



# Experimental study of a novel neuro-fuzzy system for on-line handwritten UNIPEN digit recognition

E. Gómez Sánchez<sup>a</sup>, J.A. Gago González<sup>a</sup>, Y.A. Dimitriadis<sup>a,\*</sup>,  
J.M. Cano Izquierdo<sup>b</sup>, J. López Coronado<sup>b</sup>

<sup>a</sup> Department of Signal Theory, Communications and Telematics Engineering, School of Telecommunications Engineering, University of Valladolid, C/ Real de Burgos S/N, 47011 Valladolid, Spain

<sup>b</sup> Department of Systems Engineering and Control, School of Industrial Engineering, University of Valladolid, Paseo del Cauce S/N, 47011 Valladolid, Spain

Received 6 June 1997; revised 27 October 1997

## Abstract

This paper presents an on-line hand-printed character recognition system, tested on datasets produced by the UNIPEN project, thus ensuring sufficient dataset size, author-independence and a capacity for objective benchmarking. New preprocessing and segmentation methods are proposed in order to derive a sequence of strokes for each character, following suggestions of biological models for handwriting. Variants of a novel neuro-fuzzy system, FasArt (Fuzzy Adaptive System ART-based), are used for both clustering and classification. The first task assesses the quality of segmentation and feature extraction techniques, together with an analysis of Shannon entropy. Experimental results for classification of the *train\_r01\_v02* UNIPEN dataset show real-time performance and a recognition rate of over 85%, exceeding slightly Fuzzy ARTMAP performance, and 5% inferior to the rate achieved by humans. © 1998 Elsevier Science B.V. All rights reserved.

*Keywords:* On-line handwriting recognition; UNIPEN; Fuzzy neural networks; ART

## 1. Introduction

On-line handwriting recognition has been a research focus point for a long time (Tappert et al., 1990), but specially lately due to the expansion of pen-based Personal Digital Assistants or communicators. However, lack of a common and sufficient data base posed a serious obstacle to the development of reliable systems that could handle the variability of handwriting and permit the comparison of the various methods proposed in the literature. In this sense

the UNIPEN project, that started in 1993 with the participation of 40 organizations, has already collected and organized 5 million characters for an upcoming benchmark and its public distribution (Guyon et al., 1994). We then used data from the UNIPEN project in order to test the proposed preprocessing and classification techniques.

Our recognition system uses a stroke-based model for the description of handwritten characters. According to handwriting generation models (Plamondon, 1995), components (handwritten specimen between two successive pen-lifts) are made of strokes with velocity profiles described by a given (delta-

\* Corresponding author.

lognormal) equation. Such a biologically motivated solution has already been employed in other schemes (Parizeau and Plamondon, 1995; Schomaker, 1993; Morasso et al., 1993) with promising results. In Section 2, we present a segmentation technique that permits us to obtain a consistent and short sequence of strokes for each character, that at the same time would provide efficient classification and reduced computational complexity. Furthermore, in the same section we discuss and compare candidate feature sets that would better fit to this classification problem.

The system proposed here for digit classification resembles an ART-based architecture for character recognition, previously proposed by us (Dimitriadis and López Coronado, 1995). Now we introduce a new neuro-fuzzy architecture called FasArt (Cano et al., 1996) in pattern recognition problems, and compare it with Fuzzy ARTMAP (Carpenter et al., 1992). With both architectures the STORE (Bradski et al., 1992) memory model is used to store the sequence of categories related to the sequence of strokes that form a character. The overall architecture of the proposed neuro-fuzzy system will be briefly explained in Section 3.

Classification results with the UNIPEN data are exposed in Section 4 and some common error causes are commented, while concluding remarks and current work related to the building of an allograph-lexicon are presented in Section 5.

## 2. Character segmentation and feature extraction

This recognition system parts from isolated characters, segmenting them into strokes and then extracting some features that represent these strokes. This section explains what decisions were taken with respect to segmentation and feature extraction.

