Temporal Order Recovery of the Scanned Handwriting

Abdelkarim Elbaati[1,2] , Monji Kherallah[1],Abdellatif Ennaji[2], Adel M. Alimi[1]
[1]Research Group on Intelligent Machines, REGIM,University of Sfax, Tunisia
[2]Laboratory LITIS,University of Rouen, France
{abdelkarim.elbaati,adel.alimi}@ieee.org,abdel.ennaji@univ-rouen.fr,monji.kherallah@enis.rnu.tn

Abstract

In this paper, we present a new approach to the temporal order restoration of the off-line handwriting. After the pre-processing steps of the word image, a suitable algorithm makes it possible to segment its skeleton in three types of strokes. After that, we developed a genetic algorithm GA in order to optimize the best trajectory of these segments. The repetition of a segment will be studied in a secondary algorithm so that we do not disturb the GA operations. The techniques used in GA are the selection, crossover and the mutation. The fitness function value depends on right-left direction (direction of the Arab writing), the segments repetition and angular deviation on the crossing of the occlusion stroke. To validate our approach, we tested it on the On/Off LMCA dual Arabic handwriting, the Latin IRONOFF and the off-line IFN/ENIT datasets.

1. Introduction

Few years ago, the handwriting recognition field have been challenge topics of research orientations. In this area, the on-line and the off-line handwritten script are considered. In several applications, the necessity to pass the information from an ordinary paper in a comprehensible electronic file is one of the strong motivations which push us to develop an automatic recognition system of Arabic off-line handwriting. Among these applications, we quote the automatic analysis of the administrative documents as the investigation forms, the accounts - bank checks, digitalization and safeguard of the patrimony documents, etc.
On-line handwriting arises as a mono-dimensional signal in the form of succession points whose position is function of time. Dynamic information such as the temporal order of the points as well as the speed of the pen can be used by the recognition system. For that, the performance of on line system is more encouraging than the off-line one. This interpretation encouraged us to develop an approach of extraction of a mono-dimensional signal from the off-line handwriting whose distribution of the points is identical to that on-line. This representation of the off-line handwriting enables us to unify modeling on and off-line. Works in the field of the temporal information restoration of the Arab static stroke, are few [3, 4, 1]. On the other hand, for Latin we can quote [9, 10, 2, 7].
In this paper, we present a new approach to the temporal order restoration of the off-line handwriting by using GA. GAs are a class of optimization and search methods that use randomness to avoid local extrema solutions. They are capable of an adaptive and robust search over a wide range of space topologies. GAs are distinguished from other techniques by a principal characteristic: they search in intrinsically parallel fashion from several solutions and not from a single solution [8, 11]. GA is also an iterative algorithm that depends on the generation-by-generation development of possible solutions, with selection schemes permitting the elimination of bad solutions and the replication of good ones that can be modified. There are three stages in a genetic search process: selection, crossover and mutation.

2. Description of the system

2.1. Segmentation of the skeleton

The word image, captured in level of gray from a scanner with a resolution of 300 dpi, passes by three stages of pre-processing: binarization, filteration and skeletonization. The algorithm of skeletonization is based on morphological operations of erosion. As in all the skeletonization algorithms, the skeleton can contain artefacts. For that the skeleton passes by an algorithm of filtering. Each artefact lower than 0,75 of the stroke thickness will be eliminated. The stroke thickness will be calculated by seeking the repetitive value of the vertical and horizontal projection profile on the original image.
Three types of characteristic points will be extracted from the skeleton of the tracing.

1. the end stroke point: this is the black pixel that possesses only one neighbor of the same type,

978-0-7695-3725-2/09 $25.00 © 2009 IEEE
DOI 10.1109/ICDAR.2009.266

2. the branching point: this is the black pixel that possesses three,
3. the crossing point: this is the black pixel that possesses four neighbors of the same type.

An algorithm of the skeleton follow-up permits to determine the segments of a word. These segments are classified according to three categories (figure 1):

1. segment 1: this is a stroke that is located between two end points or between an end point and a branching point (or crossing point),
2. segment 2:this is a stroke of link that is located between two branching points (or crossing) or between a branching point and a crossing point. This type of segment does not represent a contour of an occlusion,
3. segment 0: this is a stroke of link that is located between two branching points (or crossing) or between a branching point and a crossing point. It is a stroke that represents a contour of an occlusion.

This segmentation is made through a follow-up of the skeleton while departing from the characteristic points and looking every time the next neighboring point. For this reason, this segmentation allows a first organization of the points of every segment.

