http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 5 (5): 958-963, 2006
ISSN 1812-5638

A Neuro-Euclidian Approach to Handwritten Word Recognition

Mohamed Senouci, Abdelkarim Liazid, Hadj Ali Beghdadi and Djilali Benhamamouch
LTE Laboratory, ENSET-Oran, B.P. 1523 El Mnaouer 31000-Oran, Algeria

Abstract: This study addresses a segmentation technique to handwritten word recognition. This technique implements an algorithm based on analytical approach. It uses a letter sweeping procedure with a step equal to the Euclidian distance between an established reference index and the entity (the alphabet letters). Then a dissociation of this entity is achieved when this distance reached a rate of 80%. Experimenting our segmentation technique gives a rate of 81.05% of recognition whole. A neuronal multi-layer perceptron classifier confirms the extracted segment. This procedure is successively repeated from the beginning until the end of the word. A concatenation is finally used to the word reconstitution.

Key words: Pattern recognition, segmentation, neurons networks, handwritten word

INTRODUCTION

The pattern recognition is an artificial intelligence domain which treats various problems of the physical objects identification. Thus, the pattern recognition deals with man-machine communication framework. As an interesting application of pattern recognition is the handwritten word recognition. Its aim is to transform a handwritten text into a printed text (Anquetil, 1997). The letters have an infinity representations hence the handwritten word recognition is not obvious. Effectively today, there are several applications of the handwritten word recognition. We can cite for instance the automatic mail sorting, the automatic treatment of administrative files, the investigation forms, or the recording of the bank checks. These applications clearly indicate specificities of the handwritten word recognition compared to the Optical Recognition of Characters (OCR) which is related to the printed or typed characters. The progress in handwritten word recognition has mainly focused theoretical and methodological aspects. The most advanced and industrialized application concerns the mail automatic reading and the bank checks reading (Yacoubi *et al*., 1999; Anisirnov *et al*., 1999; Garcia-Salicetti, 1996; Kosrnola *et al*., 1999).

The on-line recognition field also knew the same progress for working on significant size lexicons and for omni-script writers applications (Kosrnola *et al*., 1999; Garcia-Salicetti, 1996; Ali, 2002; Kirn *et al*., 1997). A central spot in the systems of manuscript recognition is the recognition of the words. Classically, we distinguish two categories. The analytical approaches and the global solutions. The latest one requires a very significant training database. they depend on constraints corresponding to the size of the lexicon (limited or broad) and to the variability of the forms to be recognized (a number of script writers). In the analytical approach, the explicit strategies of segmentation (Dargenton, 1994) can be developed and ready to carryout multiple level of interpretation (Breuel, 2001; Kosrnala *et al*., 1998).

Two major problems arise in the two modes of recognition:

- The segmentation of the entities which belongs to the language writing.
- The characters representation way; i.e. the passage tools from the physical world to simple structures, where we can differentiate the various segments.

Particularly, these projects constitute the key stage of the automation of the information systems, the automatic reading of the mail, the bank checks, the old number plates and documentation. We propose in this study a technique based on a preliminary exploration using an Euclidean distance criterion as a pre-segmentation tool. At the laboratory level, the obtained results are with good agreement. The development of this technique at the industrial field requires certainly a some retouching. Indeed, both the data base must be intended for the neuronal network training and the reference index. Several methods are proposed in the specialized literatures which gave acceptable results. Study in this field is always active considering the difficulty of this crenel in particular, when they are the analytical methods.

Corresponding Author: Mohamed Senouci, LTE Laboratory, ENSET-Oran, B.P. 1523 El Mnaouer 31000- Oran, Algeria
Tel/Fax: (41) 41 09 16

ORGANIZATION OF THE PATTERN RECOGNITION SYSTEM

The handwritten word recognition is an application of the pattern recognition that operates according to the general diagram of the Fig. 1.

The analysis stage in this figure consists in computing the characteristics or parameters. They correspond to measurements of geometrical or topological nature this ones are usefully representing the form. This stage is named a characteristics extraction. The small space obtained is the parameters space with very small size compared to the space representation. The training is important step in recognition procedure. Its role is to configure the system and adjust its parameters to classify all examples. The decision stage is the ultimate step of the recognition. It delivers an opinion on the membership and separates the input shapes. It is question of seeking the classes in reference base which matches look like the shape in entry.

