X	76.1	74.6	70.1	79.1
Y	77.6	89.6	85.1	82.1
Z	89.6	88.1	88.1	86.6
Average	83.0	81.7	80.9	84.7

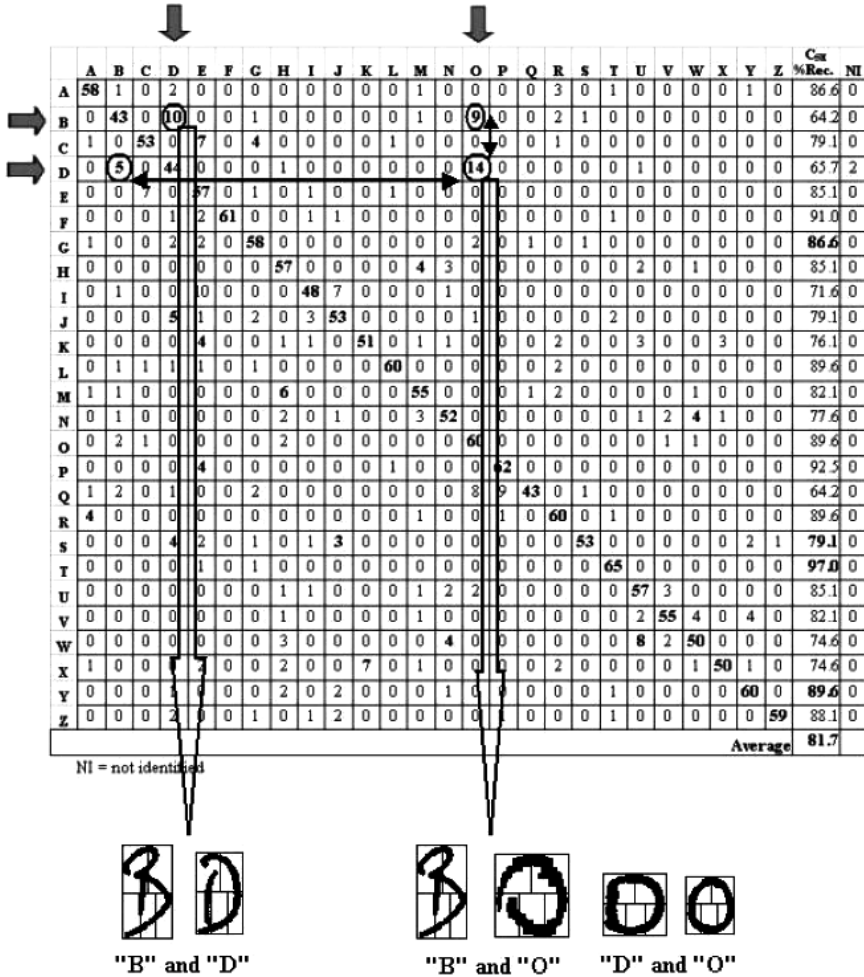


Fig. 11. Confusion matrix and parts of the letters: Z = 5H.

Figures 11 and 12 present the confusion matrices for these classifiers. Observing these figures we can extract the letters that represent the main confusions. It is important to see that these zonings have the same problems when compared to Z = 4. Now, it is important to observe that the characters that present best results (“G”, “T” and “Y”) applying Z = 5H use the information contained at the lower zone of the character to solve the confusions. On the other hand, Z = 5V just presents best results to character “O”, demonstrating that the information contained in the right half region of the character is important to discriminate “B” and “O” and “D” and “O”. It is possible to observe that Z = 5H zoning works better than Z = 5V when the confusion lies between “B” and “O” or “D” and “O”. Summarizing, the information contained on the right half region of the character is more reliable to the recognition system.

