A Handwriting Model for Syntactic Recognition of Cursive Script*

Marc PARIZEAU† and Réjean PLAMONDON‡

printed July 8, 1997

Abstract

This paper defines an operational handwriting model for on-line syntactic recognition of cursive script. To represent handwriting, it uses characteristic points linked together by segments of uniform curvature. The model acts as a guide for the extraction of attributed primitives, themselves used in shape grammars that model allographs and their adjacency rules. The handwriting model is evaluated by human readers in a comparative analysis of original cursive letter sequences versus their reconstructed traces. Also, its performance is measured in terms of mean reconstruction error and data compression rate.

I Introduction

Syntactic pattern analysis techniques have been used extensively to solve planar shape recognition problems [1]. With these techniques, complex patterns are assembled from simpler subpatterns which are themselves assemblies of even simpler subpatterns (etc...) up to the simplest of them called pattern primitives or simply primitives. One of the greatest drawback of this approach is the usual a priori hypothesis that primitives can be segmented without error.

For cursive script recognition, most approaches define several high level pattern primitives like *strokes*, *curves*, *loops* and *cups* (see for example [2, 3]). However, the problem of correct segmentation and interpretation of these kind of primitives is not trivial and is often an important source of recognition error. Recently, a model-based framework has been proposed to analyse and compare various segmentation schemes [4]. It is in the context of this framework, that a new attributed language has been developped for cursive script recognition [5].

Attributed languages have been proposed as a way to unify syntactic and statistical pattern recognition [6]. With attributed (shape) grammars, it is possible to define patterns with attributes that both describe syntactic and semantic information about the pattern. The advantage of these grammars over conventional string grammars is that relations between primitives are no longer restricted to concatenation. Any morphologically significant relation can be used to define patterns.

The object of this paper is to introduce the model based $\mathcal{A}$ttributed $\mathcal{H}$andwriting $\mathcal{P}$rimitive ($\mathcal{AHP}$) used in a fuzzy-syntactic allograph modeling approach for cursive script recognition [7]. With the $\mathcal{AHP}$, grammars can be defined to model any of the conventional handwriting primitives without making any definitive commitment to a particular segmentation scheme. The paper is organized as follows. The handwriting model is first defined in section II. Then, in section III, the $\mathcal{A}$ttributed $\mathcal{H}$andwriting $\mathcal{P}$rimitive stems from this model. Section IV proceeds with the evaluation of the handwriting model from three different criteria : average reconstruction error, data compression rate and human recognition error. Finally, section V concludes the paper.

II Handwriting Model

Many handwriting models have been proposed for analyzing or generating pieces of handwriting [8]. However, most of them are concerned mainly with temporal simulation of handwriting not directly with recognition (although some of them could be used for this purpose). The model that we propose is an opera-

*This work was published in *Proc. of the 11th International Conference on Pattern Recognition*, The Hague, August 31 to September 3, vol. II, pp. 308-312, 1992. It was supported in part by NSERC Canada under grant OGP000915, and in part by FCAR Québec under grant CRP2667. Marc Parizeau received scholarships from NSERC and FCAR.

†Marc Parizeau was with the Department of Electrical and Computer Engineering, École Polytechnique de Montréal. He is now with the department of Electrical Engineering, Université Laval, St-Foy (PQ), Canada, G1K 7P4. E-mail: *parizeau@gel.ulaval.ca*

‡Réjean Plamondon is with the Department of Electrical and Computer Engineering, École Polytechnique de Montréal, C.P. 6079 – Succ. "A", Montréal, Canada, H3C 3A7.

tional model in the sense that, while incorporating some of the features of a general handwriting model [4, 9, 10, 11], it is aimed directly at the recognition problem thus making some pragmatic simplifications.

A recognition model is a model for which all useful information for recognition purposes is preserved. We call this information *morphologically pertinent* because it is an information that is characteristic of the shape of the symbols that must be recognized.

Model Definition Let $\mathcal{T}$ denote the handwritten trace of a cursive word composed of l components[1] :

$$\mathcal{T} = \{C_1(t), \ldots, C_i(t), \ldots, C_l(t)\} \tag{1}$$

where $C_i(t) = (x_i(t), y_i(t))$ represents the trace of the i^{th} component of the word and $t \in I\!N^+$ represents the time expressed in multiples of the sampling period. A component is made up of a pair of signals representing respectively the pentip position along axes X and Y of the writing plane (i.e. axes of the digitizing tablet).