There are several reasons behind segmentation of components into strokes, mainly the following two:

- According to Plamondon handwriting generation theories (Plamondon, 1995) a trained human generates a sequence of superimposed simple movements to write a character, each of these simple movements with a delta-lognormal velocity profile. Thus, segmentation consists in fragmenting the velocity profile of a component in simpler

delta-lognormal-like profiles, each corresponding to one stroke.

- If we observe characters, such as “b”, “d”, “p” and “q”, similar segments in their trajectory can be found, e.g. a horizontal line and a left-opened semi-circumference. This can lead us to think that any character could be built with strokes from a finite base (Kerrick and Bovik, 1988). From this point of view, segmentation consists in obtaining such a base and in achieving a suitable character representation.

Based on the concept of strokes as elemental traces, minima in velocity profiles have been used as segmentation points in the literature (Schomaker, 1993; Morasso et al., 1993). These algorithms require preprocessing with real data, due to trembling, signal quantization or inaccurate sampling rate that may introduce irregularities, e.g., repeated points even if the pen is not still. Usually low-pass filters are used on the velocity signal, since these irregularities are high frequency signals. Nevertheless, this technique does not retire all repeated points, besides the fact that it cannot be interpreted geometrically.

In our system we used a geometric iterative segmentation algorithm, based on the search of velocity minima. Initially all repeated points are retired, thus eliminating redundant geometric information and avoiding proliferation of velocity minima (this process causes loss of time instant information but not of time sequential information). This produces a *generalized velocity* signal without zero values but with minima in the same locations. Then minima of the *generalized velocity* signal are considered as candidates for critical points of segmentation and iteratively these candidates (except component begin and end points) are evaluated. A candidate point  $C_i$  is retired if:

- the angle between  $\overline{C_{i-1}C_i}$  and  $\overline{C_iC_{i+1}}$  does not correspond to a significant change in trajectory direction, or
- the length of the segment  $\overline{C_{i-1}C_i}$  or  $\overline{C_iC_{i+1}}$  is smaller than a threshold.

This two-phase segmentation (selection of candidate segmentation points and iterative selection of the significant ones, in terms of global geometry) is based on (Al-Emami and Usher, 1990). It eliminates velocity minima produced by spatial quantization, selects the best segmentation point in a region with

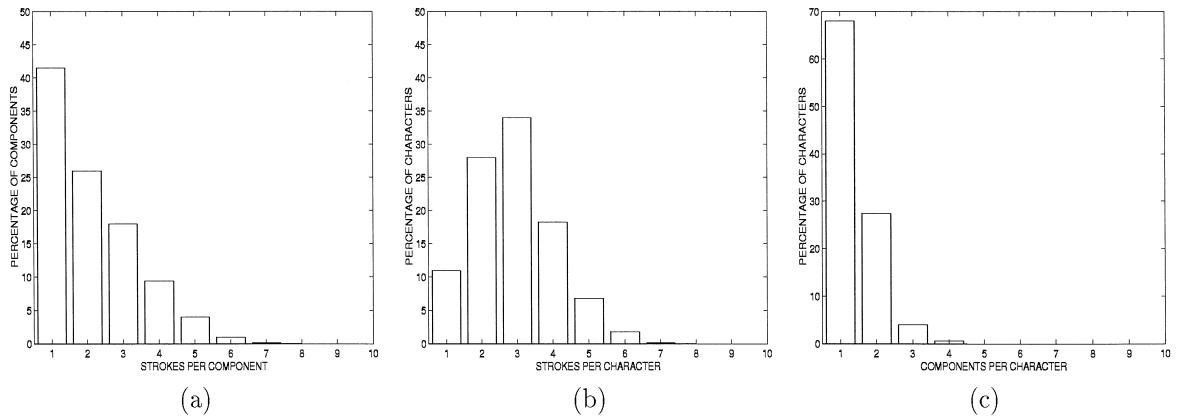


Fig. 1. Segmentation results: Distributions for characters in UNIPEN *train\_r01\_v02*. (a) Strokes per component. (b) Strokes per character. (c) Components per character.

successive changes in trajectory direction, discards spurious minima and too short artifacts in the extremes of a component. Additional enhancements implemented in our system include adaptivity for the angle and length thresholds and mechanisms to avoid proliferation of strokes in curved parts of a component (Gago, 1996).

