Allograph Adjacency Constraints for Cursive Script Recognition*

Marc PARIZEAU† and Réjean PLAMONDON‡

printed July 8, 1997

Abstract

This paper defines horizontal and vertical adjacency constraints between allographs. The constraints are used to build an allograph segmentation graph containing paths that represent character strings possibly corresponding to an unknown cursive word. Allographs are modelled by fuzzy-shape grammars and are segmented within this cursive word by an adequate parser. Results are given for multi-writer random letter sequences. For a test database containing cursive handwriting samples from 10 different writers, and without using any linguistic knowledge, average character recognition rates of 84.4% to 91.6% are obtained depending on whether only the first string outputted by the system is considered, or if the best of the top ten is accepted, the recognition system being non personalized.

1 Introduction

One of the main difficulty associated with cursive script recognition is the problem of letter segmentation [1]. Because there is no way, a priori, of determining unambiguously where each letter starts and ends, the recognition processes described in the literature usually incorporate some segmentation scheme that ultimately leads to several concurrent letter hypotheses. Adjacency constraints can be used to analyse spatial relationships between these hypotheses and generate coherent strings [2].

The adjacency constraints proposed in this paper have been used in a fuzzy-syntactic approach to allograph modelling for cursive script recognition [2, 3]. The term allograph is used here to designate an *intrinsic model* for a given letter. For example, each letter of the roman alphabet has at least one lower and one upper case model. For lower case letters, there are usually at least two allographs : the conventional cursive letter which is thought in schools[1], and the letter that resembles the printed character which is sometimes used to write more clearly.

In the above approach, allographs are modelled with fuzzy-shape grammars that define their morphological characteristics [5]. A parser is used to identify (segment) within a set of pattern primitives, those subsets that respect the fuzzy-constraints of allograph models. The object of this paper is to present the next phase of the recognition approach, where segmented allographs, called *allograph traces*, are analysed according to horizontal and vertical adjacency constraints.

The next section describes the general methodology followed for allograph building and defines a generic representation for allograph traces. Then, Section 3 proceeds with the definition of the adjacency constraints used for the construction of an allograph segmentation graph. The paths of this graph represent character strings possibly corresponding to the unknown cursive word. Experimental results for multi-writer cursive random letter sequences are presented in Section 4.

2 Allograph Modelling and Segmentation

Allograph models are constructed in a hierarchical manner, from handwriting elements that stem from

*This work was published in *Proc. of the Third International Workshop on Frontiers in Handwriting Recognition*, Buffalo, May 25-27, pp. 252-261, 1997. It was supported in part by NSERC Canada under grant OGP0915, and in part by FCAR Québec under grant CRP2667. M. Parizeau received scholarships from NSERC and FCAR.

†M. Parizeau was with the département de Génie Électrique et de Génie Informatique, École Polytechnique de Montréal. He is now with the département de Génie Électrique, Université Laval, St-Foy (PQ), Canada, G1K 7P4.

‡R. Plamondon is with the département de Génie Électrique et de Génie Informatique, École Polytechnique de Montréal, C.P. 6079 – Succ. "A", Montréal, Canada, H3C 3A7.

[1]Of course, these allographs are more or less conventional in the sense that they can vary between schoolboards or regions [4].

a pragmatic representation of handwriting. This representation is an operational model that incorporates some of the features of a general handwriting segmentation theory [6], while making some pragmatic simplifications [7]. According to this representation, handwriting components[2] are modelled as a sequence of *characteristic points* linked together by *constant curvature* segments (i.e. circular arcs). A handwriting *element* is associated with every characteristic point and represents a portion of the component that stretches from the previous characteristic point to the following one.

An allograph model, or simply an *allograph*, is the expression of an ideal concept. In general, it is represented by a set of *symbols* defined with fuzzy-shape grammars. A symbol is constructed from other (simpler) symbols or *elements*. An allograph is thus the root of a hierarchy of symbols where leaves are elements.