The proposed model represents a cursive component as a sequence of characteristic points linked together by segments of constant curvature. Characteristic points are morphologically pertinent points of the handwritten trace. The underlying hypothesis is that changes in curvature are only pertinent for recognition when they coincide with characteristic points. Otherwise they are considered an artefact of either the handwriting process itself or the data acquisition process.

The chosen characteristic points are local extrema of signals $x_i(t)$ and $y_i(t)$, and local inflexion points of $C_i(t)$. This purely static choice is motivated by the recognition objectives described previously. Two other hypotheses are also made with this model : first that the baseline of the word is approximatly oriented along the X axis of the writing plane and, second, that the inclination of ascenders and descenders in letters are approximatly oriented along the Y axis of the writing plane. If these hypotheses are not met, it is assumed that preprocessing techniques [12, 13, 14] can detect and correct these orientations.

Under these hypotheses, the chosen characteristic points correspond to a pragmatic approximation of the segmentation scheme proposed by Plamondon [4, 9] where *components* are made up of *strings*, that is, portions of components between two angular discontinuities ; each string is made up of a combination of curvilinear and angular *strokes*, that is, curvilinear or angular displacements resulting from inpulse functions applied as inputs to the proper generator ; and strokes are characterized by log-normal curvilinear and angular velocity profiles. The strings will always be delimited by local extrema of either $x_i(t)$ or $y_i(t)$. The other local extrema and local inflexion points are rough estimates of stroke boundaries which, in fact, are hidden in the signal due to a superimposition phenomena [4, 9].

Local extrema Let $E(f) = \{e_1, \ldots, e_i, \ldots, e_m\}$ be the set of local extrema indexes for signal $f(t)$ before filtering, with m being the number of extrema indexes, $e_i \in \{1, \ldots, n-2\}$ and n corresponding to the number of samples in the signal : $t \in \{0, \ldots, n-1\}$. Then, $E(f)$ is defined by :

$$E(f) = \{e_i \in \{1,\ldots,n-2\} |\ d^{e_i}_{e_i-1}(f) \times d^{e_i}_{e_i+1}(f) \geq 0\} \tag{2}$$

where $d^i_j(f) = f(i) - f(j)$ denotes the difference between sample i and sample j of signal $f(t)$. Now assume that elements of $E(f)$ are sorted in increasing order, that is, $e_i > e_{i-1}\ \forall i \in \{2, \ldots, m\}$. Then, the interesting extrema are those that are sufficiently important in the sense that they respect the following two conditions :

$$|d^{e_i}_{e_{i-1}}(f)| > \max(\delta_f, \tau\, |d^{e_i}_{e_{i-1}}(g)|) \tag{3}$$

$$|d^{e_i}_{e_{i+1}}(f)| > \max(\delta_f, \tau\, |d^{e_i}_{e_{i+1}}(g)|) \tag{4}$$

where δ_f is an amplitude constraint, τ is an angular constraint, $|d^{e_i}_{e_j}(f)|$ represents the absolute value of $d^{e_i}_{e_j}(f)$ and g corresponds to the signal orthogonal to signal f : if $f(t) = x(t)$ then $g(t) = y(t)$, otherwise $g(t) = x(t)$.

The amplitude constraint is easy to visualize : a valid extremum of $f(t)$ must possess a relative amplitude with respect to the previous and next extrema greater than a threshold δ_f. In practice, δ_f is fixed proportional to the precision of the digitizing tablet to filter quantization noise. The angular constraint is introduced to filter noise in the handwriting process. Indeed, for rectilinear trajectories, imprecisions in motor control of the hand tend to produce windings all the more important that the displacement is great. When this displacement is along the X or Y axes of the writing plane, undesirable local extrema are thus produced. The angular constraint $\phi > \tan^{-1}(\tau)$ places a lower bound on the angle ϕ between two successive

[1] 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).

