On Line Isolated Characters Recognition Using Dynamic Bayesian Networks

Redouane Tlemsani[1] and Abdelkader Benyettou[2]
[1]Departement of Transmission, National Institute of Telecommunications, Information Technologies and Communication of Oran - INTTIC, Algeria
[2]Departement of Computer Sciences, University of Sciences and Technologies of Oran USTO, Algeria

Abstract: *In this paper, our system is a Markovien system which we can see it like a Dynamic Bayesian Networks. One of the major interests of these systems resides in the complete training of the models (topology and parameters) starting from training data. The representation of knowledge bases on description, by graphs, relations of causality existing between the variables defining the field of study. The theory of Dynamic Bayesian Networks is a generalization of the Bayesians networks to the dynamic processes. Our objective amounts finding the better structure which represents the relationships (dependencies) between the variables of a dynamic bayesian network. In applications in pattern recognition, one will carry out the fixing of the structure which obliges us to admit some strong assumptions (for example independence between some variables).*

Keywords: O*n line isolated character recognition, pattern recognition, and dynamic bayesian network.*

Received July 4, 2009; Accepted October 31, 2009

1. Introduction

Since the Sixties, the Man seeks "to learn how to read" for computers. A task of recognition is difficult for the isolated handwritten characters because their forms are varied compared with the printed characters. The on line recognition makes it possible to interpret a writing represented by the pen trajectory.

This technique is in particular used in the electronic message minders of type Personal Digital Agenda (PDA). An electronic shelf and a special pen are necessary. The signal is collected in real time. It consist the succession of point's co-ordinates, corresponding to the pen position with time regular intervals. Indeed, the on line signal contains dynamic information absent in the off line signals, such as the order in which the characters were formed, their direction, the pen down and pen up position.

So that the isolated character recognition is strongly precise, it is significant to as structure characters model the usually as possible. In this work we consider that a character is composed of strokes and even their relationships were kept. The strokes are the conceptual elements and their space relations are conceptually significant and which are usually robust against geometrical and significant variations for the distinctive characters of the similar forms.

A bayesian network can model dependencies between several random variables in a probabilistic and graphic way representation.

2. Dynamic Bayesian Networks

The Dynamic Bayesian Networks (DBN) prolongs the representation of Bayesian Networks (BN) to the dynamic processes. A DBN codes the Jointed Probability Distribution (JPD) of time evolution X[t]={X 1[t]...., X N [T]} of variables. In other words, it represents the belief about the possible trajectory of the dynamic process X[t]. After a similar notation with the static representation of BN, the JPD for a finished interval of time [1,T] is factorized like:

$$p\left(X[1],\ldots.,X[T]\right)=\prod_{t=1}^{T}\prod_{i=1}^{n}P\left(X_i[t]\middle|\Pi_i[t]\right) \qquad (1)$$

Where $\prod_i[t]$ the parents of Xi[t] in the graph indicate structure of DBN.

The graphic structure of a DBN can be looked like concatenation of several dependent static BNs with the temporal arcs. We call each one of these static networks a section of time (a section of time is defined like collection of the set of X[t] in only one time T instantaneous and their parents associated $\prod[t]$ in the structure of graph) with DBN In the most general case, if no pretension are imposed on the fundamental dynamic process, the structure of the graph and the numerical parameterization of a DBN can be different for each time out in sections. In this case the DBN is regarded as BN (static) with $T\times n$ variables and the coding of the JPD can be extremely complex.

2.1. Representation

In the literature the representation of DBN generally is employed for the first stationary order Markov processes. For this case, Friedman and others described a representation simplified in terms of two static head BNs definite above the variables of a simple time section [2]. The principal representation is based on the pretension of stationnarity which implies that the structure and the parameters of DBN repeat. The JPD is coded by using a first network and an unrolled transition network.

The initial network codes the irregular structure in the border and indicates the initial states of surplus of X[1] distribution. The transition network codes the invariable probability transition time. $P\left(X[t+1] \mid X[t]\right)$ The JPD for a finished time interval is obtained by unrolling the transition network for a sufficient number of times sections. The mechanism of unfolding is composed to present a set of variables for each time out in sections and to fold up the structure and the parameters of transition network on these variables. Rearranging the limits JPD is factorized above the networks initial and transition like:

$$p(X[1],\ldots\ldots,X[T]) = P_{B_I}\left(X[1]\right)\prod_{t=2}^{T-1} P_{B\rightarrow}\left(X_i[t] \mid X[t-1]\right) \tag{2}$$

Where $P_{B_I}(.)$ and $P_{B\rightarrow}(.)$ are the densities of probability coded by the initial and transition networks, respectively.

