Segmentation and Recognition of Handwritten Numeric Chains

Salim Ouchtati, Bedda Mouldi and Abderrazak Lachouri

Abstract—In this paper we present an off line system for the recognition of the handwritten numeric chains. Our work is divided in two big parts. The first part is the realization of a recognition system of the isolated handwritten digits. In this case the study is based mainly on the evaluation of neural network performances, trained with the gradient back propagation algorithm. The used parameters to form the input vector of the neural network are extracted on the binary images of the digits by several methods: the distribution sequence, the Barr features and the centred moments of the different projections and profiles. The second part is the extension of our system for the reading of the handwritten numeric chains constituted of a variable number of digits. The vertical projection is used to segment the numeric chain at isolated digits and every digit (or segment) will be presented separately to the entry of the system achieved in the first part (recognition system of the isolated handwritten digits). The result of the recognition of the numeric chain will be displayed at the exit of the global system.

Keywords—Optical Characters Recognition, Neural networks, Barr features, Image processing, Pattern Recognition, Features extraction.

I. INTRODUCTION

WRITING, which has been the most natural mode of collecting, storing, and transmitting information through the centuries, now serves not only for communication among humans but also serves for communication of humans and machines. The handwritten writing recognition has been the subject of intensive research for the last three decades. However, the early researches were limited by the memory and power of the computer available at that time. With the explosion of information technology, there has been a dramatic increase of research in this field. The interest devoted to this field is explained by the potential mode of direct communication with computers and the huge benefits that a system, designed in the context of a commercial application, could bring. According to the way handwriting data is generated, two different approaches can be distinguished: on-line and off -line. In the former, the data are captured during the writing process by a special pen on an electronic surface. In the latter, the data are acquired by a scanner after the writing process is over. Off-line and on-line recognition systems are also discriminated by the applications they are devoted to. The off-line recognition is dedicated to bank check processing, mail sorting, reading of commercial forms, etc., while the on-line recognition is mainly dedicated to pen computing industry and security domains such as signature verification and author authentication. Optical characters recognition (OCR) is one of the successful applications of handwriting recognition; this field has been a topic of intensive research for many years. First only the recognition of isolated handwritten characters was investigated ([1], [2]), but later whole words were addressed [3]. Most of the systems reported in the literature until today consider constrained recognition problems based on vocabularies from specific domains, e.g. the recognition of handwritten check amounts [4] or postal addresses ([5], [6]). Free handwriting recognition, without domain specific constraints and large vocabularies, was addressed only recently in a few papers. The recognition rate of such systems is still low, and there is a need to improve it. Character and handwriting recognition has a great potential in data and word processing, for instance, automated postal address and ZIP code reading, data acquisition in banks, text-voice conversions, etc. As a result of intensive research and development efforts, systems are available for English language ([7]-[9]), Chinese language [10], Arabic language [11] and handwritten numerals [12]. There is still a significant performance gap between the human and the machine in recognizing unconstrained handwriting. This is a difficult research problem caused by huge variation in writing styles and the overlapping and the intersection of neighboring characters.

Dr. Salim Ouchtati is with the Electronics Department, Skikda University, Route El Hadaik, Bp: 26 Skikda 21000, Algeria. Phone:0021393935198, fax: 0021338776192; e-mail: salim21_ouchtati@ yahoo.fr.

Pr. Mouldi Bedda is with the Electronics Department, Annaba University, Annaba, 2300 Algeria. E-mail: mouldi_bedda@ yahoo.fr.

Dr. Abderrazak Lachouri is with the Electronics Department, Skikda University, Route El Hadaik, Bp: 26 Skikda 21000, Algeria. E-mail: alachouri@ yahoo.fr.