A *trace* is an instance of an allograph, as produced by a writer at a specific time. In general, it is composed of a set of *curves* that are segmented by a parser for fuzzy-shape grammars. A curve is an assembly of other (simpler) curves or *primitives*. A trace is thus the root of a hierarchy of curves where leaves are primitives.

With shape grammars, symbols are defined with a set of syntactic and semantic attributes used respectively as *attachment points* for spatial arrangement of the symbols, and as *properties* for testing the coherence of these arrangements. The grammars are said to be "fuzzy" because the applicability conditions of their production rules are expressed with *fuzzy-logic*. For more details about the grammatical formalism used or about the parsing process that enables the segmentation of traces, the reader is referred to [5].

Four allograph classes can be defined for lower case letters, depending on whether or not they possess ascenders and descenders. Figure 1 illustrates the generic representation adopted for these classes. The first class of allographs possess only a main body but no ascender nor descender (i.e. letters *a, c, e*, etc...). The second class, contains only those letters that possess a main body and an ascender but no descender (i.e. letters *b, d, h*, etc...). For the third class, they possess a main body and a descender but no ascender (i.e. letters *g, j, p*, etc...). And for class 4, they

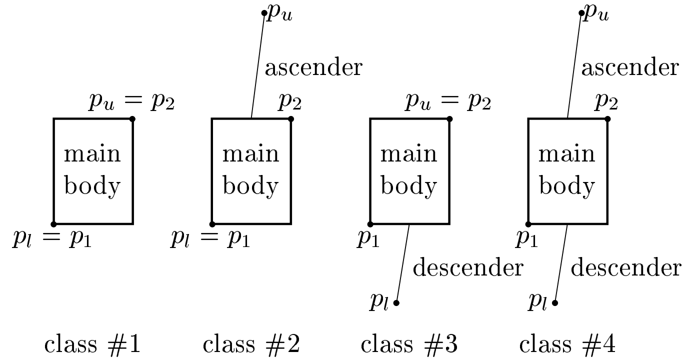

Figure 1: *Generic spatial representation of allograph traces.*

possess a main body and both an ascender and a descender (i.e. letter *f* only).

Allographs are designed in such a way that each trace is segmented with 4 attachment points $p_1(x_1, y_1)$, $p_2(x_2, y_2)$, $p_l(x_l, y_l)$ and $p_u(x_u, y_u)$ corresponding respectively to the lower left and upper right corners of the bounding box enclosing its main body, and to its lower and upper points. When the allograph doesn't possess an ascender (resp. a descender), point p_u (resp. p_l) then has the value $p_u = p_2$ (resp. $p_l = p_1$). The coordinates of these points are used to specify the adjacency constraints between pairs of traces.

3 Adjacency Constraints

Two traces are considered adjacent if, and only if, they respect both a *horizontal constraint* that measures the coherence of their horizontal spacing, and a *vertical constraint* that measures the quality of their vertical alignment. These constraints are defined and evaluated using fuzzy-logic [9, 10].

3.1 Horizontal Constraint

Let p_1^μ, p_2^μ, p_l^μ, p_u^μ denote the attachment points of trace μ, and p_1^ν, p_2^ν, p_l^ν, p_u^ν denote the attachment points of trace ν. Then, the horizontal constraint χ_h between μ and ν is expressed by the following equation :

$$\chi_h(\mu,\nu)=\begin{cases} f_h\left(\frac{x_1^\nu - x_2^\mu}{\min\left[x_2^\mu - x_1^\mu,\, x_2^\nu - x_1^\nu\right]}\right) & \text{if } x_1^\nu - x_2^\mu < 0 \\ f_h\left(\frac{x_1^\nu - x_2^\mu}{\max\left[y_2^\mu - y_1^\mu,\, y_2^\nu - y_1^\nu\right]}\right) & \text{otherwise} \end{cases} \qquad (1)$$

where f_h represents a membership function that defines a fuzzy-set associated with trace pairs that are

[2] A handwriting component is defined as a portion of the written trace between a pendown and a penlift [8] (while the pen is in contact with paper).

horizontally adjacent. In other words, function f_h puts a fuzzy-threshold on the horizontal spacing between the two traces. This threshold is applied on two different ratios depending on whether the main bodies of the two traces overlap or not. When they overlap ($x_1^\nu - x_2^\mu < 0$), the length of this overlapping, relative to the width of the thinnest main body, is used. Otherwise ($x_1^\nu - x_2^\mu \geq 0$), the horizontal distance between the two main bodies, relative to the highest one, is considered.

