INTERNATIONAL COMPUTER SCIENCE INSTITUTE

1947 Center St. • Suite 600 • Berkeley, California 94704-1198 • (510) 643-9153 • FAX (510) 643-7684

A cognitive off-line model for motor interpretation of handwritten words

Claudio M. Privitera

International Computer Science Institute

1947 Center St. Suite 600, Berkeley, California 94704-1105

TR-95-037

August 1995

Abstract

The image of a word or a generic hand made drawing on a piece of paper is usually characterized by a series of interfering zones where the cursive trace intersects itself or printed lines already present on the writing surface. In this zone, the odometric information is ambiguous and any trivial inference on the original pen tip movement cannot be done. In this article, starting from some basic cognitive considerations, a general procedure is developed to analize a generic image of a word or a common hand made scribble. This approach allows to detect each ambiguity part of the image and then interpretate them to finally recover a part of the original temporal information.

1 Introduction

The generation of a graphic shape on a writing surface during handwriting or drawing is the result of a complex motor planning process starting from an input allograph representation of the shape and then producing a sequence (partially overlapped) of primitive movements (called motor strokes) of the hand-pen tip system which traces that shape [16], [11].

This very general motor aspect of handwriting seems to play a key role also in the complementary behavior of the generation that is the off-line vision of an image representing a hand made message. In fact, the human visual system seems to exploit in a reverse operation, the same mechanism taking place during the generation: the image is analized in order to detect the corresponding sequence of primitive motor components which have then, to be interpreted on the basis of the general knowledge about motor control [1], [7]. Finally, the syntactical class of the word (or the drawing) is determined.

According to this motor approach to handwriting generation and recognition, a new paradigm for off-line cursive analysis and recognition can be set up. Unlike many of the traditional models previously published, which transform the original word image into a different representation space which is more handable by the analysis processes (wholistic approaches or thinning processes for example), this new approach tends to restore, among other things, the original odometric information by following in the image the sequence of pen movements to finally recover the original curvilinear trajectory of the pen tip [3], [12], [6]. In this case, it is possible to preserve for the subsequent analysis process (word reconstruction and recognition), part of the original information that could be irreversibly destroyed otherwise during the above mentioned traditional models for word image processing.

The major problem with this approach occurs when, parts of the words belonging to different temporal events get into contact. This is the case when different characters of the word touch each other or when additional non-textual lines such as the lines of a notebook are presented on the writing surface and obviously interfer with the handwritten trace. These interfering zones are parts of the word image where the *memory* of the original course of the pen tip trajectory is ambiguous and the process of odometric information recovery is difficult. Figure 1.*a* is a classical morphological example of this type of ambiguity: the image shows a part of a word where several strokes touch and cross each other without any explicit information about their relationship in the sense that is is not obvious for an automate system to discriminate the pairing of different components of the interfering zone.

The human visual system seems to resolve this ambiguities quite easily by exploiting general cognitive rules such as the propensity to follow a line wherever it crosses another line, considering the minimal curvature change (*good continuation* rule Figure 2.*a*) or the tendency to cluster the image into a set of closed shapes (*closure* law in Figure 2.*b*: humans recognize two shapes in the picture touching each other, rather than a line crossing itself). Another important clues for the correct interpretation of an interfering zone is the width of the stroke: two strokes section with different widths cannot belong to the same downwards or upwards primitive pen movement [6]; Figure 1.*b* is a good example of this consideration.

These cognitive rules of the visual system behavior, known as Gestalt laws [19] could provide important indications for the development of efficient algorithms to analize interfering zones: a good attempt has been studied for example by [8].

a) b)

Figure 1: Two different examples of interfering strokes. In picture b) only on the basis of the width of the interfering strokes (the bottom left counter clock wise stroke is thinner than the whole arc) it is possible to detect two different handwriting events: the two clockwise and the counter clockwise arcs.

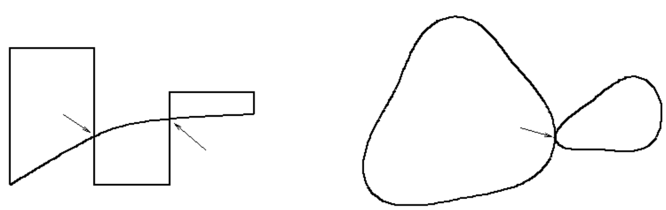


Figure 2: Gestalt laws: general cognitive properties useful for interfering interpretation.

In this article, we define a general and model for the detection and interpretation of the interfering strokes of a handwritten word (or generally speaking a generic hand made drawing): the system analizes an input bitmap image representing a generic handwritten shape and is able to detect the zones where two or more strokes interfer with each other. Consequently, each of these zones is interpreted in order to pair parts of strokes corresponding to the same original trajectory of the pen tip: in this way, the original odometric information is recovered and the ambiguity is resolved.

The paper is divided into two principal sections corresponding to the two levels of image processing: initially, the image is analyzed in order to detect every ambiguity points of a word that is the image areas where there is an interfering zone (section 2). Each interfering zone is represented by means of a small window located on the image. Finally, each of these windows is analyzed in order to evaluate, on the basis of a general mathematical model based on the above mentioned cognitive rules, the discrimination of each strokes interacting in the zone (section 3).