The Fig. 2 shows the type of writing which will be treated. Representing Latin handwritten word. No superposition or interference of the letters is allowed in order to eliminate all the problems from ambiguity or confusion which can arise. In addition, it is considered that the word is included initially in a maximum window of 350×150 pixels (simulation on a programming environment).

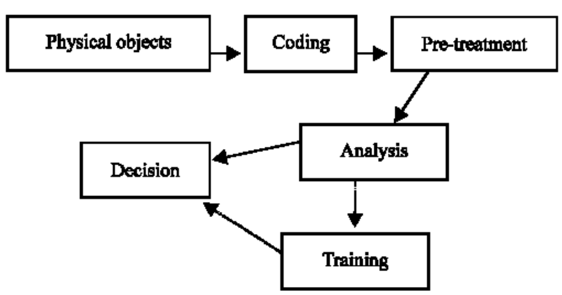

Fig. 1: Simplified structure of pattern recognition system (Xia and Cheng, 1996)

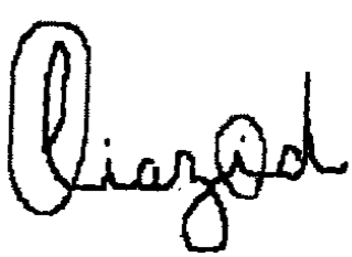

Fig. 2: Type of the processed handwritten word

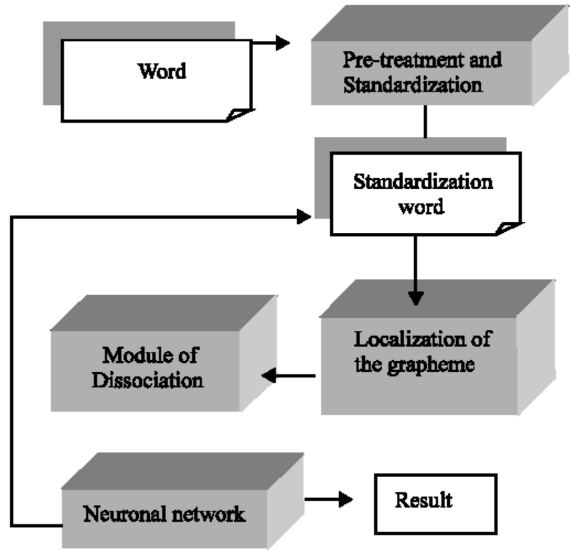

Fig. 3: Program structures view

In word recognition case, the entities are naturally the letters of the word like those surrounded in the Fig. 2. The architecture of this procedure is presented in the Fig. 3.

The system is composed of two essential layers. The first is useful for the segmentation. It's composed of two modules. The first allows the localization of the entity. It is a dynamic program able to detect the existence of the entity according to the resemblance rate compared with the reference index. Then the entity is used by the dissociation module. After detecting the entity which can be regarded as letter, it will be injected into the second layer which represents the ultimate stage of the recognition.

SEGMENTATION TECHNIQUE

The researchers are faced to the segmentation problem witch impossible to avoid. It belongs to the process of pretreatment and extraction of the information which is a precondition to any recognition.

Description of the used technique: At level one the word must be standardized according to the script form recommendations. The segmentation program generates a whole of solutions and chooses the optimal one by using an Euclidian distance, this norm uses an established reference data base. The program ensures four steps.

- Analyses the word and determines its limits.

Fig. 4: Examples of reference basis

- Stand in the beginning of the word and generates a first heuristic segmentation with the help of a HB step that represents the initial distance of dissociation. We determine HB = 30 experimentally as good value.
- Calculate the resemblance rate in relation to all elements of the reference basis and select the highest. Compare calculated rate to an experimental fixed value at 80% and

  Increase the step using a value which equalizes the maximum resemblance rate obtained and returns to (iii) else continue.

  Dissociate the segment so calculated rate ≥80% and bring back the blow in a window of 16×16 pixels.

  Repeat the operation until the end of the word.

The reference index is established in LTE laboratory. The entity-index reference comparison is not based on a directed recognition as it's the case in (Davalo and Naim, 1993). This base contains 26 characters and it is used to calculate the resemblance rate; Fig. 4 represents some examples.