3.2 Vertical Constraint

The vertical constraint χ_v between μ and ν is expressed in a similar fashion except that it is divided in two parts :

$$\chi_v(\mu,\nu) = \min \left[\begin{array}{l} f_{vu}^{mn}\left(\frac{y_u^\mu - y_u^\nu}{\max[y_u^\mu - y_1^\mu, y_u^\nu - y_1^\nu]}\right) \\ f_{vl}^{mn}\left(\frac{y_l^\mu - y_l^\nu}{\max[y_2^\mu - y_l^\mu, y_2^\nu - y_l^\nu]}\right) \end{array} \right] \quad (2)$$

where m and n correspond respectively to the class of allograph associated with traces μ and ν. The first part is concerned with the alignment of the upper zones of the two traces, while the second part deals with the alignment of their lower zones. Again, f_{vu}^{mn} and f_{vl}^{mn} represent membership functions of fuzzy-sets associated respectively to trace pairs that have vertically adjacent upper and lower zones. Three different functions are used depending to which class, traces μ and ν belong. These three functions correspond to three types of alignment : when the vertical heights of two main bodies are compared ($y_u = y_2$ or $y_l = y_1$), when the vertical heights of two upper or two lower zones are compared ($y_u > y_2$ or $y_l < y_1$) and when the vertical height of a main body is compared with the vertical height of an upper or lower zone (otherwise). The corresponding thresholds are thus applied on the height difference of the upper or lower zones of the two traces, relative to the highest one. The minimum value is retained as the final vertical adjacency membership degree.

3.3 Segmentation Graph

The adjacency relation between pairs of traces can now be used to create an allograph segmentation graph. It is expressed by the following expression :

$$\min[\chi_h(\mu,\nu), \chi_v(\mu,\nu)] > 0 \Longleftrightarrow \text{adjacent}(\mu,\nu) \quad (3)$$

where χ_h et χ_v are defined respectively by equations 1 and 2.

The segmentation graph is constituted by a set of nodes N and a set of arcs A. Each node $n \in N$ is defined by a couple $n = (\mu, \alpha_\mu)$ where μ represents a trace and α_μ its degree of membership[3]. The degree of membership α_μ is a measure of how this particular trace fits the allograph model. Let Ω be the set of traces segmented in a particular cursive word, then N is defined by the following equation :

$$N = \{(\mu, \alpha_\mu) \mid \mu \in \Omega, \alpha_\mu > 0\} \quad (4)$$

Each arc $a_{ij} \in A$ is defined by a triplet $a_{ij} = (n_i, n_j, \omega_{ij})$ where $(n_i, n_j) \in N^2$, $i \neq j$ are two nodes of the graph and ω_{ij} is the weight of the arc, which measures the adjacency between the nodes :

$$\omega_{ij} = \min[\chi_h(\mu,\nu), \chi_v(\mu,\nu)] \quad (5)$$

with $n_i = (\mu, \alpha_\mu)$ and $n_j = (\nu, \alpha_\nu)$. Then, A is defined by the following equation :

$$A = \{(n_i, n_j, \omega_{ij}) \mid (n_i, n_j) \in N^2, i \neq j, w_{ij} > 0\} \quad (6)$$

Figure 2 gives an example of a segmentation graph for cursive word *sosie* (french for a person's *double*). The nodes of N are sorted according to increassing x_1 coordinates of their attachment point p_1 (i.e. lower left corner of main body's bounding box) and are shown along a circle. The left most trace is placed at angle 0 on this circle. The other nodes are distributed uniformly counter clockwise. The degree of membership[4] associated to each trace is printed in parentheses. The superscript number is used to identify different traces of a same letter. The arcs of the graph are represented by line segments linking nodes. Their weights ω_{ij} are printed at the center of the corresponding segment.