An objective of the segmentation phase is to keep the number of strokes per component as low as possible in order to reduce the computational cost of the classification stage. This requirement was satisfied by our method, as it can be observed in the results obtained with the characters of UNIPEN second training release *train\_r01\_v02*, shown in Fig. 1. We can also see that the total number of strokes per character is kept less than 8, since there is a compromise between number of components per character and complexity of each component (e.g. contrast the one-component character “m” with the three-component character “H”).

Feature extraction or codification is necessary if we want to group strokes of similar characteristics together and therefore form a base of strokes or character primitives. Segmentation and feature extraction phases should jointly provide feature vectors for every stroke, that would be discriminant, geometrically significant and consistent for each character class. In this work we studied several codifications, containing mainly features of length and phase, but also coefficients of *wavelet* transform (WT) of important functions of the stroke trajectories. Among

other possible figures of merit, discrimination capacity can be studied through Shannon entropy as shown in Table 1 for some of the codifications. Such entropy is obtained for each feature by dividing the interval  $[0,1]$  into  $N$  subintervals and applying Eq. (1),

$$H_{\text{norm}} = - \frac{\sum_{i=1}^N P_i \log_2 P_i}{\log_2 N}, \quad (1)$$

where  $P_i$  are frequencies of appearance of the feature within interval  $i$ , and  $\log_2 N$  is the maximum Shannon entropy.

Another way to test the quality of the segmentation process and the selection of the feature set is through clustering maps. In this method, we employed the unsupervised version of FasArt, described in Section 3, to group together strokes based simply on the similarity of the corresponding feature vec-

Table 1

Normalized Shannon entropy for some codifications. The first two codifications are more discriminant, since each individual feature contains more information.

Features	Shannon entropy
Length and three phase points	0.8975 0.9609 0.8952 0.9474
Four WT coefficients	0.7758 0.9698 0.9528 0.9513
from coordinate signal	0.7902 0.9324 0.8695 0.9080
Eight WT coefficients	0.6295 0.5361 0.4667 0.5077
from velocity signal	0.3718 0.3693 0.3411 0.3192

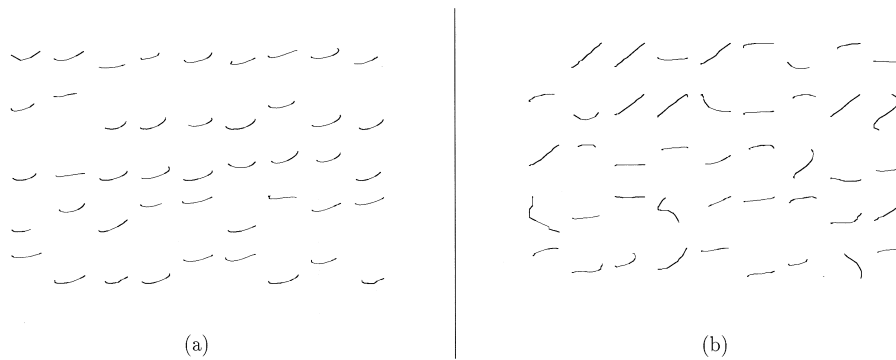


Fig. 2. Clustering maps for (a) length and three phase points, and (b) WT coefficients from coordinates. While the former produces geometrically significant clustering the latter does not, although both are discriminant according to Shannon entropy. Clustering was made with UFasArt, the unsupervised module in FasArt architecture.

tors. In Fig. 2, we can observe that the length and phase representations produce clusters of geometrically similar strokes, a desired property of our system.

The final feature set was based on length and three phase points, including also the end  $y$ -coordinate (normalized) and a discrete sequential characteristic that informed if the stroke started and/or ended component.