The whole process has been tested on a database of images corresponding to the handwriting of several writers and the results and relative discussions are reported in the last section.

2 The Detection of the interfering zones in the image

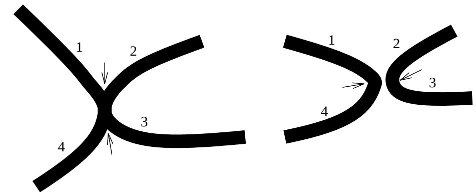
The first phase of the processing is an image scanning in order to detect the ambiguity points of the word: in other words, the system looks for the points where, the continuous trace of the pen tip on the paper intersect another trace loosing its original deterministic course.

The scanning operation is achieved by locating a window of pre-fixed dimension where the contour of handwritten line encompasses some point of maximum curvature (in elementary geometry for example, two crossing lines generate at least two acute or square angles): if the strokes conformation inside the window is a connected space and the strokes cross the boundary of the window three or more times, this means that the window is effectively located on an interfering zone (see Figure 3). Finally the rescaling of the window size is executed in order to take into consideration only the smallest part of the detected interfering zone.

2.1 Preprocessing of the image

The images composing the database are grey-level pictures acquired by a common scanner device (300 dpi) and then converted into a bitmap format by means of a Gaussian convolution filter and a threshold process.

In order to evaluate the curvature of the handwritten lines on the writing surface so that the maximum curvature points can be finally detected, we have first of all to extract the contour of the image. Using digital images, the contour of the handwritten shape is simply evaluated in terms of the discrete *closeness* evaluation between the *black* pixels belonging to the shape with the *white* pixels belonging to the respective complement: in this manner, the contour of the handwritten shape is defined by the sequence of the *black* pixels with a distance equals to 1 from the foreground (these pixels will be referred 1-pixels henceforth).



a) b)

Figure 3: Detection of interfering strokes: the window (the dotted square in the figure, see the text) scans the image looking for parts of the word where maximum curvature points are present (the arrows). In a) the window is effectively located on an interfering stroke because the strokes cross the boundary of the window four times and the space of the ink trace covering the area of the window is connected.

Several techniques do exist in literature for the distance transformation in digital image (see for example [4], [18]): the general effort is addressed to find out algorithms which provide the best approximation of the Euclidean metric with the less computational cost. The problem of Euclidean approximation is especially needed for large section shape for the most internal pixels.

A classical method is to scan all the image and for each pixel take into consideration only a small neighborhood: the distance of that pixel is consequently determinated in the basis of the pixels belonging to that neighborhood.

In our context we can assume some basic simplification since we are interested simply to detect the 1-pixels of the handwritten word border: in this case we can neglect the problems concerning Euclidean approximation and grant a privilege to faster algorithm. In this sense a neighborhood defined by the *cityblock* distance [1] seems to be the most appropriate for this pourpose [4] and the process of 1-pixels detection can be consequently defined by a scanning process which label as 1-pixels the pixels holding the following condition:

$$min_{(Q\ :\ d_c(P,Q)=1)}\ \ c_Q = 0 \tag{1}$$

where c is the *color* of the pixel that is 1 if the pixel belongs to the background and 0 for the foreground and where d_c is the cityblock distance.

At the end of the above mentioned process, the contour of the handwritten line is defined by a continuous sequence of 1-pixels which will be analyzed by the next processing step in order to finally determinate the curvature function of the line.

2.2 Evaluation of the contour curvature and maximum curvature points extraction

The curvature of the stroke contour plays a key role for two fundamental factors: firstly, the maximum curvature are important reference points to detect where an interfering zone is located in the image. Moreover, as mentioned in the introductory section about the good continuation law, the curvature of the strokes before (or after) the interfering zone, is very important for the correct interpretation of the stroke superimposition.

[1] The cityblock distance is a regular metric defined in a $2D$ discrete space by the following expression: $d_c(P,Q) = |P_1 - P_2| + |Q_1 - Q_2|$.

In the real Euclidean plane, the expression of the curvature function $C(x, y)$ of a 2D contour is defined as the rate of change of slope $\partial\varphi$ as a function of the arc length ∂s, $C(x, y) = \frac{\partial\varphi}{\partial s}$, (as shown in Figure 4) and can be expressed in terms of derivative of x and y coordinates by the following well known equation:

$$\frac{\frac{\partial^2 y}{\partial x^2}}{(1 + (\frac{dy}{dx})^2)^{\frac{3}{2}}} \tag{2}$$

Dealing with digital curves, such as handwritten traces, where the contour of the strokes is represented by a finite and discrete path of 1-pixels located such as black squares on the chessboard, the discrete interpretation of the above curvature equation is not so straightforward. In fact, simply replacing the derivatives by differences, and taking into account that the discrete approximation of the Euclidean distance between two consecutive pixels is 0 or 1, then, the computed digital curvature results such as only by a multiple of 45°. Moreover, equation 1 represents only a local definition of curvature whereas, in order to detect more general properties of the handwriting line, such as good continuation, or closure law, we need a more global evaluation of the curvature so that, the general trend of original pen-tip movement can be captured.