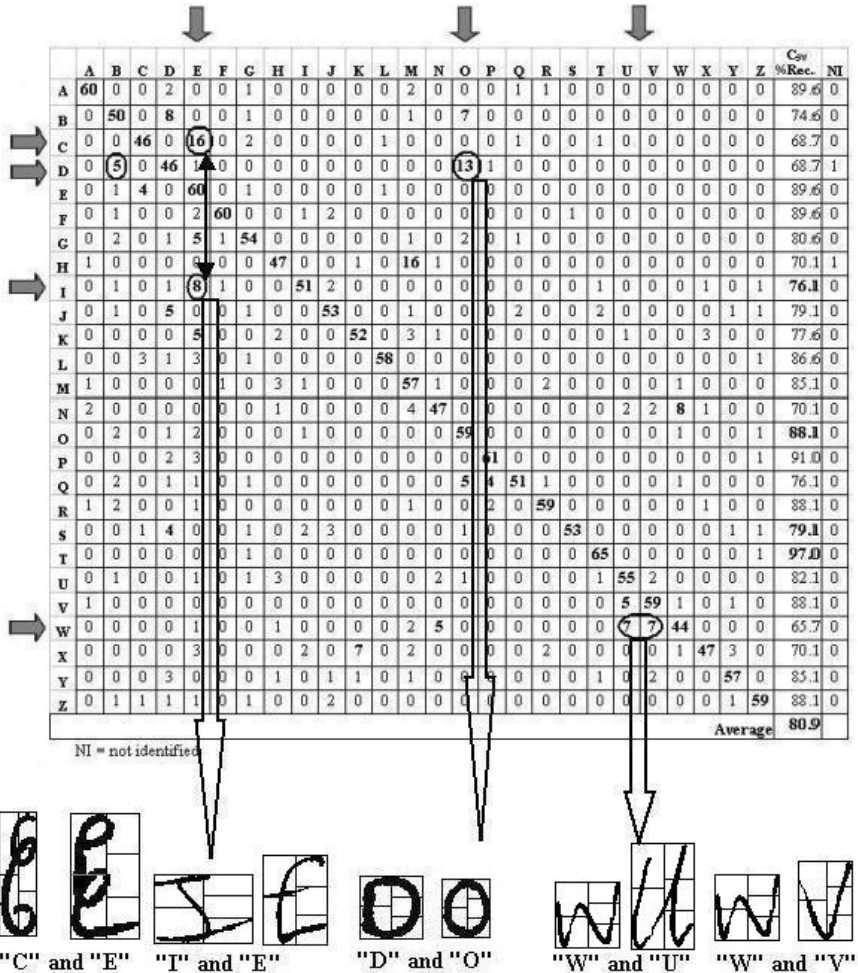
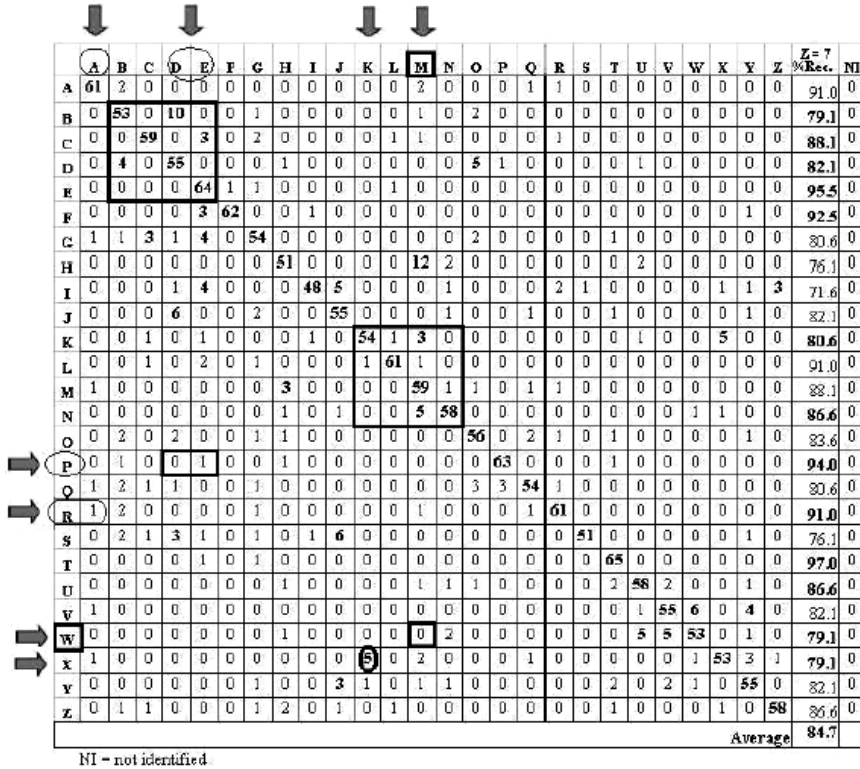