II. OPTICAL CHARACTERS RECOGNITION SYSTEMS

Today, the OCR (Optical Characters Recognition) systems are only able to recognition high quality printed or neatly handwritten documents. The current research is now basing on documents that are not well handled and including severely degraded, omnifont machine printed text, and unconstrained handwritten text. A wide variety of techniques are used to perform handwriting recognition. A general model for

handwriting recognition is used to highlight the many components of a handwriting recognition system. The model begins with an unknown handwritten character that is presented at the input of the recognition system as an image. Firstly, to convert this image into information understandable by computers, parameterization operation is needed which extracts from the image all of the necessary meaningful information in a compact form, compatible with the computer language. This involves the preprocessing of the image to reduce some undesirable variability that only contributes to complicate the recognition process. Operations like slant correction, smoothing, normalization, etc. are carried out at this stage. The second step is to extract discrimination features from the image to either build up a feature vector or to generate graphs, string of codes or sequence of symbols. However, the characteristics of the features depend on the preceding step. Features extraction method is probably the most important factor in achieving high recognition performance in character recognition systems, extracted features must be invariant to the distortions, translations, and rotations. The features vector size is also important in order to avoid a phenomenon called the dimensionality problem. Several methods for features extraction are designed for different representations of the characters, such as binary characters, character contour, skeletons (thinned characters), or even gray levels characters [13]. The features extraction methods are valued in terms of invariance properties, and expected distortions and variability of the character. Today, the studies are based not only on how to choose the appropriate features extraction methods, but also on the selection of meaningful and pertinent features from the features vector ([14]-[16]). The final step is the character recognition; most recognizers have adopted classical pattern classification methods. Major approaches are statistical based, structural analysis, template matching, and neural network approaches. Significant progress has been made in these classification methods but more work is required to match human performance. Today, some recent research has shown improved performance using a combination of several different algorithms, many examples of such multiple classifier systems can be found in character recognition publications ([17], [18]).

III. A RECOGNITION SYSTEM FOR THE HANDWRITTEN ISOLATED DIGITS

In the setting of the handwritten writing recognition, we proposed an off line system for the recognition of the isolated handwritten digits (shown in the figure 1), this system is divided in three phases:

- Acquisition and preprocessing.
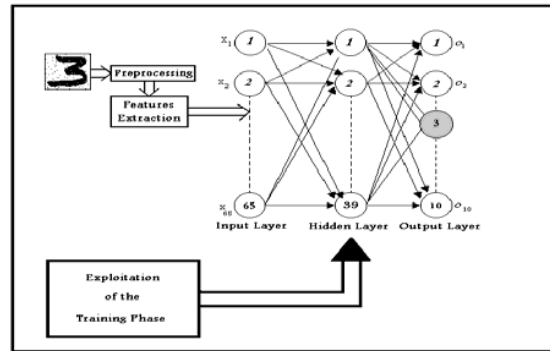- Features extraction.
- Recognition.

Fig. 1 General Schema of our handwritten isolated digits recognition system.

A. Acquisition and preprocessing

1) Acquisition

Before analyzing the different processing steps, let's recall that we are especially interested at the off line processing. For our case, the acquisition is done with a numeric scanner of resolution 300 dpi with 8 bits/pixels, the used samples are all possible classes of the handwritten digits (0,1,2,3,4,5,6,7,8,9) with variable sizes and   variable thickness, and with 100 samples for every class. Let's note that the characters images of our database are formed only by two gray levels: the black for the object and the white for the bottom. The figure 2 shows some samples of the used database.

2) Preprocessing

The preprocessing operations are classical operations in image processing, their objective is to clean and prepare the image for the other steps of the OCR system. The preprocessing attempts to eliminate some variability related to the writing process and that are not very significant under the point of view of the recognition, such as the variability due to the writing environment, writing style, acquisition and digitizing of image. For our case, we used the following preprocessing operations:

a)   Filtering and inversion of the gray levels

This operation consists in eliminating the noises in the binary image due to different reasons (bad Acquisition conditions , bad writing conditions, the writer's mood.. etc.), in our case, some digits are marked by the noise of type "peppers and salt", the application of the filter median on the digit image permitted us to eliminate easily this type of noise. Let's note that for reasons of calculation we reversed the gray levels of the character image (black for the bottom and white for the object). The figure 3 shows us the filtering and inversion operation of the gray levels of some handwritten digits.