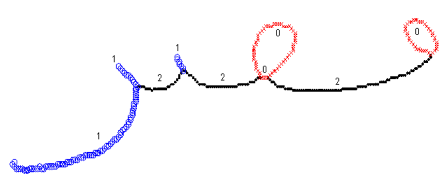

Figure 1. Segmentation of the skeleton of the word (short)

2.2. Genetic Algorithm

To restore the temporal order of the writing, the first stage consists of ordering the connected components of a word from the right to the left. In this work, we consider only the connected components which have one down-up pen. Otherwise, the writing of a connected component is done in a continuous way on the basis of a starting point and arrives at the end point without intermediaries survey-posed pen. For a related component having n segments there will be n!=1x2x.....xn possible permutations of these segments to traverse all these segments. If number n increases, the number of possibilities becomes more and more important. For that, to seek the best permutation we will use GA for each connected component. On the basis of a number of random permutations m lower than n! and by applying the genetic operations, GA converges towards the optimal solution after a certain number of generations without testing all n! permutations. The used GA is inspired from the one used to solve the problem of the commercial traveler.

2.2.1 Coding of a population

The first step in the implantation of the genetic algorithms is to create a population of initial individuals. Indeed, the genetic algorithms act on a population of individuals, and not on an isolated individual. By analogy with biology, every individual of the population is coded by a chromosome or genotype [5]. A population is, therefore, a set of chromosomes. Every chromosome codes a point of the research space. The efficiency of the genetic algorithm will depend on the choice of the coding of a chromosome. In our system, a chromosome is represented as chains of integer between 1 and the number of segments found (n). Every segment is coded by an integer between 1 and n. The coding of a segment is its order of appearance if we make a sweep from right to left (for Arabic). The position of a segment represents the position of his gravity center.
The space of research is the set of the permutations of [1,2.., n]. A point of the research space is a permutation.

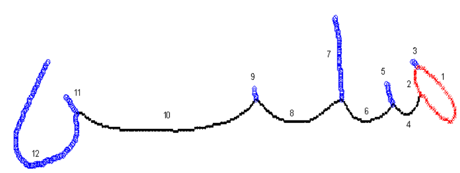

Figure 2. Coding of the segments of word of Tunisian city name (metline)

2.2.2 Initial population

The initial population is produced by an at random generation of m-1 chromosomes (m at random permutations of the n codes of the segments). To guarantee the convergence of the GA, we insert the chromosome neat [1 2 3 n]. Finally, we will have an initial population of m chromosomes.

2.2.3 Secondary algorithm

A chromosome coming from the initial population or generated by GA passes through a secondary algorithm which permits to guarantee if possible a continuity during of the segments course. To have this continuity, a segment can be:

1. Reoriented

2. Recalled (twice at maximum)

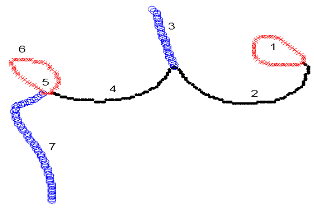

Figure 3. Coding of the segments of word (pen)

As this algorithm tries to minimize the angular variations on the occlusions during the segments reorientation.
For the example of figure 3 and for the chromosome (1 2 3 4 5 6 7) coming from the initial population or generated by GA:

1. Segment 1 is reoriented in a manner to minimize the angular deviation during the passage from 1 to 2. (figure 4)

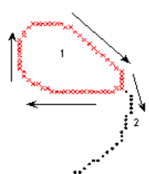

Figure 4. Reorientation of an occlusion

2. To ensure the continuity of the stroke, the segments 3 and 5 must be recalled (3 ' and 5'). The extension of the chromosome becomes: (1 2 3 3' 4 5 6 5' 7)

2.2.4 Evaluation function (Fitness)

GA seeks to find a permutation which minimizes the evaluation function value (Fitness). This is the total cost of every genetic operation. The fitness value is calculated as follow:

1. calculate the Cost of Right-Left displacement between the pixels which is named (CRL). Since Latin is written from the left to the right, this cost will be calculated conversely if Latin words will be treated,

2. calculate the Cost of the Low-High displacement between the pixels: which is named (CLH),

3. calculate the Cost of Segment Repetition (length of a course) which is named (CSR). Generally, the writer seeks to simplify and accelerate the writing stage. Thus he seeks to make the shortest way thus minimize the repititions,

4. calculate the Cost of the Angular Deviation in the crossing and branching points of the occlusions. This cost is named (CAD).

Before calculating the function of assignment of a chromosome, the secondary algorithm permits to verify the order of the pixels of the segments and to determine the segments to browse twice.