Trace s^1 (look at angle 0 on the circle) corresponds to the first letter s of word *sosie*. The bounding box around its main body is shown in the upper part of Figure 2. The next node is trace c^1 which corresponds to the left part of letter o in word *sosie*. Next node is trace o^1 which corresponds to the letter o of word *sosie*. Its bounding box is also shown in the upper part of Figure 2. Next node is trace s^2 which is a somewhat weak hypothesis of the second letter s. It is composed of the right part of letter o combined with

[3]Allographs are modelled by fuzzy-grammars and traces are thus segmented with a degree of membership to a fuzzy-set associated with the corresponding allograph [5].

[4]In fuzzy-logic, degrees of membership are usually specified in the range $[0,1]$. For practical reasons, we have used the range $[0,100]$.

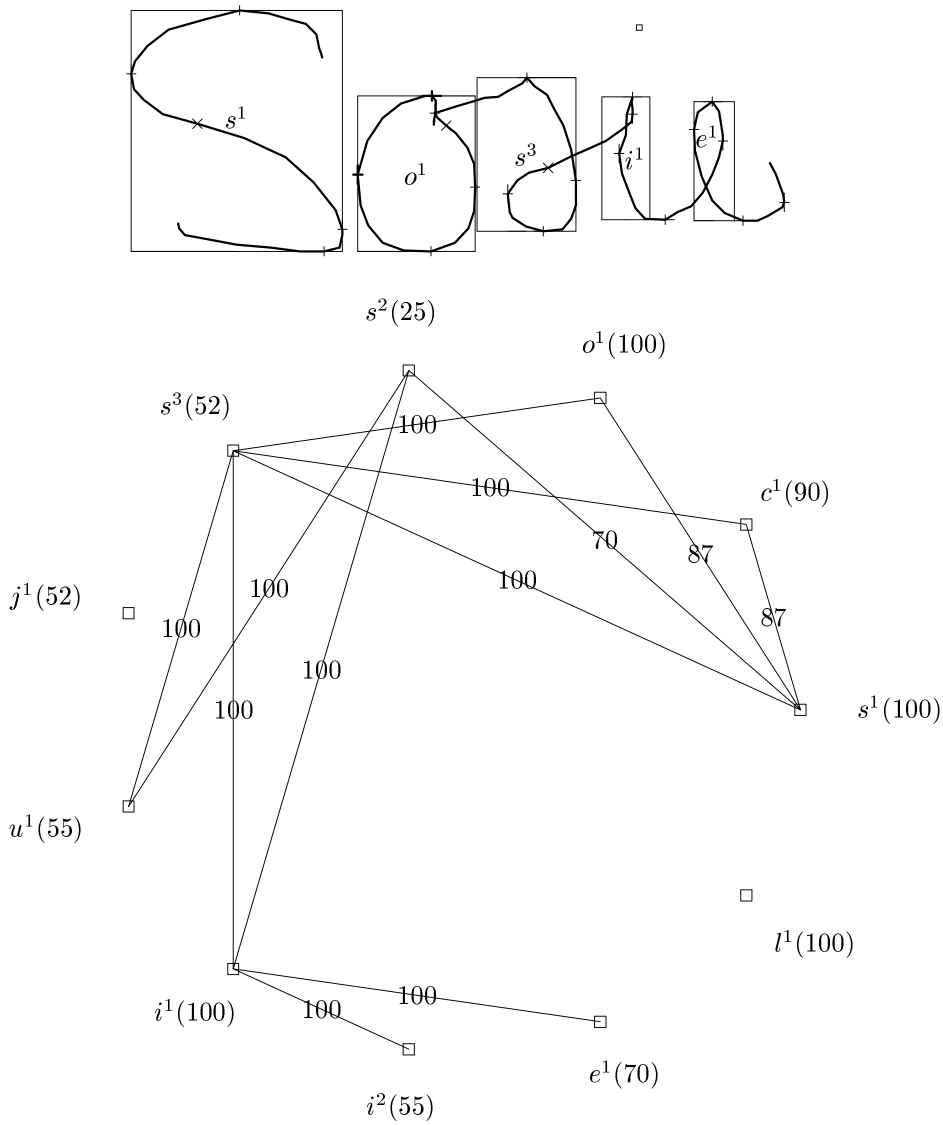

Figure 2: *Allograph segmentation graph example.*

the letter *s*. It is a weak hypothesis in the sense that its membership degree is only 25 out of 100. Next node is trace s^3 which is the correct segmentation of the second letter *s*. Its bounding box is also shown in the figure. Next node is trace j^1 which is a different interpretation of the previous trace. The reader should notice that because of vertical inconsistency, this trace is not adjacent to any other trace. Next node is trace u^1 which is a combination of letters *i* and *e* of word *sosie*. Next node is trace i^1 which corresponds to letter *i* of the same word. Again, the bounding box around its main body is shown in the upper part of Figure 2. Next node is trace i^2 which is an incorrect, although not in any way absurd, interpretation of letter *e* of the word. Next node is trace e^1 which corresponds to letter *e* of word *sosie*. As before, its bounding box is shown in the upper part of the figure. Finally, the last node is trace l^1 which is another plausible, although vertically inconsistent, interpretation of letter *e*. But like for the case of trace j^1, the reader should notice that it isn't adjacent to any other trace.

3.4 Graph Analysis