2.2. Inference in DBN

The problem of inference in DBNs is similar to the problem of BN stock inference as desired the quantity is the posterior marginal distribution of a set of hidden variables indicated an order of the observations (updated of belief): $P\left(X_h[t] \mid X_0[1],\ldots,X_0[\tau]\right)$. Where $X[t]=\{X_h[t], X_0[t]\}$ is a set of time evolution variables in which $X_o[t]$ and $X_h[t]$ indicate observed variables and hidden, respectively. Inference of the time series generally under the name filtering $(\tau=1)$, smoothing $(\tau>1)$ and the forecast $(\tau<1)$ according to the time window of observation used in calculations.

A direct approach to imply probabilities in a DBN, is to build an enormous static BN for the desired number of time sections and then to employ the general algorithms of inference for static BNs. However, this requires that the end of about a time be known a priori. Moreover, the data-processing complexity of this approach can extremely require (particularly in terms of memory). Consequently in general, the DBN inference is carried out by using the recursive operators who update the belief state of DBN while the new observations become available. The principle is similar to the message passing algorithm for static BNs. The idea is with the messages defined on a Markov cover of the variables which D-separates the past from the future and employs a process towards the procedure forward- backward to distribute all the obviousness along the DBN [2,3,4]. This technical requires only one time window of the variables to be maintained in the memory. These algorithms are indeed generalization of the algorithm (Baum-Welch) towards forward-backward well-known [5] in special HMMs and cases JLO algorithm [6].

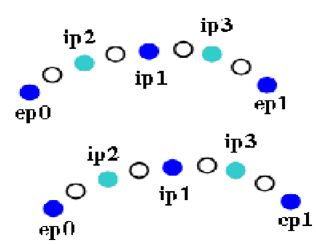
3. Modelling

In this part, we consider that a character is composed of strokes and their relationships. The strokes are direct elementary lines or almost rights which have directions distinct from the lines connected in the writing order. The relationships of the strokes indicate the dependencies of the positions between the strokes obtain an influence on the others strokes.

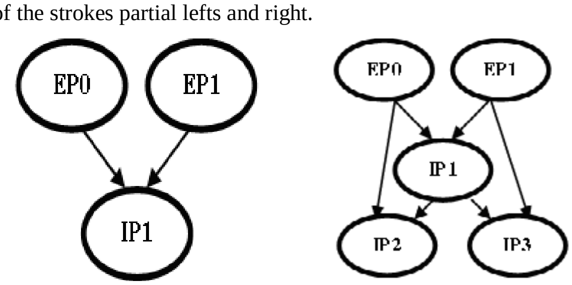
3.1 Static Model

An example of stroke is composed of points. Consequently, a stroke model is composed of point models with their relationships, called "Within Stroke Relationships" (ISRs).

Figure 1 show the recursive example of stroke construction. To the first recursive iteration (D=1), IP1 is added to median model points of all the stroke examples. It has the WSR of the final points (arcs of EP0 and EP1 with IP1). To the second recursive iteration (D=2), IP2 and IP3 are added for median points of the strokes partial lifts and right-hands side, respectively. Moreover, they have the WSR of the final points of the partial strokes. Figue.1 (c) is the prolonged model of stroke.



(a) Example for ip1's: median point of stroke ip2s et ip3's: those of the strokes partial lefts and right.



(b) Strok model depth d =1. (c) Stroke Model depth d=2.

Figure 1. The recursive construction of a stroke model.

With this recursive process, a model of stroke can as many have point models according to needs. In this part, the recursively depth d=3 is selected for all the stroke models.

It is worth the sorrow to note that the models of point to great recursively depths, do not incur the problem of non adequate model. Because when the depth is large, the partial strokes become much shorter and linear. Consequently, ISRs become much stronger and the joined probabilities of the additional point models obtain more close the probability of only one. The joined probability is obtained from those of the models of point. Let us suppose that a model S has the depth D and an example of stroke is points length T: O(1)…,O (T). To match, the example of stroke is periodically taken in the 2d-1 median points. They are indicated like IP1, IP2...., IP2d- 1 according to the order of the process of recursive taking away.