Fig. 12. Confusion matrix and parts of the letters: Z = 5V.

Figure 13 presents the confusion matrix and shows some samples of the best results obtained applying the nonsymmetrical perceptual zoning built based on the confusion matrices analysis. The average recognition rate for Z = 7 is 84.7%. The seven-part zoning is better for the following classes: "B", "C", "D", "E", "K", "N", "P", "R", "U", "W", and "X" (see Table 2). The best result observed for the character "P" (94.0%) is due to reduction of the confusion with characters "D" and "E". Another important result observed for character "R" (91.0%) is due to reduction of the confusion with character "A". In these cases, the information contained in the middle region (Z = 7) were more discriminative than the lower region (Z = 5H) or symmetrical approach (Z = 4).



Best Results: "B", "C", "D", "E", "K, "N", "P", "R", "U", "W", "X".

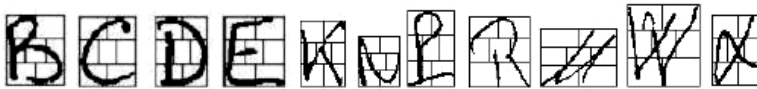


Fig. 13. Confusion matrix and parts of the letters: Z = 7.

Based on Fig. 13, it is possible to observe that some confusions were reduced (“D” and “O”) but others were increased (“H” and “M” when compared to Z = 4). Since this zoning mechanism contributes to solve the confusions based on a better representation of the middle zone, the experimental results confirm what was proposed in the methodology.

In order to support the choice of the modular NN strategy we have run an experiment using Z = 4 and a single NN (MLP, which 26 outputs). Such architecture reached a recognition rate of 64% which is a lot worse than the results achieved by the Class-Modular architecture. Some classical confusions in this neural architecture are: “B” and “D”; “E” and “F”; “Q” and “G”; “W”, “M”, and “N”. Another experiment was considered applying Z = 7 and a single NN (MLP, which 26 outputs). This architecture reached a recognition rate of 54.9% which is worse than the results achieved by the Class-Modular architecture. Some classical confusions

in this neural architecture are: “A” and “H”; “M” and “H”; “V” and “W”; “W” and “U”; “X” and “K”. These results demonstrated that a unique NN does not solve all confusions among the characters.

It can be noticed from Table 2 that a nonsymmetrical zoning mechanism yields the best recognition rate (84.7%), that demonstrates this approach could be considered as an alternative to zoning definition. As mentioned in Sec. 3, we are experts in the recognition of characters, words and digits from early childhood. However, when we observe only a part of them, its identification is not that obvious. In the first observation, we are processing global information while in the second we are processing local information. This kind of analysis is provided by this approach, when the system applies the feature extraction (labeling the background pixels) and then uses the zoning mechanism to compute the labels to each Z -part. Therefore, we have designed a strategy capable of a Global and Local analysis, as presented in Sec. 2. Due to the nonsymmetrical zoning, we provide the system a better representation of the parts where the confusions are evident.

The results of our experiments are comparable to other published methods. For example, Poisson *et al.*¹⁸ using an MLP reached a recognition rate of 87.1% for the same database.