### 3. Overview of the proposed neuro-fuzzy system

For clustering and classification we used the neuro-fuzzy architecture FasArt (*Fuzzy Adaptive System ART-based*) (Cano et al., 1996) whose structure follows the principles of ARTMAP architectures (Carpenter et al., 1992). In this system, two FasArt unsupervised modules, one grouping strokes ( $ART_a$ ) and another classifying the label ( $ART_b$ ), are joined by an inter-ART map. A *STORE* module is placed after the  $ART_a$ , in order to record the sequence of clusters into which all strokes of a character are classified. A simplified block-diagram that holds for both FasArt- and Fuzzy ARTMAP-based architectures can be seen in Fig. 3.

FasArt introduces a new activation function which can be interpreted as fuzzy membership function, and therefore it can be considered as a formal fuzzy logic system. Then, we can exploit theoretical results that confirm FasArt capability to work as a universal function approximator (Cano et al., 1996). A new

design parameter allows us specify the degree of fuzziness of the system (Cano et al., 1996). On the other hand, important properties of the neural ART architectures are kept in the new model, like its capacity for incremental learning even in the test phase.

The choice of an ART-based architecture for character recognition is based on ART's strong biological and psychological foundations that make it more appropriate for a cognitive task. Besides the general advantages of neuro-fuzzy systems, like adaptivity or parallelism, ART networks also comply to the stability-plasticity and noise-saturation dilemmas, present in various neural systems (Carpenter et

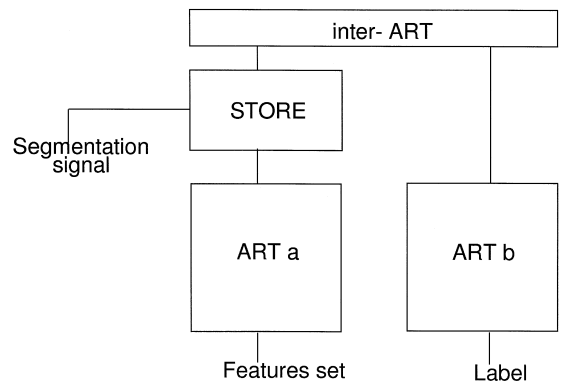


Fig. 3. Block-diagram for an on-line character recognition system based on either Fuzzy ARTMAP or FasArt. Feature vectors for each stroke are presented to  $ART_a$ , and labels to  $ART_b$ . Appropriate segmentation signals to the *STORE* module indicate end-of-component or end-of-character.

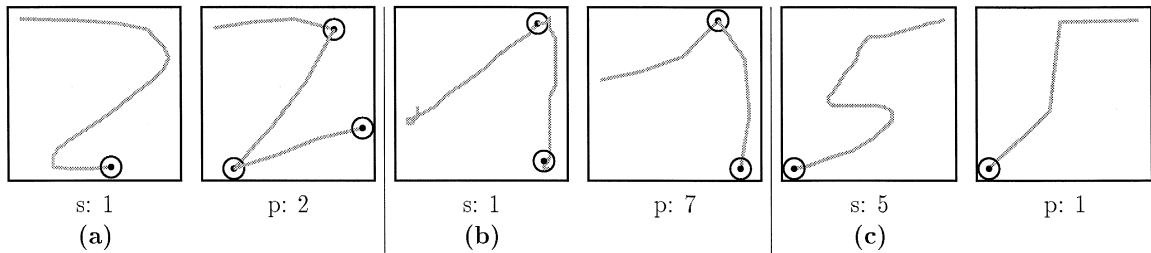


Fig. 4. Typical classification errors, where the source (s) character and the prediction (p) label are shown, with the original character and FasArt internal representation respectively. (a) Badly labeled data. (b) Ambiguous data. (c) Badly segmented data.

al., 1992). Furthermore, FasArt also incorporates the aforementioned advantages of its formulation as a fuzzy logic system. Finally, a previously developed ART architecture (Dimitriadis and López Coronado, 1995) had suggested that such systems were adequate for handwriting recognition and served as a basis of comparison with the new architecture.

Clustering is made on strokes, i.e., clusters represent categories of strokes and characters are then represented by sequences of clusters.