Then, IPi examples of point are matched with the models IPi of point. The joined probability is calculated as follows by the local Markov property of the conditional probabilities in the bayesian networks:

$$P\left(S=O(1),...,O(t)\right)=P\begin{pmatrix}EP_0=O(1),EP_1=O(t),\\ IP_1=ip_1,...,IP_{2^d-1}=ip_{2^d-1}\end{pmatrix}$$
$$=P\left(EP_0=O(1)\right)P\left(EP_1=O(t)\right)\times\prod_{i=1}^{2^d-1}P\left(IP_i=ip_i\setminus pa\left(IP_i\right)\right) \quad (3)$$

Where the pa(IPi) is the configuration of the nodes parents which the arcs of dependence like in IPi.

3.2 Dynamic Model

An example of character is composed of the strokes. Moreover, the close connections exist between them. Consequently, a character model is composed of the stroke models with their relationships, called "Inter Stroke Relationships (ISRs)".
In Figue.2, EP0 is the first point model written in a character.

The point models of first stroke are written in the order of IP1,2, IP1,1, IP1,3. Then, the models of point of the second strokes are written in the order of EP1, IP2,2, IP2,1, IP2,3. Alternatively, the following strokes are written in the same way. In conclusion, EPN is the last model of point written in a character [7,8,9].

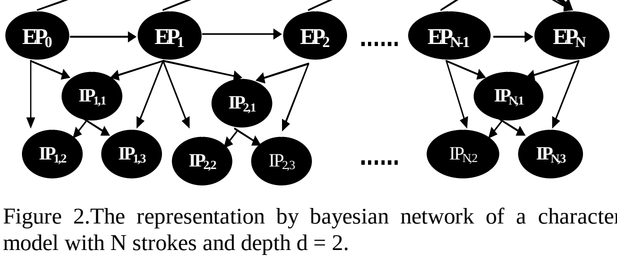


Figure 2.The representation by bayesian network of a character model with N strokes and depth d = 2.

The model of probability of a character is calculated by the enumeration of all the possible segmentations of stroke. Let us suppose that a model BN of character has N stroke model and an entry of character with T points: O(1)…,O(T). Since the entry does not have the information of border, various segmentations are possible. One poses an example of stroke segmentation by γ = (t0, T1…, tN), t0 = 1<t<… <tN = T, and the set of totality by Γ. Then the probability model of a character is given as follows:

$$\begin{aligned}&P\left(O(1),...,O(t)\setminus BN\right)\\&=\sum_{\gamma=(t_0,...,t_N)\in\Gamma}P\left(S_1=O(t_0,t_1),...,S_N=O(t_{N-1},t_N)\right)\\&=\sum_{\gamma\in\Gamma}\prod_{i=1}^{N}P\left(S_i=O(t_{i-1},t_i)\setminus S_1=O(t_0,t_1),...,S_{i-1}=O(t_{i-2},t_{i-1})\right)\\&=\sum_{\gamma\in\Gamma}\prod_{i=1}^{N}P\left(S_i=O(t_{i-1},t_i)\setminus EP_0=O(t_0),...,EP_{i-1}=O(t_{i-1})\right)\end{aligned} \quad (4)$$

Where O(ti, tj) = O(ti), O(ti+1)…,O(tj). The joined probability given by preceding strokes is calculated as follows:

$$P\left(S_i=O(t_{i-1},t_i)\setminus EP_0=O(t_0),...,EP_{i-1}=O(t_{i-1})\right)=$$
$$\begin{cases}P\left(EP_i=O(t_i)\setminus O(t_0),...,O(t_{i-1})\right)\\ \prod_{j=1}^{2^d-1}P\left(IP_{i,j}=ip_{i,j}\left(O(t_{i-1},t_i)\right)\setminus pa\left(IP_{i,j}\right)\right) \text{ if } i>1,\\ P\left(EP_0=O(t_0)\right)P\left(EP_1=O(t_1)\setminus O(t_0)\right)\\ \prod_{j=1}^{2^d-1}P\left(IP_{i,j}=ip_{i,j}\left(O(t_{i-1},t_i)\right)\setminus pa\left(IP_{i,j}\right)\right) \text{ if } i=1,\end{cases} \quad (5)$$

Where ipi, j(O(ti-1, Ti)) are the jst point sample of O(ti-1, ti).En subtituant Eq.(4) for Eq. (5), the probability of the model it is only one product of the joined probabilities of EPs and IPS:

$$\begin{aligned}&P\left(O(1),...,O(t)\setminus BN\right)=\\&\sum_{\gamma\in\Gamma}\prod_{i=0}^{N}P\left(EP_i=O(t_i)\setminus O(t_0),...,O(t_{i-1})\right)\\&\times\prod_{i=1}^{N}\prod_{j=1}^{2^d-1}P\left(IP_{i,j}=ip_{i,j}\left(O(t_{i-1},t_i)\right)\setminus pa\left(IP_{i,j}\right)\right)\end{aligned} \quad (6)$$

The joined probabilities of EPs can be interpreted by probabilities of the stroke positions total and those of IPs with probabilities of the local stroke forms.