Without any contextual information, any observer could have problems to discriminate the samples presented in Fig. 14. The fourth character, which is an “M”, can be read as an “H”, “M” or “U”. These confusions, which are difficult to solve even for human beings should be used to design reliable recognition systems. Based on confusion analysis and which regions of the characters make up these confusions, the researchers can focus their efforts on important regions to achieve better representation and recognition rates for characters.

5. Conclusion

In this paper, we have explored the perceptual zoning based on the confusion matrix and its information about the confusion regions of the characters. The perceptual regions have been verified and used to define a nonsymmetrical perceptual zoning, observing that the similarities are evident among the classes. The Gestalt principles help researchers understand how human perception can be used to improve handwriting recognition systems. The study based on zoning mechanisms presented in

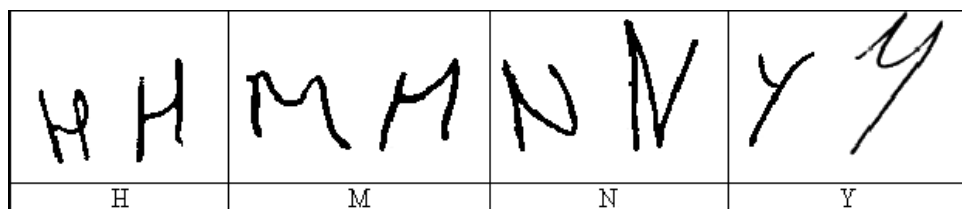


Fig. 14. Characters from IRONOFF database: confusions.

this paper can contribute to reduce some of the confusions found in handwriting recognition systems. Our approach is based on global and local representations of the characters demonstrating by experiments that the use of information contained in confusion matrix can improve the representation of the parts that make up the confusions. Therefore, this kind of approach can improve the performance of the system. Differently from Refs. 10 and 24, we have demonstrated that more cells in the zoning does not necessarily bring more confusion parts.

Finally, the experiments have shown the viability of our approach, which focuses on human visual perception. Future work will provide the validation of our approach in a different database, such as NIST or Brazilian databases. We also plan to compare the idea of perceptual zoning with search algorithms such as genetic algorithms. Based on Table 2, it is clear that there is no best classifier based on a specific zoning. These results are a good indicator for studies on ensemble of classifiers.

Acknowledgments

The authors wish to thank the Pontifícia Universidade Católica do Paraná (PUCPR, Brazil) and the Universidade Tecnológica Federal do Paraná (UTFPR-PG, Brazil) which have supported this work.

References

1. R. R. Bailey and M. Srinath, Orthogonal moment features for use with parametric and non-parametric classifiers, *IEEE Trans. Patt. Anal. Mach. Intell.* **18**(4) (1996) 389–399.
2. J. Beck, K. Prazdny and A. Rosenfeld, *A Theory of Textual Segmentation, Human and Machine Vision* (Academic Press, NY, 1983), pp. 1–38.
3. M. Blumenstein, B. Verma and H. Basli, A novel feature extraction technique for the recognition of segmented handwritten characters, *7th Int. Conf. Document Analysis and Recognition, ICDAR'03* (2003), pp. 137–141.
4. L. Heutte, T. Paquet, J. V. Moreau, Y. Lecoutier and C. Oliver, A structural/statistical feature based vector for handwritten character recognition, *Patt. Recogn. Lett.* **19** (1998) 629–641.
5. M. K. Hu, Visual pattern recognition by moment invariant, *IEEE Trans. Inform. Th.* **8** (1992) 179–187.
6. B. Julesz, Early vision and focus attention, *Rev. Mod. Phys.* **63** (1991) 735–772.
7. M. N. Kapp, C. Freitas and R. Sabourin, Handwritten Brazilian month recognition: an analysis of two NN architectures and a rejection mechanism, *9th Int. Workshop on Frontiers in Handwriting Recognition (IWFHR-9)* (2004), pp. 209–214.
8. A. L. Koerich and P. R. Kalva, Unconstrained handwritten character recognition using metaclasses of characters, *IEEE Int. Conf. Image Processing (ICIP)*, **2** (2005) 542–545.
9. V. Lecce, G. Dimauro, A. Guerriero, S. Impedovo, G. Pirlo and A. Salzo, Zoning design for handwritten numerical recognition, *7th Int. Workshop on Frontiers in Handwriting Recognition* (2000), pp. 583–588.
10. Z. C. Li, C. Y. Suen and J. Guo, A regional decomposition method for recognizing handprinted characters, *IEEE Trans. Syst. Man Cybern.* **25** (1995) 998–1010.