A possible solution could be to evaluate the curvature function of the digital curve by simpling smoothing (and then differentiate) the curve with a $2D$ filter such as a Gaussian filter [13], [10].

Another possibility is to approximate the local curvature taking into consideration a specific sufficiently large region of support. The magnitude of region of support plays obviously a key role on the balance between noise sensitivity and resolution of the curvature: too large a region will hide small changes of the line slope, whereas a small region will generate very noisy curvatures.

A general solution to the above mentioned tradeoff is taking into account a set of curvature estimations regarding all the points belonging to region of support and finally rescale the size of this region in the basis of local information about the curve [17], [15].

In Figure 5 for example, a generic piece of a digital curve is shown where, p_i represent the point where the curvature has to be evaluated and, the sequence of points $p_{i-m}, p_{i-m-1}, ..., p_i, p_{i+1}, ..., p_{i+m}$ represents a prefixed range of region of support. In order to opportunely rescale the region of support (for example to discard the right ascending part of the curve which is non correlated with the course of the curvature function in p_i) for each point of the region, an approximation of the curvature function (also called *measures of significance* [14]) with regard to the point p_i is consequently measured: for example, a good and fast measures of significance is represented by the function S_{ik} defining the perpendicular distance of the point p_i to the chord joining the points p_{i-k} and p_{i+k} (see Figure 5).

In this case, the rescaling operation of region of support is defined by means of an iterative process starting with $k = 1$ and determining the first point p_{m^*} such that $S_{im^*} > S_{i,m^*+1}$: in this manner rangem^* represents the new range of region of support.

Finally, the curvature function of the point p_i can be represented by the following average

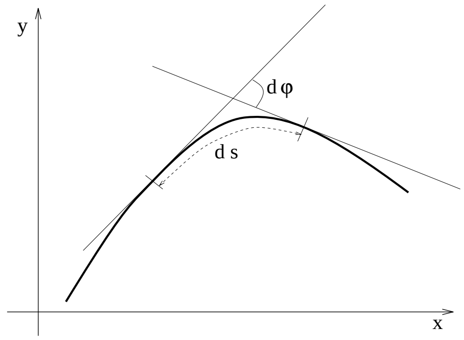


Figure 4: The definition of curvature for a $2D$ continuous function.

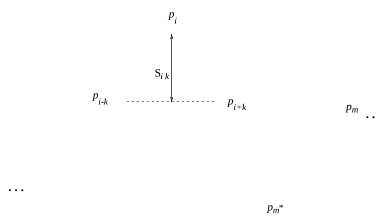


Figure 5: Curvature estimation of a generic point p_i in a digital curve: the distance S_ik represents the measure of significance with two neighborhood belonging to region of support; the point p_m represents the first approximation of the region whereas P_{m^*} the last final rescaled region.

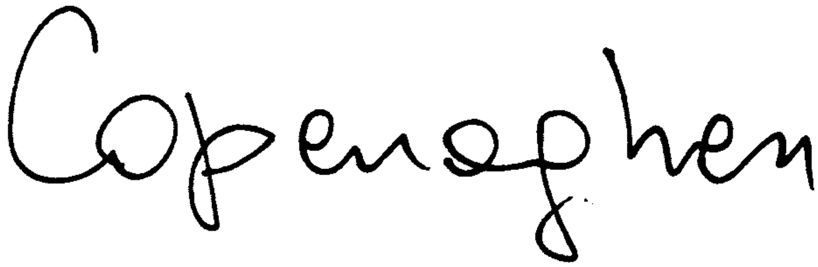
computed over all the curvature estimations belonging to the rescaled region (see also [5]):

$$\mathcal{C}(p_i) = \frac{1}{m^* - k_0 + 1} \sum_{k=k_0}^{m^*} S_{ik} \tag{3}$$

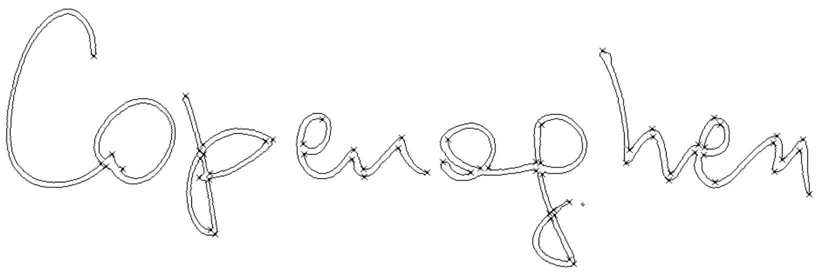
where k_0 can be more than 1 in order to discard the closest points of p_i which represent the worse curvature approximation.

Figure 6 shows the result of the above mentioned process as applied to an image displaying a handwritten word: in particular only the 1-pixels belonging to the contour are outlined in the resulting image and the X signs along the contour represent the local maximum of the evaluated curvature function. The maximum of the curvature function (or dominant points) are determinated in the basis of local informations: in particular a point p_i is a local maximum if, $\forall\ j$ belonging to the rescaled region of support ($k_0 \leq j \leq m^*$) the condition $\mathcal{C}(p_i) \geq \mathcal{C}(p_j)$ holds. Small local maxima are simply discarded with a threshold process.