4. Recognition and Training

4.1 Recognition algorithm

A handwritten character is identified by finding the model of character which produces the highest posterior probability given entry. When the Iist model of character is indicated BNi and the points entrance as O(1)…, O(T), then the recognition problem can be formulated as follows:

$$\arg\max_i P\left(BN_i \setminus O(1),...,O(T)\right)$$
$$= \arg\max_i \frac{P(BN_i)\,P\left(O(1),...,O(T)\setminus BN_i\right)}{P\left(O(1),...,O(T)\right)}$$
$$= \arg\max_i P(BN_i)\,P\left(O(1),...,O(T)\setminus BN_i\right) \quad (7)$$

The model character probability is described previously. To calculate it, all possible stroke segmentations Γ are considered. To prevent the time exponential complexity, we suppose that it can be brought closer by the character joined probability of the most probable segmentation γ^* in Γ as follows:

$$P(O(1),...,O(T)\setminus BN_i) \approx \max_{\gamma\in\Gamma} P(S_1 = O(t_0,t_1),...,S_N = O(t_{N-1},t_N)) \quad (8)$$

To carry out the probability calculation of the handy model in time, we need one pretension for research of γ^*. By matching a stroke, all the possible segmentations of its strokes preceding should be considered because dependencies of inter-stroke. For the simplicity of research, we suppose that the joined probability of a stroke is highest with the most probable configuration of the previous strokes. Then, the "dynamic programming search algorithm" can as follows be adopted:[10,11,12].

S_i **:** *$i^{ième}$ stroke model*
$\gamma_i(t)$: most prob segmentation where
S1,….,Si et O(1,t) are matched.
$\delta_i(t)$: JPD of $\gamma_i(t)$.
Initialization
$\delta_0(1)=1$, $\gamma_0(1)= \{\}$
Stroke matching
for t=2 to T
for i=1 to N
$\delta_i(t) = \max_{1\le b<t} P(S_i = O(b,t) \setminus \gamma_{i-1}(b))\,.\,\delta_{i-1}(b)$
$b^* = \arg\max_{1\le b<t} P(S_i = O(b,t) \setminus \gamma_{i-1}(b))\,.\,\delta_{i-1}(b)$
$\gamma_i(t) = \gamma_{i-1}(b^*) \cup \{t\}$
end
end
Probability of the character model
$P(O(1),….,O(T)\setminus BN_i) \approx \delta_N(T)$

Figure 3. Recognition algorithm

4.2 Training algorithm

In this part, the structure of dependence is determined by a original model starting from knowledge a priori and the experiments. The recursively depth of the stroke models is selected equal to Three (d=3). The number of models is given starting from the typical number of stroke in the character. The conditional parameters of probability are formed by training data. They are the linear regression matrixes W'S (W=[wi,J]) and covariance's Σ' S for the points models. If all point models are matched the point following the example, then they can be estimated starting from the conventional statistical algorithms of regression with the maximum object of maximum probability ML "likelihood" [13]. Let us suppose that the point P depends on P1…, PK and there are N training samples. One notes the iist sample of P have p(i) and the values of dependent variable by z(i) = [x(i)1, y(i)1…, x(i)k, y(i)k, 1]. Then, they are estimated as follows: [13]

$$\Sigma = \frac{1}{N}\sum_{i=1}^{N} p^{(i)}\left(p^{(i)}\right)^T - \frac{1}{N} W \sum_{i=1}^{N} z^{(i)}\left(p^{(i)}\right)^T$$
$$W = \left(\sum_{i=1}^{N} p^{(i)}\left(z^{(i)}\right)^T\right)\left(\sum_{i=1}^{N} z^{(i)}\left(z^{(i)}\right)^T\right)^{-1} \quad (10)$$

During the training of the character model, the Re-estimate of the parameters and it required of the most probable segmentation in strokes γ^* is repeated alternatively[11,12,13].