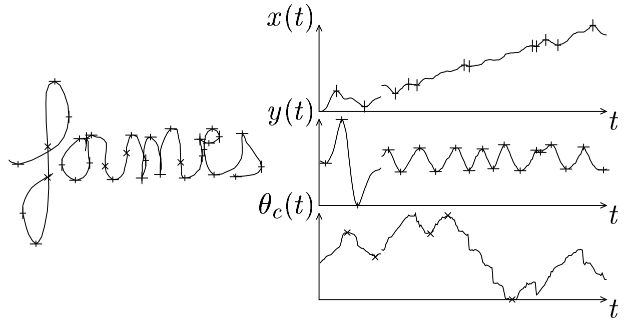

Figure 1: *Example of a cursive word with its characteristic points and the three signals $x(t)$, $y(t)$ and $\theta_c(t)$.*

extrema. Every extremum that does not respect these two conditions is eliminated from $E(f)$.

Local inflexion points Inflexion points of component $C(t)$ are located using the angle of the tangent. An inflexion point corresponds to a change in the sign of the curvature and, thus, to a local extremum in the cumulative angle signal $\theta_c(t)$ of the tangent.

Local extrema of signal $\theta_c(t)$ are obtained and filtered in the same manner as for signal $x(t)$ and $y(t)$ except that the amplitude constraint δ_θ is expressed in angle units and the angular constraint is null : $\tau = 0$. Furthermore, $f(t) = g(t) = \theta_c(t)$.

Figure 1 gives an example for the cursive word "formes" (french for "patterns"). Local extrema of position signals are marked with + and local inflexion points with $\times$. The amplitude constraints are $\delta_x = \delta_y = 0.25$ mm (i.e. twice the specified precision of the digitizing tablet), $\delta_\theta = 30°$ and the angular constraint is $\tau = \frac{1}{10}$. On the right side of the word, the three signals $x(t)$, $y(t)$ and $\theta_c(t)$ are displayed. Notice that several inflexion points were filtered because they coincided with extrema of $x(t)$ or $y(t)$.

III Handwriting Primitive

The $\mathcal{A}$ttributed $\mathcal{H}$andwriting $\mathcal{P}$rimitive ($\mathcal{AHP}$) is defined by the portion of handwriting around a characteristic point. Then, to each characteristic point corresponds a primitive to which a set of attributes is assigned. These attribute constitute the morphologically pertinent information of the corresponding handwriting segment.

Two types of attributes are considered : attachment points used for syntactic arrangement of the primitives ; and properties for their semantic coherence.

Attachment points Three attachment points are used : a *starting point*, a *characteristic point* and an *ending point*. Because only characteristic points of the handwriting model are considered morphologically pertinent, the starting and ending points are respectively chosen as the previous and next characteristic points.

This choice of attachment points imply that two successive primitives of the same component will always have two common attachment points and a common segment of the cursive trace. This is consistent with the proposition that all non characteristic points are only a consequence of the relation that unite together characteristic points. It is also consistent with the handwriting model of Plamondon where strokes usually overlap [4, 9, 11].

Properties Properties associated to primitives stem directly from the handwriting model. A primitive is represented in the $x{\cdot}y$ plane as two circular arcs linking three attachment points. In the $t \cdot \theta_c$ plane it corresponds to two line segments since the curvature of a circle is constant.

Let t_s, t_c and t_e be the time indexes for samples of the component corresponding respectively to the starting, characteristic and ending points of a primitive. Then, let $\omega_s = a_s t + b_s$ and $\omega_e = a_e t + b_e$ be the two regression lines corresponding to the line segments preceding and following the primitive's characteristic point in the $t \cdot \theta_c$ plane, with a_s and a_e being the slopes of the lines, and b_s and b_e the origin's ordinates.

From these two lines, three types of properties are inferred : measures of *discontinuity*, *tilt* and *curvness*[2].

The first property is a measure of *angular discontinuity* at the characteristic point of the primitive. Figure 2 gives three typical examples of primitives in the $x \cdot y$ and $t \cdot \theta_c$ planes. Two variables of angular variation are used for the discontinuity property : a local difference $\phi_l =< \omega_e(t_c) - \omega_s(t_c) >$ and a global difference $\phi_g =< \omega_e(t_e) - \omega_s(t_s) >$. In the case of Figure 2a, we have perceptibly, a typically continuous primitive where the local variation is small and the global variation large. Figure 2b, shows an intermediate case with to straight line segments where the local variation is important but equal to the global variation. Finally, in Figure 2c, the primitive is typically discontinuous with a large local variation and

[2] The term *curvness* is used as a heuristic curvature measure.