The algorithm uses the VH2D (vertical horizontal 2 diagonals) projection defined below and the Euclidian norm defined by the Eq. 1:

$$d(X_i - X_k) = \left[\sum_{j=1}^{p} (X^j_i - X_k{}^j)^2 \right]^{\frac{1}{2}} \qquad i = 1....26; \qquad k = 1....NR \tag{1}$$

Where: NR is the number of segments constituting the word.

Xi is the vector of the dissociated elements

X_K is the elements of the reference index vector.

Figure 5 shows the synoptic mode of the treatment.

The dissociation of the entity uses a program which will receive the coordinates of the segment in progress. The Vertical-Horizontal 2Diagonal (VH2D) method is adopted to carry out the tests at each level i. The following operations are then made:

- Obtain a segment of level i
- Apply the VH2D projections, to have the comparison vectors in a plan of 16×16 pixels.
- Compute the Euclidian distances d_i with all reference-index.
- Repeat the process so much that the distances d_i is large.

The VH2D approach proposed in (Xia and Cheng, 1996) consists in projecting every character on the abscissa, on the ordinate and the diagonals 45 and 135°. The projections take place while calculating the sum of the values of the pixels i_{xy} according to a given direction.

The dissociated entity is normalized and projected on a 16x16 pixels window.

Presentation of the VH2D method

Vertical projection: The vertical projection of an image $I = I_{xy}$ (of dimension NR × NR) representing a character is indicated by:

Horizontal projection: Flat projection of an image $I = I_{xy}$ (of dimension NR × NR) representing a character is indicated by:

$$P^v = \left[P^v_1, P^v_2, ..., P^v_N\right] \qquad P^v_y = \sum_{x=1}^{N} i_{xy}$$

$$P^h = \left[P^h_1, P^h_2, ..., P^h_N\right] \qquad \text{Where} \quad P^v_y = \sum_{x=1}^{N} i_{xy}$$

Diagonal projection (45°): Projection on the diagonal 45° of an image $I = I_{xy}$ (dimension NR x NR) representing a character C is indicated by:

$$P^{d1} = \left[P^{d1}_1, P^{d1}_2, ..., P^{d1}_{2N-1}\right] \qquad \text{Where:}$$

$$P^{d1}_m = \begin{cases} \displaystyle\sum_{l=N-m+1}^{N} \sum_{k=1}^{m} i_{lk} & 1 \le m \le N \quad \text{et} \quad l = k + N - m \\ \displaystyle\sum_{l=1}^{2N-m} \sum_{k=m-N+1}^{N} i_{lk} & N+1 \le m \le 2N-1 \quad \text{et} \quad l = k + N - m \end{cases}$$

Projection on the diagonal 135°: Projection on the diagonal 135° of an image $I = I_{xy}$ (dimension NR × NR) representing a character C is indicated by:

$$P^{d2} = \left[P^{d2}_1, P^{d2}_2, ..., P^{d2}_{2N-1}\right] \qquad \text{Where:}$$

$$P^{d2}_m = \begin{cases} \displaystyle\sum_{l=1}^{m} \sum_{k=l}^{m} i_{lk} & 1 \le m \le N \quad \text{et} \quad k = m - l + 1 \\ \displaystyle\sum_{l=m-N+1}^{N} \sum_{k=m-N+1}^{N} i_{lk} & N+1 \le m \le 2N-1 \quad \text{et} \quad k = m - l + 1 \end{cases}$$

ALGORITHM FORMULATION

```
Begin
read the following image  (H, W):
 /* determination of the lower and higher limits
For I = 0  until H;
For J = 0 until W;
Jmax ← word Beginning ();
Jmin ← end of word (); / * reading the heuristic
parameter HB, and the threshold
of convergence Ψ */
word:  string;
word = ' ';
HB = 30;
K = 1;
Ψ  = 80%;
X = Jmin;
Y = X + HB;
While (Y ≤ Jmax) do:
Disos K ← dissociate (X, Y);
VH2D (disos K) ;
For i =1 to I = 26; do
Example I  = read Ref(i) base;
RESS [ I ] = Euc (Example I , Disos K ) ;
End for.
For i =2 to I = 26; do
Rate = Ress[1 ];
Rate = Min (Rate, RESS[i ]);
End for.
If (Rate ≥ Ψ )
RN ← Disos K  // neural network
Word = word + RN;  / / concatenation.
K ← K+1;
End of while.
End if
Else do
X = X + euc;
End of while
End
```

This algorithm generates the graphemes of the character which will be classified by a neuronal network classifier.