It is noteworthy the concentration of maximum curvature points in those parts of the word, where two or more strokes interfer each other. This strategic concentration of maximum curvature points outlines the presence of ambiguity points in the word and it will be determinating for a fast and efficient detection procedure of the interfering zones of the word.



a)



b)

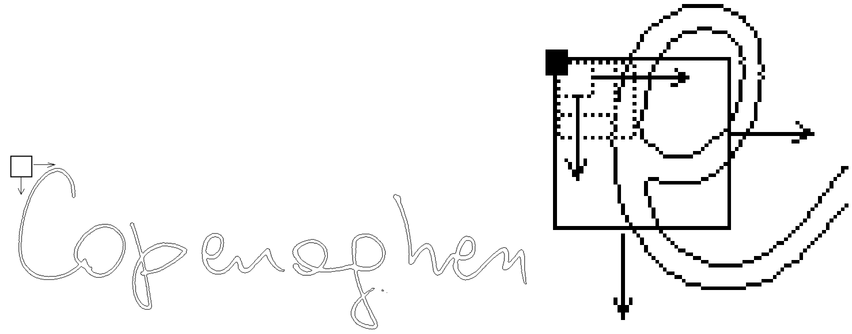
Figure 6: The result of curvature evaluation and maximum detection processing: in particular only the contour of the word is outlighted (the 1-pixels in b)) where the X sign along the contour represent the maximum curvature points.

2.3 Image scanning and interfering zones detection

The maximum curvature points of the handwritten word contour extracted during the previous process represent important key references for the detection of which part of the handwritten word could include ambiguity points: in fact, as shown in the previous section, the crossing of two planar shape (in our case, the development of a continuous ink trace) necessarily produces a concentration (at least two) of acute or square angles that, in a such way, outline the conformation of the intersection.

In order to effectively detect the conformation of the interfering zone to accurately define the interfering strokes, our algorithm scans the image with a moving window of a pre-fixed initial size from the top left side of the image (see Figure 7) and, each time that the border window enclose a zone of dominant points concentration, a series of checking processes take place.

The first operation is to evaluate the connectivity of the part of the image defined by the window border and then the second one is compute the crossing number of handwritten line through the window border: if the space representing the handwritten line inside the surface of the window is connected (see Figure 3), and the strokes cross the border of the window more than two times (for example in Figure 3, the strokes cross the window contour four times), this means that the window is effectively covering an interference part of the handwritten word. Finally, the window dimensions are scaled in order to improve the representation of the interference zones detected by that window without mislead any informations about the conformation of the handwritten lines in that area.



a) b)

Figure 7: Detection of interference area. a): a maximum size window scans all the image; b): the window is located on an interference area (the continuos square) and keeping constant the top-left vertix of the original window a series of optimal smaller window (the dotted squares) are hypothesized in order to improve the resolution of the interference.

Let represent the perimeter of a window $\square$ with a function $\mathcal{P}(\square)$ defined by a minimum ($\mathcal{P}_{min}$) and maximum ($\mathcal{P}_{max}$) value: in general, the maximum dimension (which is the only input parameter needed for this process, whereas the minimum value could correspond just to the dimension of the single pixel) cannot be too big because some false crossing stroke would be detectable. On the other hand, if the maximum dimension is less than the section of the handwritten line, the window would not be able to circumcise the interfering zone. The exploitation of a general heuristic function of the height of the word, has been sufficient for a good performance of the whole algorithm: in details, the maximum side $\mathcal{P}_{max}$ of the window is set to $\frac{Height}{6}$ where $Height$ represents the height of the handwritten shape.

Let $\mathcal{R}$ be a boolean function representing the above mentioned checking process so that $\mathcal{R}(\square)$ is true if the window $\square$ circumcises an interference area (the internal area is connected and the window border is crossed more or equal to three times by the handwritten line) and let $\mathcal{M}(\square) = p_{max1}, p_{max2}, ...$ be the set of the dominants points circumcised by the window. Let finally the set $\mathcal{D} = \square_1, \square_2, ...$ be all the windows representing the interference zones detected during the scanning process.

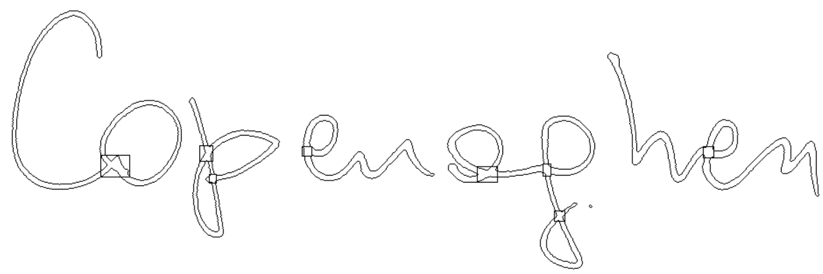


Figure 8: The small windows along the development of the image represent the ambiguity points of a handwritten word detected by the algorithm.