This approach is similar to the training algorithm EM (Expectation Maximization). Being given the parameters (W and Σ), γ^* is updated. Then, with the news γ^*, the parameters Re-are estimated. The detailed algorithm is as follows:

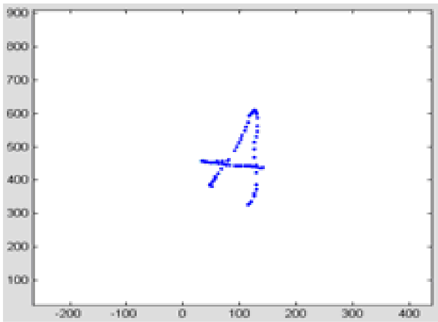
- Step 1: To initialize the character model with the initial data (part of the examples of the manually segmented strokes).
- Step 2: To seek the most probable segmentation γ^* of the totality of the characters of training not segmented by using the algorithm of required the previous one.
- Step 3: To estimate the parameters (W and Σ) on the examples partitioned by γ^*.
- Step 4: To repeat stages 2 and 3 until the sum of probabilities of the model will not change any more (stability).

5. Experiences, Results and Analyses

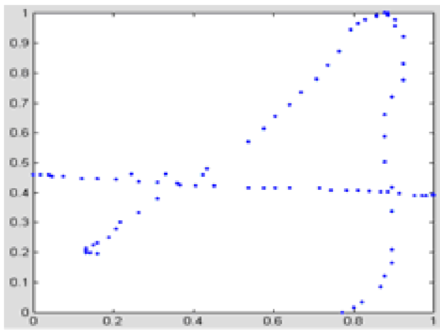
UNIPEN data base is the reference index for the development and the comparison of writing recognition systems. We use here the part of this data base containing the isolated characters, digits, small letters and capital letters. This base contains layouts of more than 200 script writers. The difficulty in this base is due mainly to the number of script writers and thus to the many allograph which they employ.

5.1 Manual segmentation

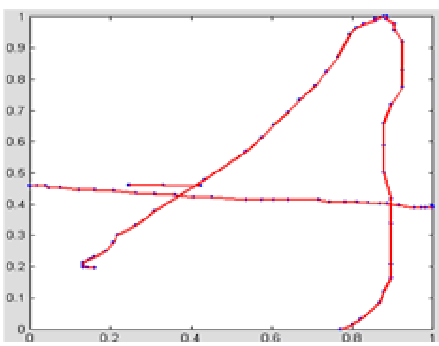
The aquisition made dynamically using a graphics tablet is had to digitalize. This one has a resolution specifies and samples at a speed selected writing. A time of adaptation is necessary to the script writer to be able to write has little close correctly despite everything for its treatment the character will have to be segments in trace i.e. in strokes elementary. The figure above shows the various changes implemented and applied for example to character "A".



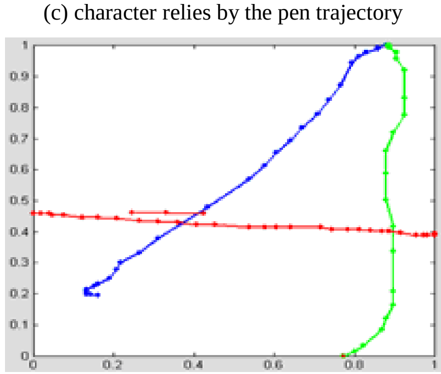
(a) Original character.



(b) normalize character.



(c) character relies by the pen trajectory



(d) Segmented character.

Figure 4. Treatment and segmentation of character "A .

5.2 DBN Experience

For our experiments one implemented and checking our bayesian model on a set of letters of A to H of the Latin alphabet. Table 1 shows the confusion matrix of the recognition rates and they are relatively significant rates on such under corpus and it is an effectiveness returns to the safeguarded elements space and Gaussian probabilities of each point model. The total rate of recognition attempt 70,24%.

The results show rates increased for letters A, B and D. of the rates relatively acceptable for the letters E and F and of the rate lower and equal to 50 for the letters C, G and H.

Table 1. Bayesian Model: Recognition Rates of the letters A to H.

%	A	B	C	D	E	F	G	H
A	98,33	-	-	-	-	01,66	-	-
B	-	97,50	-	-	-	02,50	-	-
C	-	-	44	-	44	02	-	-
D	-	-	-	100	-	-	-	-
E	-	-	-	-	70	20	-	10
F	-	-	-	-	28,15	69,55	-	2,21
G	03,22	-	-	-	25,81	38,71	29,03	3,22
H	08,88	-	-	-	11,11	26,66	-	53,33

5.3 Neuronal verifier

The method implemented in this article is a probabilistic model which is summarized with the concept of the dynamic bayesian networks and after having obtained γ^* the vector of probabilities by the dynamic programming algorithm, instead of taking the maximum value, the researchers in pattern recognition often use a verifier operator on this level to adapt the system. We prefer integrate the neural networks like tool for checking of the forms.