The Hb parameter is fixed in experiments. According to the script writer used the user to configure his HB.

NEURONAL NETWORKS

The second tool is a neuronal network Perceptron Multi-Layers type. These classifiers is trained with the Back propagation (Torres-Moreno, 1997; Lorette and Lecourtier, 1992; Bozinivic and Sriari, 1989; Belaid and Belaid, 1991), Fig. 5.

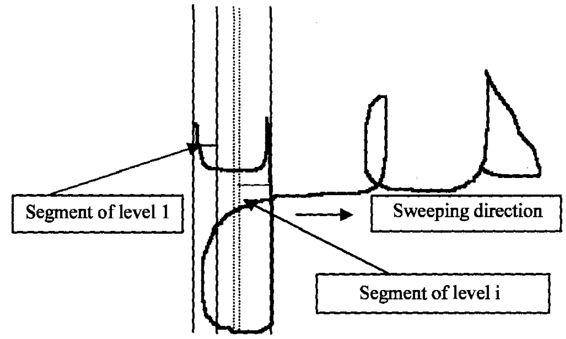

Fig. 5: Segmentation process

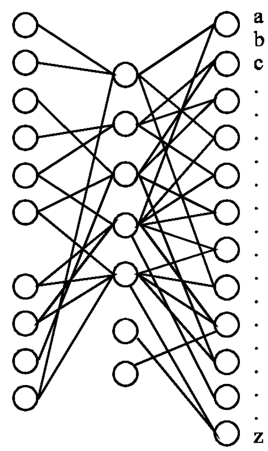

Fig. 6: Neuronal network architecture

Table 1: Results obtained by using neuronal network

No. of hidden layers	Iteration count	Rate of training (%)	Rate of recognition (%)
47	5000	85.20	75
53	2000	84.35	60
50	2000	90.40	80.50
50	5000	99.70	91.36

The selected neural network adopts the supervised training, uses the prototype data base LDB-LTE (Learning Data Base from LTE Laboratory). This training data base represents the Latin letters. It can be used as an experiment tool at the laboratory level. It contains 26,000 characters with 1000 characters by type, collected by 26 hand writers. The results obtained using this data base are resumed on the Table 1 where it is noted that the best rate of training is 99.70%. In addition, the best rate of the neuronal network recognition is 91.36% from a sample of 5200 letters written by 52 different persons. The network architecture is showed in the Fig. 6.

As the Fig. 6 shows this network:

- Has only one inlet layer.

- Include only one exit layer.
- Can include one to several hidden layers.
- Allows every neuron must be joined exclusively to all neurons of the following layer.
- Use a sigmoid transfer function f(x) = 1 / (1 + exp(-x))

Our network uses:

- Ninty four neurons at the inlet layer due to the use of VH2D method for the extraction of the characteristics.
- Fifty neurons at the hidden layer fixed experimentally.
- Twenty six neurons at the exit layer that correspond to the number of Latin small letters.

RESULTS AND DISCUSSION

The use of a neuronal network is justified by the following results: The segmentation algorithm determines the existence of an entity starting from the proposed reference index, but it is likely to fall into conflicts, in particular with the case when the letters have the same form (i.e., similar Euclidian criterion value) according to the reference script writer. The Table 2 shows our suggested classification based on observations of some standardized script writers. The intervention of the neuronal network confirms the decision on the one hand and corrects on the other hand.

To try out the algorithm, we used handwritten word samples written by 26 persons and adding up a whole of 10,400 words. Figure 7 shows some examples. The exactitude of the segmentation technique suggested amounts to 81.05% of recognition. This rate is obtained after a re-segmentation (second master key) of 4700 words taken arbitrarily as a sample of the initial ensemble. The complete process of the recognition (integrating the neuronal classifier), has given an 88.43% as an exactitude rate which is very satisfactory.

The results obtained depend completely by quality of the data bases used for the neuronal training or the reference index. The percentages which we present justify the use of the Euclidienne distance and the power of the hybridization of the networks of neurons with the Euclidienne distance. The integration of this technique in marketed codes requires certainly an update of the data bases.

Table 2: Classification of Latin letters

Classification	Latin letters
Class A	h, O, S, N, E, X
Class B	U, v, W, m
Class C	H, L, k,f
Class D	P, Q, D, B, T
Class E	y, G, Z, J
Class F	I, R, C

Fig. 7: Example of results: Segmentation-recognition of the words: yes liazid

Note: These results are obtained on open words: according to proposed critters.