The following steps represent the exploited general procedure for interference zone detecting and window size rescaling:
1) move the window $\Box$ (with the maximum dimension) as shown in Figure 7;
$if\ \mathcal{R}(\Box)\ then$
 2) keeping constant the top-left vertex of the window (see Figure 7):
 $for\ \mathcal{P}(\Box) = \mathcal{P}_{min}\ to\ \mathcal{P}_{max}$
 $if\ \mathcal{R}(\Box)$
 2a) $if\ \exists\ \ \Box_i \in \mathcal{D} : \mathcal{M}(\Box) \cap \mathcal{M}(\Box_i) \neq \{\}\ and\ \mathcal{M}(\Box) \neq \mathcal{M}(\Box_i)$
 $then$ discard $\Box$ and $go\ to$ 1
 2b) $if\ \exists\ \ \Box_i : \mathcal{M}(\Box) \supseteq \mathcal{M}(\Box_i)$
 $then$
 $if\ \mathcal{P}(\Box) < \mathcal{P}(\Box_i)\ then$ discard $\Box_i$ and keep $\Box$
 $else$ discard $\Box$

In order words, the process of window rescaling is performed by an iterative procedure that firstly detect the interference area (step 1) and then try to locate the smallest window in that area (step 2): step $2a$) is necessary to avoid overlapping between windows of different interference areas. In step $2b$), two windows are located over the same interference area (the windows have some dominant points in common): in this case, only the the smaller window is kept to represent the interference.

Figure 8 shows an example of the final result of this interfering detection algorithm: as can be noted, all the parts of the word containing an overlapping between different strokes, are detected. Each square represents the final dimension of the windows at the end of the algorithm.

It worth noting that this general approach to handwritten image analysis is consistent with the cognitive behavior of the visual system: in particular, human visual system seems to have a strong specificity to the points of a picture where the curvature values of the boundary are high, so that a complex input shape is initially analized on the basis of these maximum concavity regions [9], [2].

In the following section, each stroke arriving at or departing from an interfering window is analized in terms of its direction, bend value and section, in order to finally pair strokes belonging the same movement and to recover the original trajectory.

3 The recovery of the original movement direction

The procedure is based on parameters that seem to be involved in the human visual processing of hand made graphic forms. These parameters are: the section of the stroke (strokes belonging to the same primitive movement have to have approximately the same width section); the good continuation rule (the handwriting process is a *smooth* movement where the curvature cannot abruptly change but rather, the writer tends to keep the same value of curvature for each primitive movement) and finally the closure law (the tendency of handwriting generation system to alternate tract of opposite curvature during the generation of a sequence of primitive strokes, think for example the movement of the pen tip during a repeated generation of $\ell aeo...$). In the off-line analysis of the ambiguity points of a word, these *perceptivo-motor* constraints of the handwriting generation have to be taken into account for a plausible characterization of the original movement.

3.1 The analysis of interfering strokes

For each window representing an interference, a specific series of measurements have to be made: the most important among these is the curvature evaluation of the departing and arriving strokes, because, as previously discussed, the curvature represents a key parameter for the correct pairing of strokes belonging to the same movement. In other words, we attempt to evaluate the level of mechanical inertia associated with each possible pair of strokes presents in the interference zone.

The evaluation of the curvature is performed as in the preprocessing stage: the only difference is that, each stroke contour getting in the window is paired with the corresponding contour of another strokes of the interference (see Figure 9) generating a set (all possible combinations) of *pairing hypothesis*. For each of these pairing hypothesis, the corresponding curvature function is then computed. In this way, the curvature function can provide an important clue about how the curvature would change, if the corresponding pair of stroke contour would be effectively associated.

Figure 9 shows the curvature function for three pairing hypothesis: the strokes numbered 0 with the remaining strokes of the window. In particular, each curve display the mean curvature function evaluated taking into consideration the two quasi-parallel contours characterizing each single strokes. The peak of the curvature in two of these case underlines quite an abrupt change that could be represent an unnatural movement.

The evaluation of the curvature function stability can be defined by the difference between the maximum and minimum value of the function in the definition interval. In particular, in order to better represent smooth changing of the function, for each couple of points $x_i x_j$ a sort of discrete derivative is consequently computed:

$$\wp(x_i, x_j) = \frac{\mid f(x_i) - f(x_j) \mid}{\mid x_i - x_j \mid} \tag{4}$$

The parameter $\wp(\cdot)$ is the Lipschitz parameter and represents the level of the uniform continuity of the function. Associating to each function f the maximum value $\wp_m$ such that $\wp_m = max_{x^i, x^j} \wp(\cdot)$ then, the smaller is this value the smoother is the function. In other words, curvarture functions with abrupt changes (for example a peak such as in a