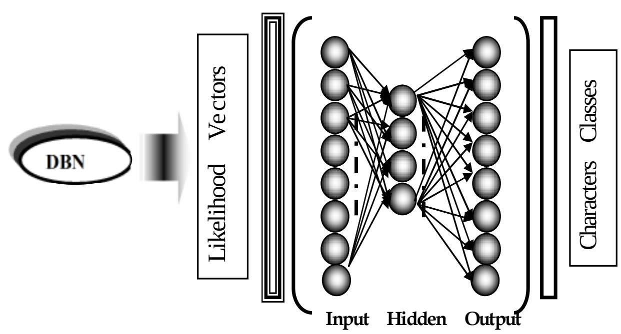


Figure 5. Neuronal verifier operator.

Neuronal architecture used is shown in the figure above (8 neurons in entry, 4 in the hidden layer and 8 neurons in the layer of exit), in entry assignment of the vectors of probabilities obtained by the dynamic network bayésien. At exit a binary vector indicates the class of associated nature.

In figure 6, one clearly notices the convergence of the average quadratic error obtained starting from training neuronal.

Table 2 shows the confusion matrix of the recognition rates for the bayesian model with a neuronal verifier operator and one notices on the level of some characters such as A and B there is a total

performance, but there are also non given characters such as 55% of the letter C and 65% of the letter E and that return to the probabilistic mechanism of the bayesian networks.
The total rate of recognition is 66.87%.

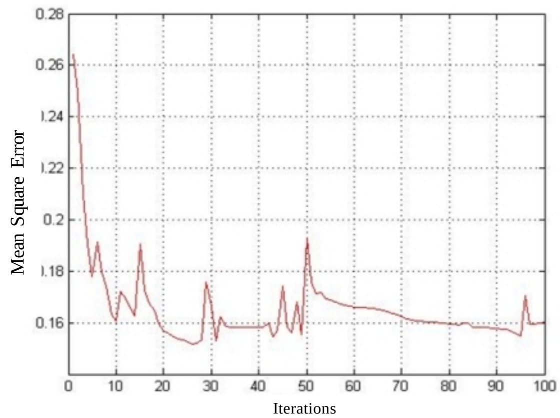


Figure 6. Graph of the error evolution by iteration

Table 2. Neuronal verifier operator: Recognition Rate of the letters of A to H.

%	A	B	C	D	E	F	G	H
A	100	-	-	-	-	-	-	-
B	-	100	-	-	-	-	-	-
C	-	-	15	-	-	-	85	-
D	-	-	-	100	-	-	-	-
E	-	-	-	-	-	25	65	10
F	-	-	-	-	-	70	30	-
G		-	-	-	-	5	90	-
H	20	-	-	-	-	10	10	60

5.4 Novel algorithm using segmentation by points parts

The idea of this algorithm is to divide the writing signal in sections of points, whereas this algorithm gives a segmentation to the each level section of points, the old algorithm consists in assigning to the level of each point in the signal, a segmentation of this last, which is very slow for a on line recognition system. Two approaches:

- Segmentation using Static part of points: One fixes a number of static points for all the observations, (example: the total number of the points of the signal T=48, the beach P=3, old algorithm DRA buckles 47 times to give the final segmentation whereas the new loop requires that the number ((T/P)+1) is rounded with the higher close entireties)
- Segmentation using Dynamic part of points: One fixes here a percentage of points instead of a fixed number for all the observations, thus the number of points of the beaches changes according to the number T total of the points of one observation. (Example: the total number of the points of the T=48 signal, percentage of the Pr=5% points then beach P=((T . Pr)/100), this number is rounded with the higher close entireties). The idea of this new concept is summarized in figure 7.

One implemented three approaches (DRA, DRA by beach of static points and DRA by beach of dynamic points) applied to the set of the characters of A to H and at exit one obtained the time of segmentation of the observation indicated (see table III). The result shows effectively that the method of segmentation based on the beach of static points is faster compared to old algorithm DRA. Also this approach by beach of points showed a better overall noticed speed. One can show this effectiveness because of the basic concept which gives a possibility of partition by a set of points and not a single point.