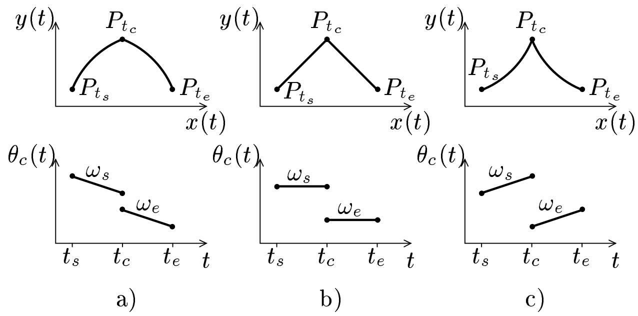

Figure 2: *Discontinuity model : a) typically continuous primitive, b) intermediate case, c) typically discontinuous primitive.*

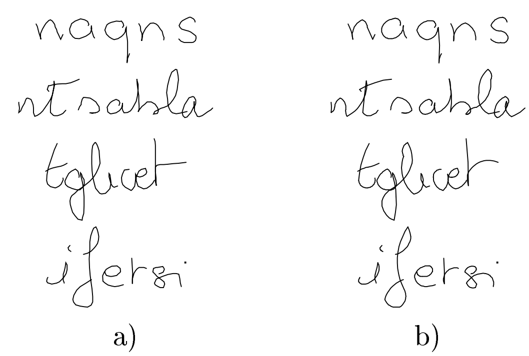

Figure 3: *Examples of cursive letter sequences. a) original traces, b) reconstructed traces.*

small global variation.

Hence, the local variation corresponds well to the discontinuity measure when it is equal to the global variation. However, a global variation that exceeds the local variation tends to reduce the importance of the latter and, reciprocally, a smaller global variation reinforces it. The proposed discontinuity measure ϕ_c thus weights the local angular variation by the square root of the ratio of local and global variations to produce a more robust measure : $\phi_c = \phi_l\sqrt{|\phi_l/\phi_g|}$.

Three properties of *tilt* are defined : the average tilt $\gamma_c = \frac{1}{2}(\omega_s(t_c) + \omega_e(t_c))$ of the primitive around the characteristic point, the tilt $\gamma_s = \omega_s(t_s)$ at the start of the primitive and the tilt $\gamma_e = \omega_e(t_e)$ at the end of the primitive.

Two properties of *curvness* are defined : the average curvness $\sigma_s = a_s(t_c - t_s)$ before the characteristic point and the average curvness $\sigma_e = a_e(t_e - t_c)$ after the characteristic point.

IV Model Evaluation

A special cursive handwriting data base was constructed to validate the handwriting model. This data base consists of 100 handwritten letter sequences written by 4 different writers. The letter sequences were generated at random while respecting the letter frequencies of a 32 000 words french dictionary and word length was varied from 5 to 7 letters. The letter sequences were assigned at random to the 4 writers (25 each) and they were asked to write them with their normal, although clean, handwriting.

The first writer naturally detaches every letter in his words. The fourth writer has a tendency to detach most letters but not all. And the two others write normal cursive script, that is, have a tendency to attach letters but not all. These subjects were chosen because they had different handwriting styles.

Table 1: *Average reconstruction error (in %).*

writer				average error
# 1	# 2	# 3	#4	
1.6	1.8	1.8	1.8	1.7

Data acquisition was conducted using a Penpad 300 digitizing tablet with a sampling frequency of 100 Hz and resolution of 0.001 inch.

Reconstruction Error Handwriting reconstruction is carried out by interpolating between successive characteristic points with circular arcs of angle σ corresponding to the curvness property of the primitive. Figure 3 gives one example of a cursive letter sequence for each of the 4 writers. On the left side of the figure, original digitized points are linked together with lines. On the right side, characteristic points are linked with circular arcs.

Table 1 gives average reconstruction errors for each writer. The reconstruction error for a given cursive word is evaluated by an objective spatial measurement defined by Maarse [15]. It computes the average reconstruction error for each *stroke* (in this case, the segments between characteristic points) by measuring the area between the original and reconstructed strokes and dividing it with the square of its length (in this case, the distance between the two characteristic points) so has to make it independent of handwriting size.