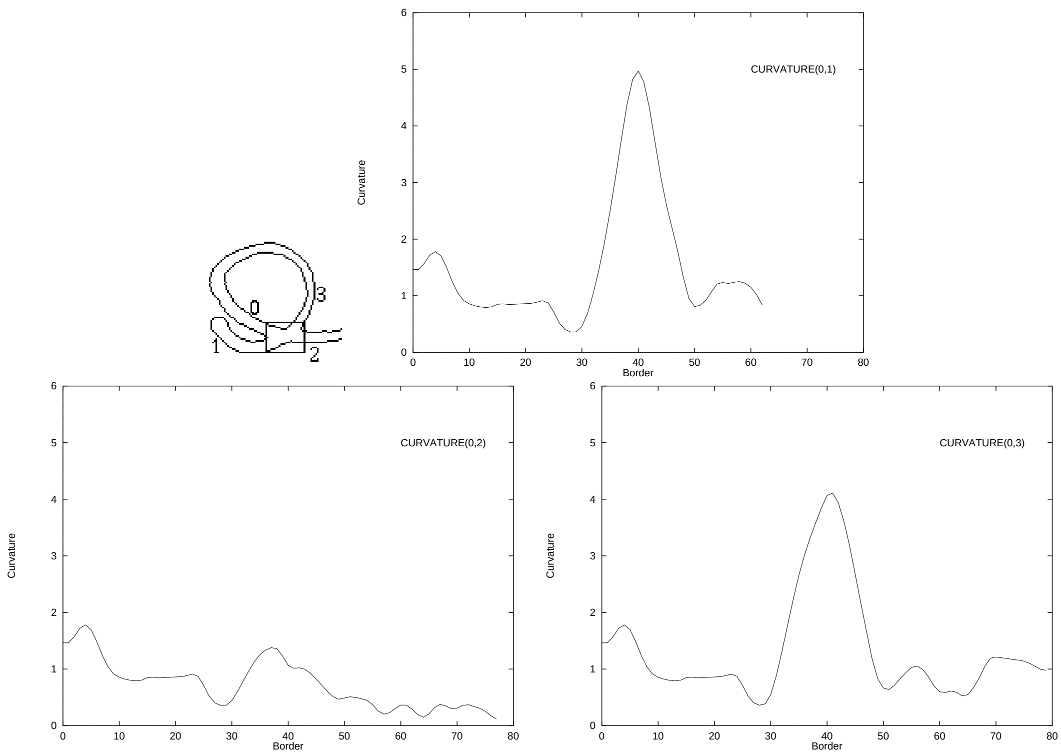


Figure 9: A small section of the word containing an interference (top left: the square represents the interference window defined in the previous processes) and the evolution of the curvature function for three different coupling hypothesis. In particular, the stroke numbered 0, with all the other strokes. The graphs show the curvature function along an interval of 60 1-pixels around the interference window.

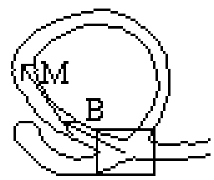


Figure 10: Bend direction evaluation.

gaussian with a small variance) or large alternation (even if the function is continue) will be represented by a value of $\wp_m$ bigger than smooth (even if non monotonic) curvature function.

The value $\wp_m$ associated with each curvature function is consequently a good parameter for the representation of the good continuation paradigm; in the case of the example shown in Figure 9, the values of $\wp_m$ are: 0.65 for $Curvature(0,1)$, 0.24 for $Curvature(0,2)$ and finally 0.51 for $Curvature(0,3)$.

On the other hand, while the parameter $\wp_m$ represents a good continuity criterion for the pairing hypothesis, we need also a measure for making hypothesis on the level of closure. The bend direction of the two corresponding strokes of a pair hypothesis is used for this purpose. The direction of the curvature $\Im_i$ (clock wise or counter clock wise) associated with each stroke l_i of the interfering zone is represented by the sign of the z projection (the normal axis to the image plane) of the following vectorial product of the two vector $\vec{B}_i$ (connecting the center of the window and a point along the stroke) and $\vec{M}_i$ (connecting the previous point with another point along the stroke, this latter point is located roughly at the level of the first 1-pixel of the curvature function, see Figure 10):

$$\Im_i = (\vec{B}_i \times \vec{M}_i)_{\mathbf{z}} \tag{5}$$

Finally, the third parameter that has to take part in the interpretation of the interfering is the relationship $S(l_i, l_j)$ between the section of the strokes for each pairing hypothesis; the section of the two strokes d_i and d_j is easily evaluated along the border of the window (see Figure 10) and the evaluation method is a gaussian $S(l_i, l_j) = exp(\frac{-(d_i - d_j)^2}{\sigma})$ with a fixed variance

With respect to the cognitive issues raised in few part of this paper, the three parameters, $\wp(m)$, $\Im_i$ and $S(l_i, l_j)$ represent the basic informations that are needed for a correct and plausible interpretation of the interfering zone.

All the above listed parameters are condensed in the following global interpretation function $\Omega(l_i, l_j)$ which practically assigns a score for each pairing hypothesis of the two strokes l_i and l_j:

$$\Omega(l_i, l_j) = \begin{cases} \wp_m(1 - S(l_i, l_j))\eta & \text{if } sign(\Im_i) \neq sign(\Im_i) \\ \wp_m(1 - S(l_i, l_j)) & \text{otherwise} \end{cases}$$

where $\eta > 1$ is an amplification factor that improve the evaluation of that hypothesis when the corresponding bend direction are in opposite sense ($sign(\Im_i) \neq sign(\Im_i)$: that means the closure law).

The global interpretation function tends to follow the same indication as the curvature function: small values have to be interpreted such as a good probability of association between the two corresponding strokes.

The final interpretation of the interfering zone takes place by associating each other the strokes of the interference so that the sum of the score is the optimum over all the possible combinations.