The rate of segmentation 81.05%, Rate of recognition from the reference index (without neuronal network 60.48%), Rate of recognition of the system 88.43% .

CONCLUSIONS

A segmentation technique was proposed in this paper that consists in a vertical letters cutting based on an Euclidian pre-recognition approach. Our system recognizes all the words adapted to the script form presented in Fig. 2. The use of the VH2D technique is justified by the reason that the construction of the characteristics vector, witch decreases the global recognition time. Our system considers that the VH2D is sufficient because the recognition will be confirmed with two classifiers. The Euclidian distance pre-recognition helps the second classifiers to give exact segmentation results and eliminates the problems of slants. The results given depend on the quality of the bank samples. The prototype system has provided a good result at the laboratory level. It was used to segment the handwritten words of the learning database (LDB-LTE). This database is made up of 10,400 words written by 26 persons. The exactitude rate of the segmentation is 81.05%. For complete recognition process, the exactitude rate reached 91% which is very satisfactory. In the future we will tackle the problem of script forms adaptation.

REFERENCES

Ali, 2002. Nosary reconnaissance automatique de textes manuscrits par adaptation au scripteur. Thèse de doctorat de l'université de Rouen,

Anisimov, A., N.F. Gorski, O. Augustin. D. Baret, J. Price and C. Simon, 1999. A2IA Check Reader: A family of bank check reeognition systems. ICDAR '99. Bangalore, pp: 523-526.

Anquetil, E., 1997. Modélisation et Reconnaissance par logique floue; Application a la reconnaissance de lecture en ligne de lecture manuscrite omni-scripteurs, thèse de doctorat de l'université de rennes 1.

Belaid, A. and Y. Belaïd, 1991. Reconnaissance des formes, méthodes et applications. InterEditions.

Bozinivic, R.M. and S.N. Sriari, 1989. off-line cursive script world recognition. IEEE Trans. Pattern Analysis and Machine Intelligence, 11: 68-83.

Breuel, M., 2001. Segmentation of Hand printed Letter Strings using a Dynamic Programming Algorithm Xerox PARC, Palo Alto, CA, USA.

Dargenton, 1994. Patrice « Contribution à la segmentation et a la reconnaissance de l'écriture anuscrite » thèse de doctorat de l'institut du science appliqué de Lyon.

Davalo, E. and P. Naïm. 1993. Des réseaux de neurones. 2nd Edn., Eyrolles. Réseaux de neurones et reconnaissance visuelle application aux chiffres manuscrits, mémoire d'ingénieur de l'université d'Oran, 2001.

Garcia-Salicetti, S., 1996. Une approche neuronale rédictive pour la reconnaissance en-ligne de l'écriture manuscrite. Thèse de doctorat de L'université de Paris 6.

Kim, G., V. Govindaraju, A. Lexicon, 1997. Driven approach to handwritten word recognition for real-time applications. IEEE-Pattern Anlysis and Machine Intelligence. April, 19: 366-379.

Kosmala, A., G. Rigoll and D. Willet, 1998. A new hybrid approach to large vocabularv cursive handwriting recognition. ICPR 98. Brisbane, pp: 1512-1514.

Kosmola, A., and D. Willvett. 1999. Gerhard Rigoll. Advanced State Clustering for Large Voeabulary IIMM-based On-Line Handwriting Reeognitiotx JCDAR '99. Bengalore, pp: 442-445.

Lorette, G. and Y. Lecourtier, 1992. Reconnaissance et interprétation de textes manuscrits hors-ligne: Un problème d'analyse de scène BIGRE N°80 - CNED'92, Colloque National sur l'Ecrit et le Document, Nancy, pp: 109-135.

Torres-Moreno, J.M., 1997. Apprentissage et généralisation par des réseaux de neurones: étude de nouveaux algorithmes constructifs". Thèse de Doctorat, Institut National Polytechnique de Grenoble,

Xia, D.C. and H.D. Cheng, 1996. A Novel Parallel Approach to Character Recognition and its VLSI Implementation. Pattern Recognition, 29: 97-119.

Yacoubi, A., M. Giiloux, R. Sabourin and C.Y. Suen, 1999. An hmm-based approach for off line unconstrained I landwritten word modeling and recognition. IEEE-Pattern Analysis and Machine Intelligence. August, 21: 752-760.