The average result of 1.7% shows that reconstruction is very good, in fact better than the results reported in [8] for all surveyed models. However, contrary to most other models, one must remember that this model is not concerned with temporal simulation

Table 2: *Average compression rates (in %).*

writer				average
# 1	# 2	# 3	# 4	rate
40.4	37.4	37.1	37.7	38.2

(which is a more difficult task), only with spatial reconstruction.

Data Compression Rate Table 2 gives data compression rates for the 4 writers. These rates are computed by assigning 2 storage units for each point digitized by the tablet and 3 storage units for each characteristic point (2 for the point itself and 1 for the "curvness" information) of the model. The rates are obtained by computing the ratio of storage units used for the model over the storage units needed for the digitized handwriting (at 100 Hz).

The average compression result of 38.2% is good considering that to obtain an equivalent performance by reduction of the sampling frequency, one would need to sample at 38 Hz which is approximatly 2 to 2.5 times the highest frequency found in handwriting signals (assuming the generally accepted higher bound in the area of 15 to 20 Hz). Futhermore, compression rates are quite stable across writers although the higher number of components for writer #1 could explain its somewhat lower performance.

Human Recognition Experiment For the recognition experiment, 4 readers were selected and each was assigned at random 25 of the 100 letter sequences in the data base. These readers were chosen with no formal pattern recognition knowledge (just their natural intuitive reading experience) and did not have any known link with the four writers. For each of them, the 25 letter sequences were printed at a scale of 1:1 in both original and reconstructed form. Each letter sequence was always printed on one sheet of paper in its original form and the corresponding reconstruction on another sheet of paper, but they were assigned randomly either to the first sheet or the second one. These two sheets of paper were presented to each reader in two different sessions, and they were asked simply to translate in carefully written block letters what they could read for each letter sequence. In the second session, the readers did not have access to their first sheet and they were never told about the handwriting model. They were told that the experiment was to compare human and machine performance.

Table 3: *Human recognition results by reader for letter sequences and their reconstructed trace (in %).*

reader								average	
# 1		# 2		# 3		#4		error	
ϵ_o	ϵ_r	ϵ_o	ϵ_r	ϵ_o	ϵ_r	ϵ_o	ϵ_r	ϵ_o	ϵ_r
9.3	9.3	4.0	4.0	10.4	13.6	1.9	2.6	6.4	7.3

Table 4: *Human recognition results by writer for letter sequences and their reconstructed trace (in %).*

writer								average	
# 1		# 2		# 3		#4		error	
ϵ_o	ϵ_r	ϵ_o	ϵ_r	ϵ_o	ϵ_r	ϵ_o	ϵ_r	ϵ_o	ϵ_r
2.0	3.3	7.9	8.6	5.2	6.5	10.5	11.1	6.4	7.3

The motivation behind this experiment is that humans are, by far, the best cursive script recognizers and, in the current state of the art, we can only hope to acheive comparable results. The proposed handwriting model is used to extract the $\mathcal{AHP}$ of our recognition system. Hence, we seek to demonstrate that, if the model based reconstruction of handwriting is just as good for human readers, then the $\mathcal{AHP}$ contains all morphologically pertinent information needed for recognition.

Table 3 summarizes the results for the readers and table 4 for the writers. The error rates are computed using the Wagner-Fischer algorithm [16] for string to string editing distance with equal cost functions for insertion, deletion and substitution. The results are expressed in percentage of the ratio of number of errors over the number of characters. ϵ_o and ϵ_r represent respectivelly the error rates for the original and reconstructed letter sequences.

Average results give a 0.9% difference in error between reconstructed and original letter sequences. Using a Student reference distribution with 3 degrees of freedom, this difference is significant at the 15% level for reader average performances and at the 1% level for writer average performances. Furthermore, compared to the 6.4% average error for original letter sequences, this difference is relatively small. Here are some other miscealeneous statistics :

1. number of errors per letter sequence vary from 0 to 3 ;
2. worst error is not always for reconstructed letters : in 7 cases, it was for the original letter sequence compared with 11 cases for the recon-

struted one ;

3. in 5 cases, the original letter sequence was read erroneously while the reconstructed one was read correctly compared with 10 cases for the reciprocal ;
4. 26% of original letter sequences were read erroneously compared with 31% for reconstructed ones.