In the next section, taking into consideration a database of handwritten word images, the efficiency of this general equation will be evaluated.

3.2 Experiments

The interpretation model has been applied on a database of 200 images representing names of international cities wrote by 6 different writers: each writer was asked to write six different names on a $A4$ piece of paper. Each picture was first preprocessed in order to extract the contour and localize the maximum curvature points. Consequently, in the basis of the detected maximum curvature points, the procedure for the interfering zone detection was applied. Finally, considering each interference zone, the interpretation rule was applied in order to detect the best combination among the pairing hypothesis. The best combination was computed by finding the minimum of $\Omega(l_i, l_j)$ for each hypothesis.

Figure 11 shows some typical results of the algorithm: the double arrows represent the paired stroke whereas the single arrow a stroke that ends in the interference zone: this happens for instance in the first e of the word shown in Figure 11.a, where the interfering stroke is composed of three strokes and necessarily one of these has to finish there. Also in the case of the last e (Figure 11.b) the model cuts one of the pairing hypothesis (stroke number 0 and 2) This is because the minimum sum score $\Omega(1,2) + \Omega(0,3) = 0.63$ overtakes a prefixed maximum threshold (0.5) and consequently only the best hypothesis pair score is taken into consideration ($\Omega(1,3) = 0.12$) cutting the remaining strokes. It should be noted that, this cutting does not force an incorrect interpretation of the interference zone because the final hypothesis produced by the process is consistent and plausible from a handwriting point of viwe.. The scribble in Figure 11.c represents another case for testing the model: in particular the crossing in the middle of the picture is interpreted by giving priority to the closure law.

The results of the model on our database are quite encouraging: the error during the interfering zone detection is 89.% which is not so bad considering that it consists of the whole image and that each image is characterized by several interference zone. This error represents indeed the percentage of images where at least one interference zone was not detected. Taking into consideration only the detected interfering zones, the model was able to correctly interpretate the 94.% of the interfering strokes in the sense that the results were effectively correlated with the original movement of the pen tip during the word generation.

4 Conclusions

Although, the movement of the pen-tip on the writing surface during the handwriting process is a continuous movement assumed to be composed by a series of distinct and primitive movements (called strokes) partially superimposed to each other, the graphical

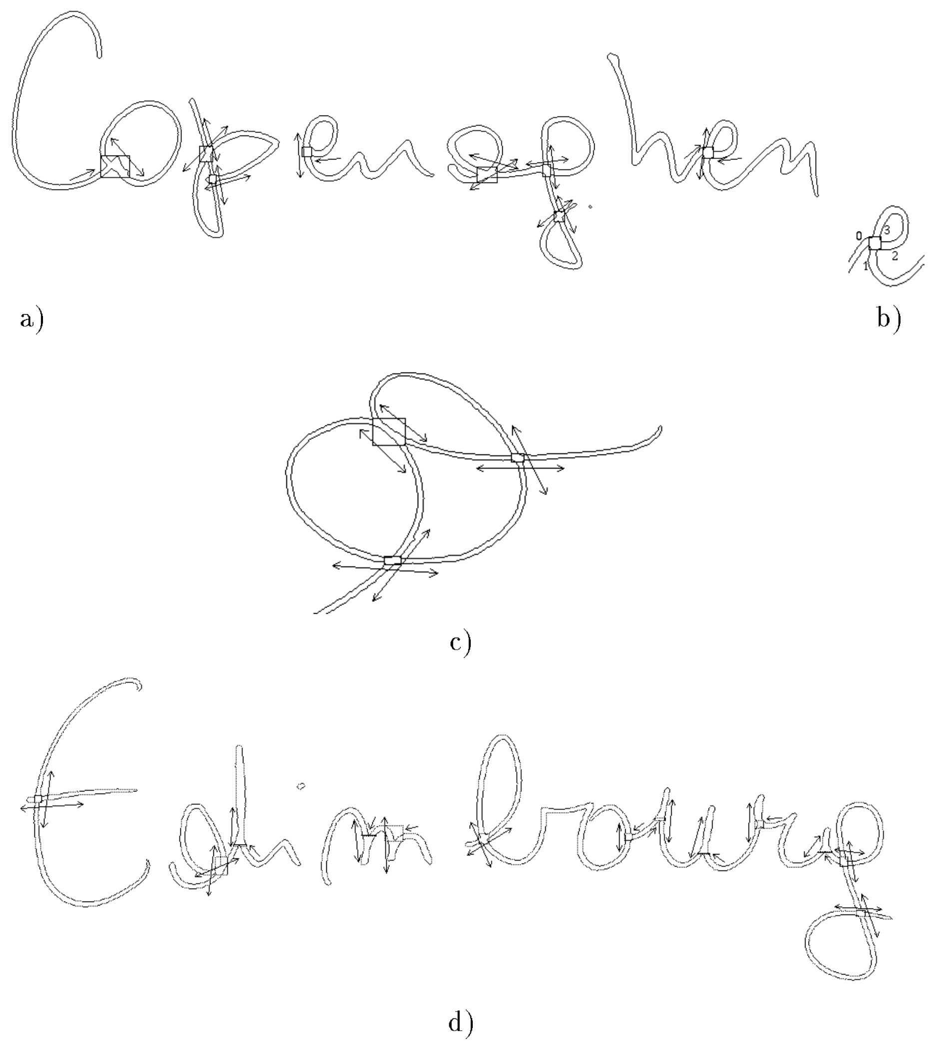


Figure 11: The figures represents three different examples of ambiguity points interpretation performed by the model: the double arroes represent a pairing between two different parts of the same stroke whereas the single arrows a stroke that presumly start or end in the interfering zone.

result of the handwriting or, in other words, the off-line vision of the word image, does not explicitly contain any indication about the original odometric information.