A path in the segmentation graph from the node associated with trace μ to the node associated with trace ν, is defined by a sequence of k nodes $n_1, n_2, \ldots, n_k$ such that $n_1 = (\mu, \alpha_\mu)$, $n_k = (\nu, \alpha_\nu)$ and that n_i and n_{i+1} are adjacent nodes $\forall\, i \in \{1, \ldots, k-1\}$.

A character string can be associated with every path in the graph. These strings then correspond to different, possibly incomplete, interpretations of the unknown cursive word. Three criteria are used to sort these interpretations. The first concerns the quality of segmented traces in the path. The higher their degree of membership, the more likely the interpretation. The second deals with adjacency between traces. Again, the higher their adjacency relations are, the more likely becomes the interpretation. Finally, the third criterion takes into account the proportion of primitives associated with the traces in the path, relative to their total number in the cursive word. Obviously, the greater that proportion is, the more representative the resulting path.

These criteria are combined using the following ranking coefficient Γ :

$$\Gamma = \frac{100}{\eta} \sum_{i=1}^{k} \mathrm{card}(n_i) \left[\frac{1}{k} \sum_{i=1}^{k} \alpha_i + \frac{1}{k-1} \sum_{i=1}^{k-1} \omega_{i,i+1} \right] \quad (7)$$

where n_i is the i^{th} node in the path, α_i is the degree of membership for the trace associated with n_i, $\omega_{i,i+1}$ is the weight associated with arc $a_{i,i+1}$, η is the total number of primitives in the unknown cursive word, and $\mathrm{card}(n_i)$ represents the number of primitives contained in the trace associated with node n_i.

The paths of the graph are explored using a dynamic programming algorithm that maximizes Γ. For the graph of Figure 2, the first six character strings are listed below with the values obtained by the three criteria in parentheses :

1: sosie (81,84,97)	2: sosii (81,81,97)
3: sosu (81,77,96)	4: scsie (69,82,97)
5: scsii (69,79,97)	6: scsu (69,74,96)

The first number is the term $\frac{100}{\eta} \sum_{i=1}^{k} \mathrm{card}(n_i)$, the second corresponds to the term $\frac{1}{k} \sum_{i=1}^{k} \alpha_i$ and the third to the term $\frac{1}{k-1} \sum_{i=1}^{k-1} \omega_{i,i+1}$.