Looking at individual observed reader performances, considerable fluctuations can be seen. The best reader outperforms the worst by a factor greather than 5. Also, the observed difference of error rates between original and reconstructed letter sequences vary from 0% to 3.2%. Similarly, observed writer performances, that is readability, have shown important fluctuations.

V Conclusion

The main purpose of this paper was to present and support our model based $\mathcal{A}$ttributed $\mathcal{H}$andwriting $\mathcal{P}$rimitive used in a fuzzy-syntactic allograph modeling approach to cursive script recognition [7]. With this primitive, allographs can be modeled using fuzzy-attributed shape grammars and segmented by an adequate parser [5]. The proposed handwriting model was found to give good data compression (38% of digitized data) and very good reconstruction quality (1.7% error). On average, human readers have made 1% more errors with the model based reconstructed letter sequences, although this difference is not very significant (at the 15% level). Also, it was found that readers have made between 1.9% to 10.4% of errors while reading cursive script with no available lexical context.

Another purpose of this paper was to propose a protocol for evaluating the quality of the primitive extraction phase of any structural cursive script recognition method. By using human readers and reconstructing handwriting only from the information of the primitives — the information available to the recognition algorithm — it is possible to approximatly estimate the maximum percentage of errors that can be directly attributed to the primitive extraction phase, assuming that linguistic knowledge can only reduced or maintain this percentage.

References

[1] Fu K.S., *Syntactic Pattern Recognition and Applications*, Prentice-Hall, Englewood Cliffs, 1982.

[2] Bozinovic R.M., Srihari S.N., "Off-Line Cursive Script Word Recognition", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. PAMI-11, no. 1, pp. 68-84, jan. 1989.

[3] Higgins C.A. and Ford D.M., "A New Segmentation Method for Cursive Script Recognition", *Proc. of* 2$^{\text{th}}$ *International Workshop on Frontiers in Handwriting Recognition*, Chateau de Bonas, France, 23-27 September 1991, pp. 241-252.

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

[5] Parizeau M., Plamondon R. and Lorette G., "Fuzzy-Shape Grammars for Cursive Script Recognition", *Advances in Structural and Syntactic Pattern Recognition*, H. Bunke (editor), World Scientific Publishing, August 26-28, pp. 320-332, 1992.

[6] Fu K.S., "A Step Towards Unification of Syntactic and Statistical Pattern Recognition", *IEEE trans on Pattern Analysis and Machine Intelligence*, vol. PAMI-5, no. 2, pp. 200-205, 1983.

[7] Parizeau M., "Système de reconnaissance d'écriture cursive et bloc-notes électronique", Ph.D. Thesis, Ecole Polytechnique de Montréal, to be published summer 92.

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

[9] Plamondon R., "Handwriting Control : A Functional Model", *Models of brain function*, Cotterill R.M.J. (editor), Cambridge University Press, 1989, pp. 563-574.

[10] Plamondon R., "Steps Towards the Production of an Electronic Pen Pad", *Proc. International Conference on Document Analysis and Recognition*, St-Malo, September 30 to October 2, 1991, pp. 361-371.

[11] Plamondon R., "What Does Velocity Information Reveals about Handwriting Generation", *Proc. of the* 5$^{\text{th}}$ *International Handwriting Conference*, Tempe Arizona, October 28-30, 1991.

[12] Brown M.K., Ganapathy S., "Preprocessing Techniques for Cursive Script Word Recognition", *Pattern Recognition (GB)*, vol. 16, no. 5, pp. 447-458, 1983.

[13] Guerfali W., "La restauration de l'écriture cursive", M.Sc.A. Dissertation, Ecole Polytechnique de Montréal, 1990.

[14] Guerfali W., Plamondon R., "An Editing Tool for Handwritten Text", *Proc. of the* 5^{th} *Handwriting Conference*, Tempe Arizona, October 28-30, 1991.

[15] Maarse F.J., *The Study of Handwriting Movement : Peripheral Models and Signal Processing Techniques.* Lisse, The Netherlands : Swets & Zertlinger, 1987, p.160.

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