11. S. Madhavanath and V. Govindaraju, Preceptual features for off-line handwritten word recognition: a framework for prediction, representation and matching, *Advances in Pattern Recognition* (1998), pp. 524–531.
 12. M. E. Morita, R. Sabourin, F. Bortolozzi and C. Y. Suen, Segmentation and recognition of handwritten dates: an HMM-MLP hybrid approach, *Int. J. Docum. Anal. Recogn.* **6** (2004) 248–262.
 13. I.-S. Oh and C. Y. Suen, A class-modular feedforward neural network for handwriting recognition, *Patt. Recogn.* **35** (2002) 229–244.
 14. J. J. Oliveira Jr, M. N. Kapp, C. O. A. Freitas, J. M. Carvalho and R. Sabourin, Handwritten month word recognition using multiple classifiers, *XVII Brazilian Symp. Computer Graphics and Image Processing (SIBGRAPI)* (2004), pp. 82–89.
 15. N. Otsu, A threshold selection method from gray-level histograms, *IEEE Trans. Syst. Man. Cybern.* **9**(1) (1979) 63–66.
 16. J. R. Parker, *Algorithms for Image Processing and Computer Vision* (John Wiley, 1997).
 17. S. Pittner and S. V. Kamarthi, Feature extraction from wavelet coefficients for pattern recognition tasks, *IEEE Trans. Patt. Anal. Mach. Intell.* **21**(1) (1999) 83–88.
 18. E. Poisson, C. Viard-Gaudin and P. M. Lallican, Multi-modular architecture based on convolutional neural networks for online handwritten character recognition, *Int. Conf. Neural Information Processing* **5** (2002), pp. 2444–2448.
 19. J. R. Pomerantz, Perceptual organization in the information processing, in *Perceptual organization*, eds. M. Kubovy and J. R. Pomerantz (Hillsdale, NJ, Erlbaum, 1981), pp. 141–180.
 20. P. V. W. Radtke, L. S. Oliveira, R. Sabourin and T. Wong, Intelligent zoning design using multi-objective evolutionary algorithms, *7th Int. Conf. Document Analysis and Recognition (ICDAR2003)* (2003), pp. 824–828.
 21. R. Sekuler and R. Blake, *Perception*. 3rd edn. (McGraw-Hill Inc., 1994).
 22. L. Schomaker and E. Segers, A method for the determination of features used in human reading of cursive handwriting, *6th Int. Workshop on Frontiers in Handwriting Recognition* (1998), pp. 157–168.
 23. M. Shridhar and A. Badreldin, High accuracy character recognition algorithm using fourier and topological descriptors, *Patt. Recogn.* **17**(5) (1984) 515–524.
 24. C. Y. Suen, J. Guo and Z. C. Li, Analysis and recognition of alphanumeric handprints by parts, *IEEE Trans. Syst. Man Cybern.* **24** (1994) 614–631.
 25. C. Y. Suen, Réflexions sur la reconnaissance d’écriture cursive, *CIFED’98*, 1998, pp. 1–18.
 26. O. D. Trier, A. K. Jain and T. Taxt, Feature extraction methods for character recognition — a survey, *Patt. Recogn.* **29**(4) (1996) 641–662.
 27. E. Valveny and A. López, Numeral recognition for quality control of surgical sachets, *7th Int. Conf. Document Analysis and Recognition, ICADR’03* (2003), pp. 379–383.
 28. C. Viard-Gaudin, The ironoff user manual, IRESTE, University of Nantes, France (1999).
 29. C. Viard-Gaudin, P. M. Lallican, S. Knerr and P. Binter, The ireste on/off (ironoff) dual handwriting database, *IEEE Int. Conf. Document Analysis and Recognition* (1999), pp. 455–458.
 30. P. Xiang, Y. Xiuzi and Z. Sanyuan, A hybrid method for robust car plate character recognition, *IEEE Int. Conf. Syst. Man Cybern.* (2004), pp. 4377–4737.
-