For this example, the right choice is ranked in first place. This is not always the case but one should not forget the aim of this method : to have the right choice ranked, for example, in the top ten, knowing that linguistic knowledge will be needed to disambiguate otherwise morphologically coherent alternatives.

4 Experimental Results

To evaluate the performance of the adjacency contraints presented in this paper, a special test database of cursive random letter sequences was constructed.

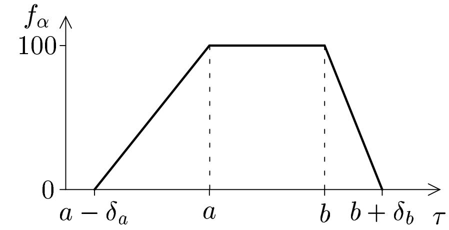

Figure 3: *Membership function used for fixing fuzzy-thresholds.*

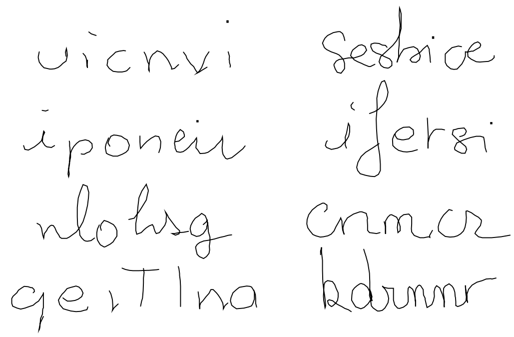

Figure 4: *Samples of cursive random letter sequences found in the test database.*

This database was not used in any way for the developpment of allograph models, nor for fixing the various parameters associated with each adjacency constraint. The membership functions used for fixing the corresponding fuzzy-thresholds have the general shape of a trapezoid like the illustration of Figure 3. Parameters a, b, δ_a and δ_b for this membership function were ajusted for each constraint using a separate training dataset.

The test database contains 100 cursive random letter sequences, each written by 10 different writers. Writers were asked to perform random letter sequences instead of dictionary words because it was planned to compare machine and human reading performance on a common dataset and at the morphological level only (i.e. without linguistic knowledge) [3]. Hence, 100 character strings were generated randomly, each containing from 5 to 7 lower case characters, while respecting the letter frequencies of a 32 000 words french dictionary. Figure 4 gives a few samples in this database. Writers were instructed to used their natural, although clean, handwriting.

Handwriting data were digitized using a Pencept tablet, model *Penpad 300*, with a sampling frequency of 100Hz, a resolution of $0.001''$ and a precision (as specified by the manufacturer) of $0.005''$.

Table 1: *Recognition rates for cursive random letter sequences (test database), in %.*

accepted ranks	writer #1	#2	#3	#4	#5	#6	#7	#8	#9	#10	average rate
1	91.0	91.7	84.5	78.5	92.2	68.2	82.7	75.5	89.6	89.9	84.4
1-2	93.1	95.8	88.9	81.9	94.8	73.4	85.5	80.1	94.1	93.8	88.1
1-3	93.1	96.7	90.5	83.4	95.8	75.2	85.8	83.0	95.3	94.6	89.3
1-5	93.5	97.7	93.1	85.3	95.9	77.5	86.9	85.5	96.1	95.4	90.7
1-10	93.5	98.2	93.8	87.1	96.2	78.6	87.3	88.4	96.7	96.1	91.6

Table 1 gives recognition results for the test database. They are obtained by computing error rates using the Wagner-Fischer algorithm [11]. The first line of Table 1, gives recognition rates when considering only the character string ranked first by the system. The other lines of this table enumerate results when considering respectively the best of the first 2, 3, 5 and 10 strings output by the system.