Classification of a pattern is performed according to the degree of activation of fuzzy rules generated during learning phase. Since values of membership functions for a fuzzy set can also be interpreted as classification to a class with a certain confidence degree, other fuzzy sets (classes) can also be considered for postprocessing modules that may correct errors. Additionally, the interpretation of categories as fuzzy sets let us recover the prototypes codified in the network weights, very useful in the analysis of classification performance, as it can be seen in Fig. 4.

Significant design parameters are vigilance parameters,  $\rho_a$ ,  $\rho_b$  and  $\rho_{ab}$  for the unsupervised modules  $ART_a$ ,  $ART_b$  and the inter-ART map respectively. These parameters control granularity of built

clusters, as well as handling of exceptions. Learning rate parameters  $\beta_a$ ,  $\beta_b$  and  $\beta_{ab}$  control how fast new learning adapts previous knowledge. A new statistical module has been added in our system, that rejects a network prediction if the edit distance between the sequence that represents the input character/pattern and the learned prototypes is high enough. Such a distance is based on the number of insertions, deletions and substitutions between two sequences.

All parameters are fixed for the experiments based on the knowledge about the statistical distribution of the data and a validation phase. Although we selected heuristic values, we could have employed a genetic algorithm, such as that proposed by Renders and Flasse (1996) in order to tune them. However, it was experimentally found that parameters could be safely varied in  $[0.5, 1]$  without undergoing a recognition rate of 78%, as long as vigilance in the training phase was equal or higher to that in the test.

#### 4. Analysis of the classification results

Experimental results shown in this paper use all 4020 digits of the second UNIPEN training release,

Table 2

FasArt and Fuzzy ARTMAP main parameters (their values were fixed before the experiments). It was experimentally observed that better results were obtained with lower vigilance parameter and learning rates during the test phase (this change was made externally).

	Train			Test		
	Inter-ART	$ART_a$	$ART_b$	Inter-ART	$ART_a$	$ART_b$
$\rho$ (vigilance)	0.70	0.90	1.00	0.60	0.60	1.00
$\beta$ (learning rate)	1.00	0.80	1.00	–	–	–
$\beta^C$ (only FasArt)	1.00	0.80	1.00	–	–	–
STORE attenuation	0.75	–	–	0.75	–	–

Table 3

Recognition rates with FasArt and Fuzzy ARTMAP after five training cycles for the second UNIPEN training release.

Architecture	Correct	Wrong	Unclassified
FasArt	85.66%	6.55%	7.79%
Fuzzy ARTMAP	83.68%	9.28%	7.04%

distributed as *train\_r01\_v02*, and divided randomly in equivalent training and test sets, assuming that all the authors provide characters to both sets. The main network parameters are shown in Table 2, and five training cycles were employed.

Global recognition rates for both architectures can be seen in Table 3, while the confusion matrix is shown in Table 4. It can be observed that FasArt performance is slightly better than that of Fuzzy ARTMAP. Although relative enhancement is not significant, it can be further improved because class labels together with confidence degrees provided by FasArt fit well to a postprocessing module of error detection and correction (Merino Pastor et al., 1996).

It is important to assess the reasons of errors or no-decisions, that come out of a global recognition system, with various modules. The following main error sources were detected after an analysis of the global confusion matrix, as well as of characteristic special cases, such as those shown in Fig. 4.

- *Erroneously labeled data* (as judged by human recognition), since when this character set was released not all characters had been checked by human recognizers. This is the case of Fig. 4(a), where the correct label should have been “2” instead of “1”, as predicted by our system.

Table 4

Confusion matrix for numeral recognition using FasArt. High confusion exists between “1” and “7”.

	0	1	2	3	4	5	6	7	8	9
0	<b>163</b>	4	1	0	2	0	4	0	7	0
1	0	<b>166</b>	3	4	3	1	0	<b>14</b>	3	1
2	1	3	<b>177</b>	2	0	0	0	0	0	0
3	0	0	5	<b>156</b>	0	4	0	5	0	0
4	1	5	0	0	<b>171</b>	1	3	1	0	5
5	0	2	0	6	0	<b>135</b>	1	0	1	5
6	2	2	0	0	0	1	<b>188</b>	0	1	0
7	0	<b>22</b>	3	0	0	0	0	<b>157</b>	0	2
8	5	0	1	0	0	0	1	0	<b>177</b>	0
9	0	0	0	2	0	3	0	2	1	<b>177</b>

Table 5

Recognition rates for human testers for the same second UNIPEN training release.