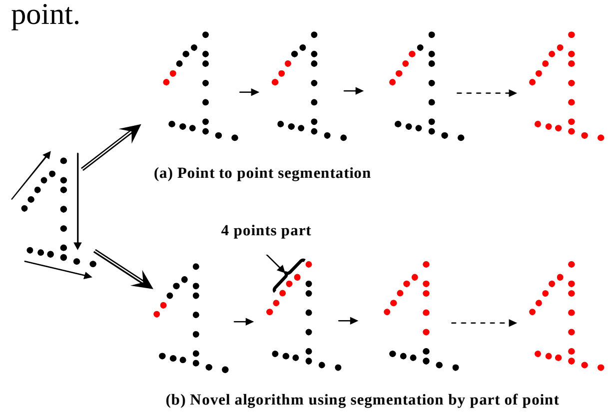


Figure 7. Dynamic Research algorithm (DRA) by part of points.

Table 3. Segmentation observations time for the three algorithms.

	DRA	Static DRA	Dynamic DRA
« A »	93.015 s	4.985 s	1.36 s
« B »	188.468 s	9.266 s	1.734 s
« C »	37.797 s	2.329 s	1.407 s
« D »	22.468 s	1.625 s	1.172 s
« E »	77.718 s	4.235 s	1.735 s
« F »	32.078 s	2.063 s	1.641 s
« G »	71.485 s	3.766 s	1.781 s
« H »	45.203 s	2.375 s	1.375 s
Means	71.029 s	3.830 s	1.526 s

6. Conclusion

To facilitate the use of the computers more, new approaches were proposed and create. They do not use any more the keyboard or the mouse but a stylet connected to an electronic TABLET.

In handwritten recognition there are two distinct recognitions, with problems and solutions different: on line and off line handwriting recognition but there are approaches combined between the two as in [14].

So that the character recognition isolated is strongly precise, it is significant to modelling the characters structure usually as possible. In this work we considered that a character is composed of strokes and even their space relationships were kept. The segmentation in stroke is not single [15, 16] but it gave effectiveness.

The use of the graphic models, such as the dynamic bayesian networks, us a made it possible to effectively treat the characters isolated set by keeping their spatial information and the writing order from each character.

The Bayesian Networks are not the only[17,18,19] used approaches in this research orientation and one can quote others at base modeling but the points are modelled by the conditional probability and distributions for the positions of pen (pen down/pen up) which specify relative information of points.

The dependencies modeling are a bayesian formalism which presented advantage of keeping spatial information between two models of point with a conditional probability of Gaussian distribution.

The goal of our work was to conceive and carry out an on line automatic recognition system of isolated characters without passing to the recognition words and sentences which can be treated soon in future work.

The Arab recognition manuscript also interests us[20,21,22] and creation Arabic data base is one of our major research objectives.

References

[1] Deviren M., Dynamic Bayesian networks for speech recognition. In Proceedings of AAAI 2002, SIGART/AAAI Doctoral Consortium, Edmonton, Canada, 2002.

[2] Zweig G., Speech Recognition with Dynamic Bayesian Networks. PhD thesis, University of California, Berkeley, Spring 1998.

[3] Friedman N., Murphy K, and Russell S. Learning the structure of dynamic probabilistic networks. In UAI'98, Madison, Wisconsin, 1998.

[4] Murphy K.P, Dynamic Bayesian Networks: Representation, Inference and Learning. PhD thesis, UC Berkeley, Computer Science Division, 2002.

[5] Kjaerulff U., A computational scheme for reasoning in dynamic probabilistic networks. In Proceedings of the Eighth Conference on Uncertainty in Artificial Intelligence, pages 121-129, San Mateo, 1992. Morgan Kaufmann.

[6] Rabiner L.R and Schafer R.W., Digital Processing of Speech Signals. Prentice Hall, Englewood Cliffs, NJ, USA, 1978.

[7] Jensen F.V., Lauritzen S.L., and Olesen K.G. Bayesian updating in recursive graphical models by local computations. Computational Statistics and Data Analysis, 1990.

[8] Nathan K., Beigi H., Subrahmonia J, Clary G.J, Maruyama H., Real-time on-line unconstrained handwriting recognition using statistical methods, Proceedings of IEEE ICASSP, Detroit, USA, Vol. 4, 1995, pp. 2619-2622.