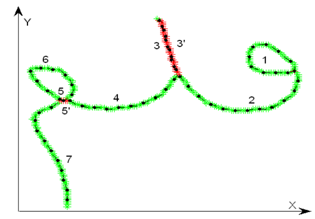

Figure 5. Example of a rebuilt signal of word (pen)

The curvature function is calculated on the connection point between the two segments by using the method developed in [9] Eq 1.

$$C(pi) = \frac{1}{m^* - k_0 + 1} \sum_{k=k_0}^{m^*} Sik \tag{1}$$

For example: for the word (pen) (figure 3), the cost of angular variation during the passage (from 4 to 6) is more important than that of (from 4 to 5) (figure 6).

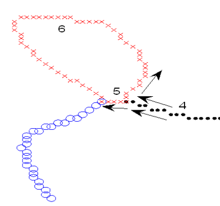

Figure 6. Angular variation in an occlusion

The fitness function is a sum of all these costs.

2.2.5 Selection

To determine which individuals are more inclined to obtain the best results, a selection is operated. We use here the method of selection by caster. One calculates initially the median value of the function of evaluation in the population Eq. 2.

$$\overline{f} = \frac{1}{m} \sum_{i=0}^{m-1} f(P_i) \tag{2}$$

Where P_i is individual i of the population and m the size of the population.The place of an individual P_i in the caster is proportional to $\frac{\bar{f}}{f(P_i)}$. Then, we select $\frac{m}{2}$ individuals for the reproduction.

2.2.6 Crossover

During this operation, two chromosomes exchange parts of their chains to give new chromosomes. Once the intermediate population has been selected, we supplement the population with the children of the intermediate population. Two combining chromosomes give rise to two other chromosomes in the following way:

1. A point of hybridization is by chance given (between 1 and the length of the chromosome n).

2. We copy in the first son the indices of the first father until the point of hybridization.

3. Then, we supplement with the indices of the second chromosome father not being already in the first son. These indices should be remained in the same order found in the second father.

4. We reiterate (2) and (3) with the same point of hybridization, but by reversing the role of the two parents.

For example, for n=8, there are the two following chromosomes which are hybridist:
C1=(2 3 5 1 4 7 6 8)
C2=(4 6 1 3 5 2 8 7)
We hybrid starting from row 4. Then we obtain the two new chromosomes:
C1'=(2 3 5 1 4 6 8 7)
C2'=(4 6 1 3 2 5 7 8)

2.2.7 Mutation

In a random way, a gene can substituted for another. One of the simplest methods in order to transfer a chromosome is to reverse the positions of two genes. However, this mutation changes much the adaptation function. If we want to reverse two non-contiguous gene i and j, the evaluation function changes much. A solution which disturbs less the individual consists in reversing all the genes between i and j.
For example:
C=(1 6 3 5 7 8 2 4)
becomes
C'=(1 6 2 8 7 5 3 4)

3. Experimentation and results

To validate this approach we have to test it on several words of the On/Off LCMA dual Arabic handwriting [6], the Latin IRONOFF and the off-line IFN/ENIT databases. The evaluation with the three databases is done visually.

Table 1. Temporal order restoration results

Database	Nbr of words	Nbr of pseudo-words	Handwritten quality	% for the pseudo-words	% for the words
IFN/ENIT	230	1000	Random	92%	72%
IFN/ENIT	100	450	Good	99%	97%
LMCA	36	86	Good	96,51%	94,44%
IRONOFF	40	87	Random	87,35%	72,5%

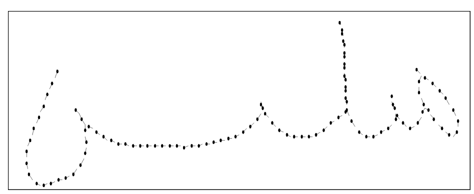

Figure 7. Restoration of the temporal order of the Tunisian city names (metline) (from IFN/ENIT database)

In the example presented in figures 2 and 7 , the GA gives the chromosome: P= (1 3 2 4 5 6 7 8 9 10 11 12) After the passage by the secondary algorithm this chromosome becomes: PF= (1 3 3' 2 4 5 5' 6 7 7' 8 9 9' 10 11 11' 12).

The performance evaluation of the system carried out on a set of words of the On/Off LMCA dual Arabic handwriting, the Latin IRONOFF and the off-line IFN/ENIT databases give a following encouraging preliminary result. The comparison of the rates of the temporal order rebuilding of various approaches is difficult because the difference in the methods of test and the used databases. The experimental evaluation shows several advantages of this approach:

1. Use of an intelligent approach contrary to the majority of preceding work for Arabic [3, 4, 1], and for Latin [9, 10] which use approaches,

2. the nature of the skeleton segmentation approach facilitates the temporal order restoration stage since it offers a first organization of the pixels belonging to the same segment,

3. the approach does not impose the knowledge of the starting segment.