Cinthia O. A. Freitas received the B.S. degree in civil engineering in 1985 from Universidade Federal do Paraná (UFPR - Curitiba - Brazil), M.Sc. degree in electrical engineering and industrial informatics from Centro Federal de

Educação Tecnológica do Paraná (CEFET/PR-Curitiba-Brazil) in 1990, and Ph.D. degree in applied computer science from Pontifícia Universidade Católica do Paraná (PUCPR-Curitiba-Brazil) in 2001. Since 1985, she has been a professor in Computer Science and Computer Engineering Departments at PUCPR. Currently, she is a Full Professor and researcher in the post-graduated program in applied computer science (PPGIA) at PUCPR. She is a member of Brazilian Computer Society and IAPR.

Her research interests are in handwriting recognition, symbol recognition, document image analysis, and forensic science.



Luiz S. Oliveira received the B.S. degree in computer science from UnicenP (Curitiba-Brazil), M.Sc. degree in electrical engineering and industrial informatics from Centro Federal de Educação

Tecnológica do Paraná (CEFET/PR-Curitiba-Brazil), and Ph.D. degree from the Ecole de Technologie Superieure — Université du Quebec (Montreal-Canada) in 1995, 1998, and 2003, respectively. From 1994 to 1998, he was a system analyst at HSBC Bank (Curitiba-Brazil), where he worked in financial systems. From 2000 to 2003, he was a visiting scientist at the Centre for Pattern Recognition and Machine Intelligence. In 2005, he joined the Pontifícia Universidade Católica do Paraná (PUCPR-Curitiba-Brazil), where he is currently associate professor of computer science. He is a member of Brazilian Computer Society and IAPR.

His research interests include pattern recognition, computer vision, and evolutionary computation.



Flávio Bortolozzi

obtained the B.S. degree in mathematics in 1977 from Pontifícia Universidade Católica do Paraná (PUCPR-Curitiba-Brazil), a B.S. degree in civil engineering in 1980 from PUCPR, and a Ph.D.

degree in system engineering (computer vision) from the Université de Technologie de Compiègne, France, in 1990, where he worked on the trinocular vision. From 1994 to 1999, he was the head of the department of informatics, and the dean of the college of exact sciences and technology at PUCPR. From 2000 to 2005, he was the Vice-rector for Research and Postgraduate Programs at PUCPR. Currently, he is a Full Professor at the Computer Science Department at PUCPR. Since 1996, he is a member of CNPq — Conselho Nacional de Pesquisa Científica e Tecnológica in Brasil. He is the author (and co-author) of more than 200 scientific publications including journals and conference proceedings. He was the Symposium Chair of BSDIA'97 (Brazilian Symposium on Document Image Analysis, Curitiba, Brazil). He was nominated as Conference Chair of the next ICDAR'07 (9th Int. Conf. Document Analysis and Recognition) that will be held in Curitiba, Brazil in 2007.

His research interests are in computer vision, handwriting recognition, document image analysis, educational multimedia, and hypermedia.



Simone B. K. Aires

received the B.S. in informatics in 1999 from Universidade Estadual de Ponta Grossa (UEPG-Ponta Grossa-Brazil), M.Sc. degree in applied computer science from Pontifícia Universidade Católica

do Paraná (PUCPR-Curitiba-Brazil) in 2005. In 2002, she joined the Universidade Federal Tecnológica do Paraná (UFTPR/Ponta Grossa-Brazil), where she is currently associate professor of technology of information system.

Her research interests are in handwriting recognition, symbol recognition, document image analysis, and forensic science.