[9] Starner T., Makhoul J., Schwartz R, Chou G, On-line cursive handwriting recognition using speech recognition methods, Proceedings of IEEE ICASSP, Adelaide, Australia, Vol. 5, 1994, pp. 125-128.

[10] Sin B., Kim J., Ligature modeling for online cursive script recognition, IEEE Trans. Pattern Anal. Mach. Intell. 19 (6) (1997) 623-633.

[11] Kim H., Kim J., Hierarchical random graph representation of handwritten characters and its application to Hangul recognition, Pattern Recognition 34 (2) (2001) 187-201.

[12] Liu C., Kim I., Kim J., Model-based stroke extraction and matching for handwritten Chinese character recognition, Pattern Recognition 34 (12) (2001) 2339-2352.

[13] Kim I., and Kim J., Statistical utilization of structural neighborhood information for oriental character recognition, Proceedings of the Fourth IAPR International Workshop on Document Analysis Systems, Rio de Janeiro, Brazil, 2000, pp. 303-312.

[14] Vinciarelli, A., Perrone, M. Combining Online and Offline Handwriting Recognition. Proceedings of the Seventh International Conference on Document Analysis and Recognition, (2003) 844-848

[15] De Stefano C., Guadagno G. and Marcelli A. A Saliency-Based Segmentation Method for On-Line Cursive Handwriting, International Journal of Pattern Recognition and Artificial Intelligence, vol.18, no. 6 (2004), 1139-1156.

[16] Elanwar R., Rashwan M., and Mashali S., Simultaneous Segmentation and Recognition of Arabic Characters in an Unconstrained On-Line Cursive Handwritten document. Proceedings of World Academy of Science, Engineering and Technology Vol 23 August 2007 Issn 1307-6884

[17] El-Hajj R., Likforman-Sulem L., Mokbel C., Arabic Handwriting Recognition Using Baseline Dependent Features and Hidden Markov Modeling", Proc. of ICDAR 2005, Seoul, Korea, August 29.-September 1.2005

[18] Touj S. and Ben Amara N., Arabic Handwritten Words Recognition based on a Planar Hidden Markov Model, The International Arab Journal of Information Technology, vol. 2, no. 4, October 2005.

[19] Al-Taani A., Maen H., Recognition of On-line Handwritten Arabic Digits Using Structural Features and Transition Network. Informatica journal Vol 32 (2008) 275–281

[20] Mezghani N., Mitiche A. and Cheriet M., A New Representation of Shape and its Use for Superior Performance in OnLine Arabic Character Recognition by an Associative Memory,

International Journal on Document Analysis and Recognition, Vol. 7, No. 4, pp. 201-210, 2005.

[21] Nourouzian E., Mezghani N., Mitiche A., and Robert Johnson B., Online Persian/Arabic character recognition by polynomial representation and a Kohonen Network. IEEE International conference on pattern recognition, ICPR. Singapour 2006.

[22] Mezghani N., Mitiche A., and Cheriet M., Bayes classification of online Arabic characters by Gibbs modelling of class conditional densities", IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol 30, No. 7, pp. 1121-1131, 2008.

Redouane Tlemsani is actually preparing a PhD in Arabic Handwriting Recognition, young researcher in pattern recognition and artificial intelligence and Master in data processing and computer sciences from the university of Sciences and Technologies of Oran USTO, Algeria. He is a member of the research team of laboratory SIMPA (Speech, Image Signal) since 2002.

He also teaches modules at both Engineering and LMD levels in computer science and software engineering.

He is actually at the National institute of Telecommunications, Information Technologies and Communication of Oran - INTTIC, Algeria.

Abdelkader Benyettou received the engineering degree in 1982 from the Institute of Telecommunications of Oran and the MSc degree in 1986 from the University of Sciences and Technology of Oran-USTO, Algeria. In 1987, he joined the Computer Sciences Research Center of Nancy, France, where he worked until 1991 on Arabic speech recognition by expert systems (ARABEX) and received the PhD in electrical engineering in 1993 from the USTOran University. From 1988 throught 1990, he has been an assistant Professor in the department of Computer Sciences, Metz University, and Nancy-I University. He is actually professor at USTOran University since 2003. He is currently a researcher director of the Signal-Speech-Image–SIMPA Laboratory, department of Computer Sciences, Faculty of sciences, USTOran, since 2002. His current research interests are in the area of speech and image processing, automatic speech recognition, neural networks, artificial immune systems, genetic algorithms, neuro-computing, machine learning, neuro-fuzzy logic, handwriting recognition, electronic/electrical engineering, signal and system engineering.