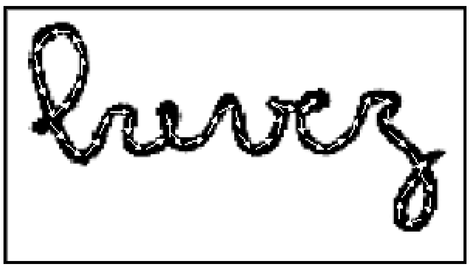

Figure 8. Restoration of the temporal order of the off-line word (buvez) (from IRONOFF database)

4. Conclusion and prospects

In this paper, we have presented a new approach to rebuild the temporal order of the off-line handwritten script. The first stage consists in segmenting the skeleton of a word in three types of the segments by using a method of follow-up. These segments are applied to the input of a genetic algorithm in order to optimize the course of the word by calculating a fitness function. Before calculation the fitness of a given course, a secondary algorithm makes it possible to find the segments to be traversed twice. A validation of the approach on words of the On/Off dual Arabic handwriting LMCA, the Latin IRONOFF and the off-line IFN/ENIT databases shows the effectiveness of our approach. As a perspective of our work is to improve the stage of the stroke extraction and correction. The restoration of dynamic information (temporal order) enables us to profit from the modeling of the on-line writing to improve the stage of recognition. Our work quoted in [4] explains well this passage of the word image towards on-line modeling. Thus our system of the temporal order restoration is promising for any application of off-line handwritten documents recognition such as the recognition of signatures, automatic tri of mail, the automatic treatment of the administrative files, investigations forms, accounts - bank checks, digitalization and saving of written heritage . etc.

Acknowledgments

The authors would like to acknowledge the financial support of this work by grants from the General Direction of Scientific Research and Technological Renovation (DGRST), Tunisia, under the ARUB program 01/UR/11/02. Also, this work was developed in part, when the first author did a scientific stay in the LITIS laboratory (Rouen, France).

References

[1] I. Abuhaiba, M. Holt, and S. Datta. Straight line approximation and 1d representation of off-line handwritten text. In *Picture and Vision Computing*, pages 648–659, 1994.

[2] G. Boccignone, A. Chianese, L. Cordella, and A. Marcelli. Recovering dynamic information from static handwriting. In *Pattern Recognition*, volume 26, pages 409–418, 1993.

[3] A. Elbaati, M. Charfi, A. Alimi, and A. Ennaji. Extraction and organization of the useful signal of arabic off line handwritten tracing. *ACIDCA-ICMI*, 9:715–722, 2005.

[4] A. Elbaati, M. Kherallah, A. M. Alimi, and A. Ennaji. De l'Hors-Ligne Vers un Système de Reconnaissance En-Ligne : Application à la Modélisation de l'Écriture Arabe Manuscrite Ancienne. In *Semaine du document numrique, SDN (ANAGRAM)*, 2006.

[5] J. Holland. Adaptation in natural and artificial systems. In *MIT Press*, 1975.

[6] M. Kherallah, L. Hadded, A. Mitiche, and A. M. Alimi. On-line recognition of handwritten digits based on trajectory and velocity modeling. In *International journal of Pattern Recognition Letter*, volume 29, pages 580–594, 2007.

[7] P. Lallican, C. Gaudin, and S. Knerr. From off-line to on-line handwriting recognition. In *in Inter. Workshop on Frontiers in Handwriting Recognition (IWFHR)*, pages 303–312, 2000.

[8] F. Pernkopf and D. Bouchaffra. Genetic-based em algorithm for learning gaussian mixture models. In *IEEE Trans Pattern Anal Mach Intell*, volume 27, pages 1344–8, 2005.

[9] R. Plamondon and C. Privitera. The segmentation of cursive handwriting : an approach based on off-line recovery of the motor-temporal information. In *IEEE Transactions on Image Processing*, volume 8, pages 80–91, 1999.

[10] L. Rousseau, E. Anquetil, and J. Camillerapp. Reconstitution du parcours du trac manuscrit hors-ligne de caractres isols. In *Colloque International Francophone de lEcrit et du Document CIFED04*, 2004.

[11] G. Schneider, H. Wersing, B. Sendhoff, and Al. Evolutionary optimization of a hierarchical object recognition model. In *IEEE Trans Syst Man Cybern B Cybern*, volume 35, pages 426–37, 2005.