b)   Normalization of the digit image

Knowing that the digits images have variable sizes, this

operation consists at normalizing the image size at 64 * 64 pixels (see figure 4).
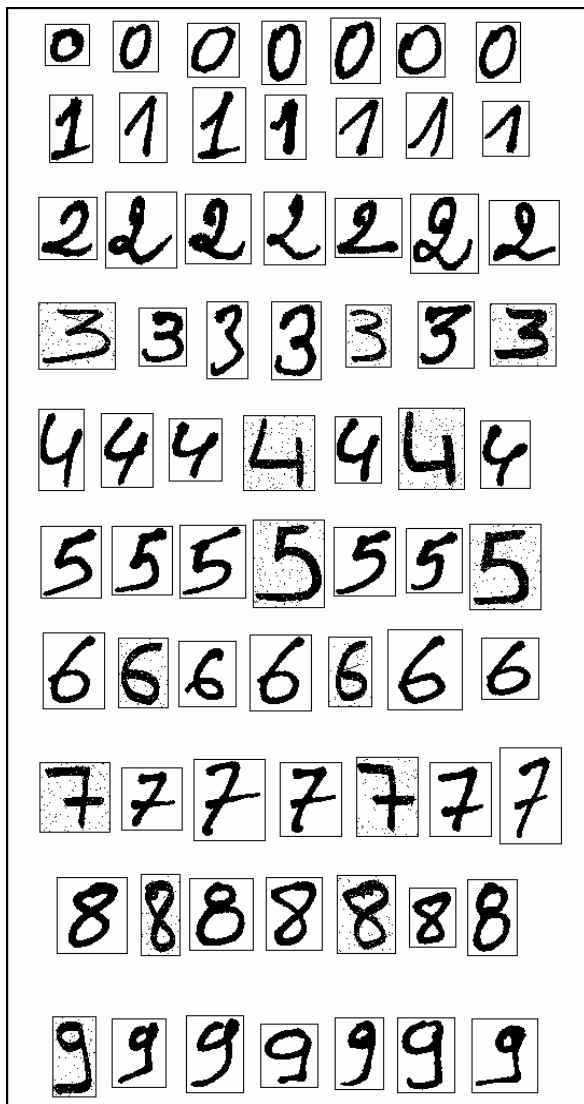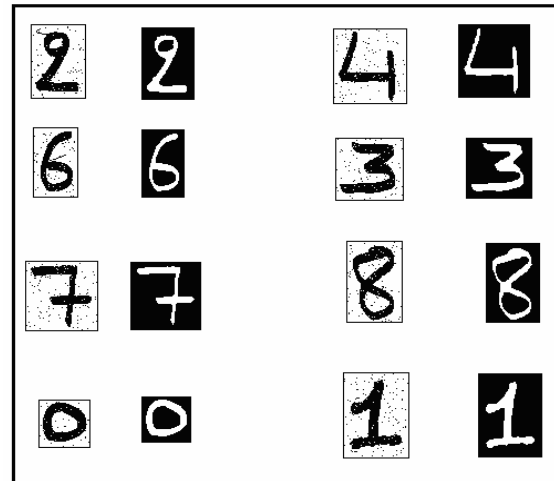

Fig. 2   Some samples of the used database.

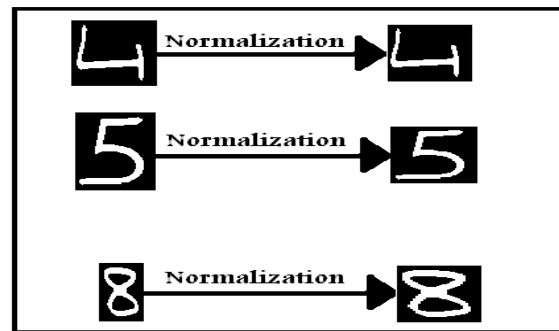Fig. 3   Filtering and inversion of the gray levels of some handwritten digits.

Fig. 4   Normalization of some handwritten digits

B. Features extraction

Features extraction is an important step in achieving good performance of OCR systems. However, the other steps also need to be optimized to obtain the best possible performance, and these steps are not independent. The choice of features extraction method limits or dictates the nature and output of the preprocessing step and the decision to use gray-scale versus binary image, filled representation or contour, thinned skeletons versus full-stroke images depends on the nature of the features to be extracted. Features extraction has been a topic of intensive research and we can find a large number of features extraction methods in the literature, but the real problem for a given application, is not only to find different features extraction methods but which features extraction method is the best?. This question led us to characterize the available features extraction methods, so that the most promising methods could be sorted out. In this paper, we are especially interested in the binary image of the digits, and the methods used to extract the discrimination features of are the following:

1) The distribution sequence

While dividing the digit image at a determined number of zones, the distribution sequence characterizes a number of the object pixels in relation to the total pixels number in a given zone. For our application, the digit image is divided in 64 zones (see figure 5), and the values of the distribution sequence are defined by:

$$x_i = \frac{N_i}{N} \tag{1}$$

With:

x_i : is the ith value of the distribution sequence.
N_i : is a number of the object pixels in the ith zone.
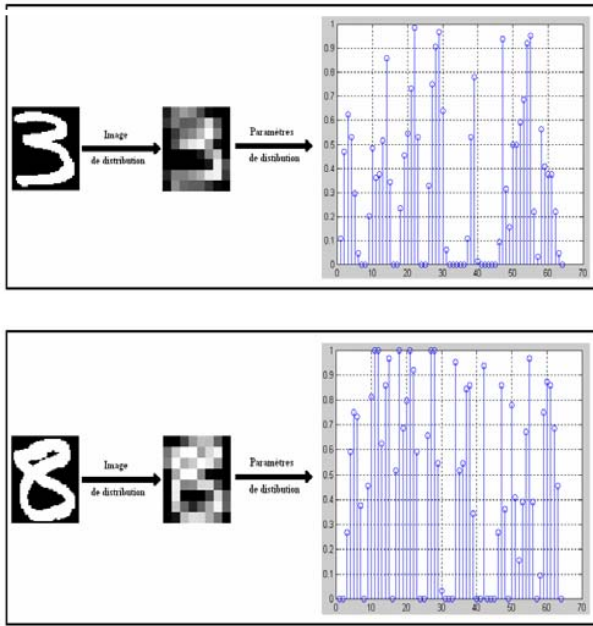N: is a total pixels number in the ith zone.

Fig. 5 The digits three and eight and their distribution sequences.

2) Barr-features

The Barr features have been used with success in several works [19], they are calculated on the binary digits images. Firstly, four images parameters are generated, and every image parameter corresponds to one of the following directions: east (e), North (n), Northeast (ne), Northwest (nw). Every image parameter has a whole value representing the Barr length in the direction in question. The features are calculated from the images parameters using zones that overlap to assure a certain degree of smoothing. Fifteen rectangular zones are arranged in five lines with three zones for every line; every zone is of size [(h/3)*(w/2)] where h and w are respectively the height and the width of the image. The high corners on the left of the zones are at the positions {(r0, c0): r0=0, h/6, 2h/6, 3h/6, 4h/6 and c0=0, w/4, 2w/4}. The

values in every zone of the parameters images are added and the sums are normalized, and the dimension of the features vector is 15* 4=60. If we suppose f_1, f_2, f_3, f_4 are the images parameters associated at a shape in entry and Zi (i=1,2… .15) is an rectangular zone of size [(h/3)*(w/2)] with the top corner on the left is (r_0, c_0), the value P_{ik} of the parameter associated to the Zi zone for the image parameter f_k (k=1,2,3,4) is given like follows:

$$P_{ik} = \frac{1}{N} \sum_{r=r_0}^{r_0+\frac{w}{2}} \sum_{c=c0}^{c0+\frac{h}{3}} f_k(r,c) \tag{2}$$

With N is a factor of normalization (for our application N=64).

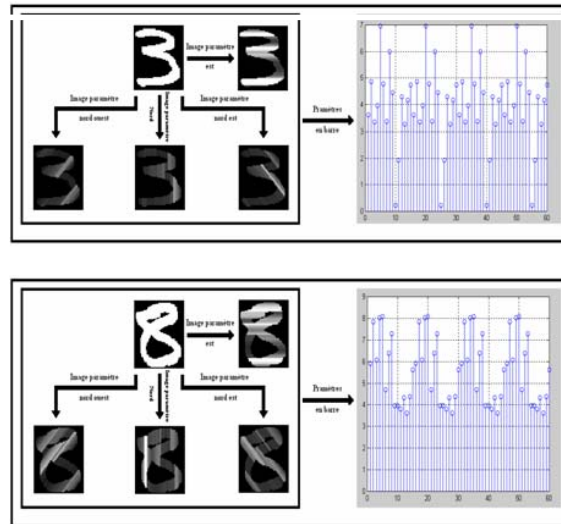The figure 6 shows the images parameters of the digits three and eight and their Barr features.