This is due to two fundamental causes. First, the presence on the writing surface of pre-printed lines, such as the horizontal lines on the page of a notebook, so that the ink trace left by the writing device can often overlapping these non-textual contour. Second, there is an overlapping between two different parts of a word due to accidental contacts between different character or for the effective allographic conformation of a character (the classical example is the crossing strokes composing the character ℓ).

Both from a cognitive point of view concerning the visual process used by humans and in the light of a recent trend in off-line handwriting analysis and recognition where a word is analyzed in order to recover the whole motor temporal information of the generating process, the detection of the ambiguity parts of a word and their corresponding interpretation, is a fundamental initial preprocessing for an automatic recognition system. In fact, resolving all the ambiguity parts, the word can be decomposed into the original sequence of motor strokes (such as in on-line methods).

In this paper, we have proposed a general automatic procedure that first detect the ambiguity parts and then interpret them so that, each interfering stroke can be associated with its corresponding continuation. The followed paradigm is based on some biological properties regarding human vision performance easily implementable from a mathematical point of view.

The computational cost of the whole system is very low and the time processing for each word very short (few seconds, depending on the size of the word); the whole procedure is also very easy to implement. All these characteristics underline an interesting versatility of the system in many aspects of the off-line handwriting analysis.

The model has been tested on database of handwritten word images and the results are quite encouraging, especially considering the fact that, in the context of a cursive off-line recognition system, our approach would represent a basic preprocessing stage as well as a knowledge data base for the following prediction-verification steps. In this context, a 100.0% accuracy is not necessary at the preprocessing level since hypothesis raised at higher levels in the recognition process could be used to check for missing or missinterpreted interference zones.

References

[1] M.K. Babcock and J.J. Freyd. Perception of dynamic information in static handwritten forms. *American Journal of Psychology*, 101:111–130, 1988.

[2] I. Bierderman. Recognition-by-components: a theory of human image understanding. *Psychological Review*, 94:115–147, 1987.

[3] G. Boccignone, A. Chianese, L.P. Cordella, and A. Marcelli. Recovering dynamic information from static handwriting. *Pattern Recognition*, 26(3):409–419, 1993.

[4] G. Borgefors. Distance transformation in digital images. *Computer Vision, Graphics, and Image Processing*, 34:344–371, 1986.

[5] L.S. Davis and A. Rosenfeld. Detection of step edges in noisy one-dimensional data. *IEEE Trans. Comput.*, C-24:1006–1010, 1975.

[6] D.S. Doermann and A. Rosenfeld. Recovery of temporal information from static images of handwriting. *International Journal of Computer Vision*, 15:143–164, 1995.

[7] J.J. Freyd. Representing the dynamics of a static form. *Memory and Cognition*, 11 (4):342–346, 1983.

[8] V. Govindaraju and N. Srihari. Separating handwritten text from overlapping non-textual contours. In *Proc. of International Workshop on Frontiers in Handwriting Recognition*, pages 111–119, Chateau de Bonas, France, 1991.

[9] D.D. Hoffman and W. Richards. Parts of recognition. *Cognition*, 18:65–96, 1985.

[10] F. Mokhtarian and A. Mackworth. Scale-based description and recognition of planar curves and two-dimensional shapes. *IEEE Trans. PAMI*, 8(1):34–43, 1986.

[11] R. Plamondon and C.M. Privitera. A neural model for learning and generating rapid movement sequence. *in press on Biological Cybernetics.*

[12] C.M. Privitera and R. Plamondon. A general system for off-line handwriting segmentation. In *ICDAR'95*, Montreal, Quebec, 1995.

[13] A. Rattarangsi and R.T. Chin. Scale-based detection of corners of planar curves. *IEEE trans. on PAMI*, 14(4):430–449, 1992.

[14] R. Rosenberg. The analysis of convex blobs. *Comput. Graphics Image Processing*, 1:183–192, 1972.

[15] A. Rosenfeld and E. Johnston. Angle detection on digital curves. *IEEE Trans. Comput.*, C-22:875–878, 1973.

[16] R.Plamondon. A kinematic theory of rapid human movements. part i: Movement representation and generation. *Biological Cybernetics*, 72:297–307, 1995.

[17] C. Teh and R.T. Chin. On the detection of dominant points on digital curves. *IEEE trans. on PAMI*, 11(8):859–872, 1989.

[18] B.J.H. Verwer. Local distances for distance transformations in two and three dimensions. *Pattern Recognition Letters*, 12:671–682, 1991.

[19] D.K.W. Walters. Selection of image primitives for general-purpose visual processing. *Computer Vision and Image Processing*, 37(3):261–298, 1987.