	Correct	Wrong	Unclassified
Human	90.21%	3.77%	6.02%

- *Ambiguous data* (even to human recognition), as the characters shown in Fig. 4(b). This is the case of “1” and “7”, a very frequent error confirmed by the confusion matrix.
- *Segmentation errors*, that failed to detect geometrically significant strokes, as shown in Fig. 4(c). For such complex strokes, current codification was not a fair description, since it is mainly based on phases.

A reasonable performance limit of our system can be obtained if we compare its performance with that of humans. We can then estimate indirectly how many errors were due to badly labeled or ambiguous data. The same experiments were also carried out with five volunteers not related to this research project. Isolated digits were presented randomly to the human testers, who had to label as digits from “0” through “9” or “unknown”. Average recognition rates shown in Table 5 indicate a fine performance of our system for this difficult classification task.

As we explained in the previous section, clustering is performed on strokes while characters are represented by sequences of clusters. The number of clusters may serve as a measure of the system complexity, while the number of sequences of clusters for each digit represents the total number of allo-graphs considered by our system. In our experimental studies with FasArt, 372 clusters were created for all 6055 strokes present in the 2016 digits of the training set built from the *train\_r01\_v02* data set. On the other hand, a typical representation through

Table 6

Recognition rates of our system for the most recent UNIPEN *train\_r01\_v07* set.

Architecture	Train	Test	Correct	Wrong	Unclassified
FasArt	7139	7239	82.36%	9.67%	7.97%
Fuzzy ARTMAP	7139	7239	81.59%	11.95%	6.46%

Table 7

Processing time, measured with the Linux utility `time`, for a PC with a Pentium 120MHz processor.

Stage	Average time
Segmentation	6 ms
Codification	5 ms
Training (1 cycle)	158 ms
Test	265 ms
Average (incl. training)	434 ms
Average (not incl. training)	276 ms

allographs takes 16 allographs to represent more than 95% of the samples of digit ‘‘1’’. The number of clusters and allographs is rather high but it can be reduced using other preprocessing techniques, vigilance parameters, or rule fusion techniques.

Using the same experimental conditions and with just a training cycle, we tested our system with the most recent UNIPEN seventh training release distributed as *train\_r01\_v07*. The capacity of our system with other data sets can be assessed with the recognition rates shown in Table 6, slightly worse than those of the second release set, due to the increased variability of the new set.

One of the most important features for the applicability of an on-line handwriting recognizer in a pen-computer is real-time performance. Although our system is not optimized in terms of access to hard drive or monitor, processing time is low enough, as we can observe from the results shown in Table 7. We should remind that handwriting is a low velocity process, especially if we compare it with the speed of 0.28 sec/char achieved by our system.

## 5. Conclusions

A new handwriting recognition system was proposed in this paper. Instead of testing and validating it in a home-made set, we performed extensive experimental work on massive data collected within the UNIPEN project. The main interesting point of UNIPEN data is the fact that they are not controlled, come from various authors, graphic tablets, cultures, etc. and they can serve for comparison among various systems. The achieved recognition rate of almost 86% on the digits of the second training set has been

shown to be close to the rate of 90% achieved by independent human testers. Then several of the errors or no-decisions are due to badly labeled or ambiguous data even for human readers. At the same time its low computing requirements make it valid for a real-time implementation in a pen-computer.

The proposed system considers characters as sequences of strokes, inspired on biological models of handwriting generation, and therefore it is heavily influenced by the technique that segments handwritten components into strokes. The implemented algorithm avoids problems of local processing and enhances existing techniques based on detection of velocity minima. It satisfies the requirements for a reduced number of geometrically significant strokes, that could form a base of elemental movements, capable of describing characters in a consistent way.