The average recognition rate thus varies from 84.4% to 91.6% depending on whether only the string ranked first is considered, or if the best of the top 10 strings is accepted. These performances are juged very good considering the fact that the recognition system is truly multi-writer, in the sense that system parameters are the same for all writers, and that no linguistic knowledge is used for recognition. Furthermore, writers #6, #8 and #10 were completly unknowned by the system (no participation in the training datasets). This might explain in part their somewhat lower performance (especially for writer #6 who uses several allographs that weren't modelled).

5 Conclusion

This paper has presented adjacency constraints that can be used to analyse the spatial relationships between concurrent letter hypotheses. These constraints can be applied as long as the segmentation/recognition method can isolate the bounding box around the main body of the different hypotheses, and their upper and lower points. They are then used to create a segmentation graph from which paths correspond to possible interpretations of the unknown cursive word.

The mean character recognition rates (from 84.4% to 91.6%) demonstrate the feasability of recognizing natural cursive script with adjacency constraints. Moreover, these results are obtained without the use of any linguistic knowledge. By integrating, for example, information on acceptable *n-grammes* in the construction of the segmentation graph, higher recognition rates should be achieved.

The idea of using adjacency constraints for analysing the relationships between different letter segmentation hypotheses is attractive for bypassing the conventionnal ill-posed problem of zone estimation for cursive handwriting. Indeed, because the frontiers between the three handwriting zones (upper, middle and lower zones) often fluctuate, there is no global solution to this problem. By using adjacency constraints, combinations of different hypotheses can be accepted or rejected with a local criteria.

References

[1] **Tappert** C.C., **Suen** C.Y., **Wakahara** T., "The State of the Art in On-Line Handwriting Recognition", *IEEE trans. on Pattern Analysis and Machine Intelligence*, vol. PAMI-12, no. 8, p. 787-808, August 1990.

[2] **Parizeau** M., "Reconnaissance d'écriture cursive par grammaires floues avec attributs : étape vers la conception d'un bloc-notes électronique", Ph.D. Thesis, Ecole Polytechnique de Montréal, 1992.

[3] **Parizeau** M., **Plamondon** R., "A Fuzzy-Syntactic Approach to Allograph Modelling for Cursive Script Recognition", *in preparation.*

[4] **Suen** C.Y., "Handwriting Education — A Bibliography of Contemporary Publications", *Visible Language*, vol. 9, p. 145-158, 1975.

[5] **Parizeau** M., **Plamondon** R., **Lorette** G., "Fuzzy-Shape Grammars for Cursive Script Recognition", *Proc. of the IAPR International Workshop on Structural and Syntactic Pattern Recognition*, Bern, august 26-28, 1992, in Press.

[6] **Plamondon** R., "A Model-Based Segmentation Framework for Computer Processing of Handwriting", *Proc. of the* 11$^{\text{th}}$ *International Conference on Pattern Recognition*, The Hague, August 30 to September 3, pp. 303-307, 1992.

[7] **Parizeau** M., **Plamondon** R., "A Handwriting Model for Syntactic Recognition of Cursive Script",

Proc. of the 11th *International Conference on Pattern Recognition*, The Hague, August 31 to September 3, vol. II, pp. 308-312, 1992.

[8] **Plamondon** R., **Maarse** F.J., "An Evaluation of Motor Models of Handwriting", *IEEE trans. on Systems, Man and Cybernetics*, vol. SMC-19, no. 5, pp. 1060-1072, 1989.

[9] **Zadeh** L.A., **Fu** K.S., Tanaka K., Shimura M., (editors) : *Fuzzy Sets and their Applications to Cognitive and Decision Processes*, Academic Press, 1975.

[10] **Dubois** D., **Prade** H., *Théorie des possibilités — Applications à la représentation des connaissances en informatique*, Masson, Paris, 1985.

[11] **Wagner** R.A., **Fischer** M.J., "The String to String Correction Problem", *Journal of the ACM*, vol. 21, no. 1, pp. 168-173, 1974.