The feature vectors for each stroke have been selected using the Shannon entropy criterion, together with an analysis of the produced clustering maps, that confirms that geometrically similar strokes belong to the same cluster. The recognition system was based on an ART-based neuro-fuzzy system, applied for the first time to pattern recognition problems. Its neural nature permits learning from examples, as well as construction of class prototypes, while its formal definition as a fuzzy logic system allows a more efficient postprocessing scheme and extraction of fuzzy rules.

As an evidence of such capacity of rule extraction, our current work focuses on the creation of an allograph lexicon, that would provide us insights about the way characters are built from primitives.

## Acknowledgements

We would like to thank the members of the Document Processing Group (A. Grande, G.I. Sainz, P. Sánchez, O.M. González, M. Sánchez) for their contributions to the preparation of this paper and all related work made within the group.

## References

- Al-Emami, A., Usher, M., 1990. On-line recognition of handwritten arabic characters. *IEEE Trans. Pattern Anal. Machine Intell.* 12 (7), 704–710.

- Bradski, G., Carpenter, G.A., Grossberg, S., Rosen, D.B., 1992. Working memory networks for learning temporal order, with applications to three-dimensional visual object recognition. *Neural Comput.* 4 (2), 270–286.
- Cano, J.M., Dimitriadis, Y.A., Arauzo, M., López Coronado, J., 1996. FasArt: A new neuro-fuzzy architecture for incremental learning in system identification. In: Proc. IFAC World Congress, IFAC96, San Francisco, vol. F, pp. 133–138.
- Carpenter, G.A., Grossberg, S., Markuzon, N., Reynolds, J.H., Rosen, D.B., 1992. Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps. *IEEE Trans. Neural Networks* 3 (5), 698–713.
- Dimitriadis, Y.A., López Coronado, J., 1995. Towards an ART-based mathematical editor, that uses on-line handwritten symbol recognition. *Pattern Recognition* 28 (6), 807–822.
- Gago, J.A., 1996. Development of a segmentation module for a neural system of handwriting recognition. Engineering Degree Thesis, School of Telecommunications Engineering, University of Valladolid, Spain (in Spanish).
- Guyon, I., Schomaker, L., Plamondon, R., Liberman, M., Janet, S., 1994. UNIPEN: Project of on-line data exchange and recognizer benchmarks. In: Proc. 12th International Conference on Pattern Recognition, Jerusalem, Israel, pp. 9–13.
- Kerrick, D.D., Bovik, A.C., 1988. Microprocessor-based recognition of hand-printed characters from a tablet input. *Pattern Recognition* 21 (5), 525–537.
- Morasso, P., Barberi, L., Pagliano, S., Vergano, D., 1993. Recognition experiments of cursive dynamic handwriting with self-organizing networks. *Pattern Recognition* 26 (3), 451–460.
- Merino Pastor, F., Dimitriadis, Y.A., García García, R., López Coronado, J., 1996. A dictionary-based neural network for on-line handwriting recognition. In: Simner, M.L., Leedham, C.G., Thomassen, A.J.W.M. (Eds.), *Handwriting and Drawing Research: Basic and Applied Issues*. IOS Press, Amsterdam, The Netherlands, pp. 343–357.
- Plamondon, R., 1995. A kinematic theory of rapid human movements. Part 1: Movement representation and generation. *Biological Cybernet.* 72, 295–307.
- Parizeau, M., Plamondon, R., 1995. A fuzzy-syntactic approach to an allograph modeling for cursive script recognition. *IEEE Trans. Pattern Anal. Machine Intell.* 17 (7), 702–712.
- Renders, J.M., Flasse, S.L., 1996. Hybrid methods using genetic algorithms for global optimization. *IEEE Trans. Systems Man Cybernet.* 26 (2), 243–258.
- Schomaker, L.R.B., 1993. Using stroke- or character-based self-organizing maps in the recognition of on-line connected cursive script. *Pattern Recognition* 26 (3), 443–450.
- Tappert, C.C., Suen, C.Y., Wakahara, T., 1990. The state of the art in on-line handwriting recognition. *IEEE Trans. Pattern Anal. Machine Intell.* 12 (8), 787–808.