Fig. 6 The images parameters of the digits three and eight and their Barr features.

3) The centered moments of the different projections

In this case, the discrimination parameters are the centered moments extract from the follow projections (see Figure 7):

The vertical projection: for a given column, the value of the projection is equal to the number of object pixels in this column.
The horizontal projection: for a given line, the value of the projection is equal to the number of object pixels in this line.
The diagonals projections: for a given diagonal, the value of the projection is equal to the number of object pixels according the direction in question.

While the centered moments of the projection are given by the following formulas:

$$u_k = \sum_{i=1}^{M} (x_i - \bar{x})^k . p(x_i) \qquad (3)$$

$$p(x_i) = prob[x = x_i] \qquad (4)$$

$$\bar{x} = \sum_{i=1} x_i p(x_i) \qquad (5)$$

With

- u_k : is the centered moment of k order

- $\bar{x}$: is the mean value of the projection in question.

- $p(x_i)$: is the probability of 1 'element x_i in projection sequence.

- M: is projection sequence size.

For our application, we chose the first six moments for every projection.
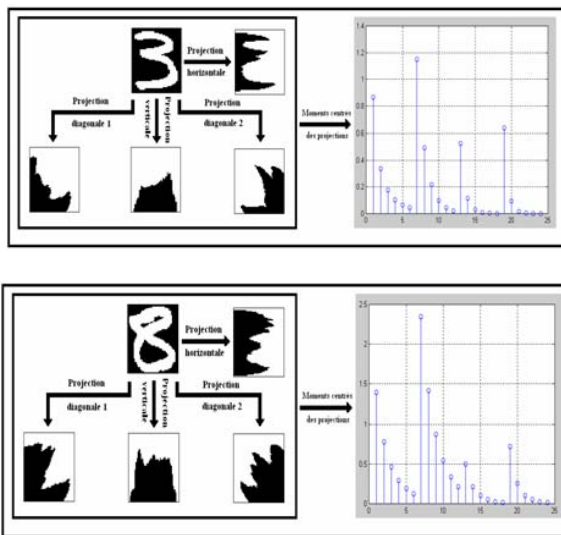

Fig. 7 The digits three and eight and their different projections.

4) The centered moments of the different profiles

By definition, the image profile in a given direction is the set of the object pixels seen while standing in this direction. From this point of view, we can define the following profiles (see Figure 8):

- Low profile: it is the set of the object pixels seen while standing below the image of the number
- High profile: it is the set of the object pixels seen while standing over the image of the number
- Left profile: it is the set of the object pixels seen while standing on the left of the image of the number

- Right profile: it is the set of the object pixels seen while standing on the right of the image of the number

While the centered moments of the different profiles are calculated by formulas: 3, 4 and 5.

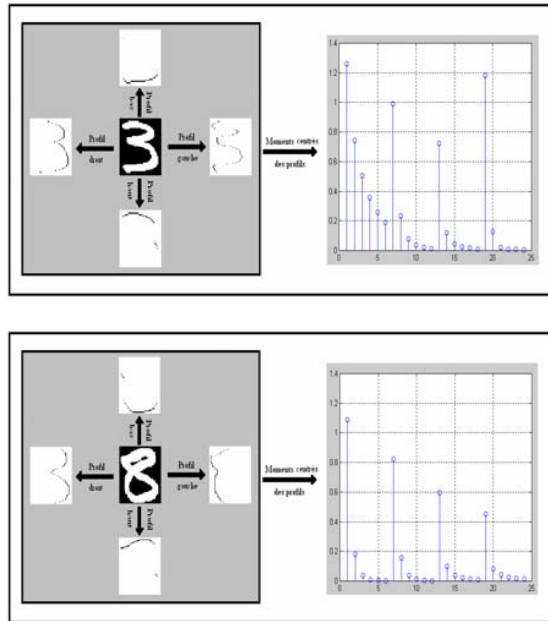For our application, we chose the first six moments for every profile.

Fig. 8 The digits three and eight and their different profiles

C. Digit recognition

The handwritten digits recognition is a problem for which a recognition model must necessarily take in account an important number of variabilities, dice at the time, the recognition techniques based on the neural networks can bring certain suppleness for the construction of such models. For our system, we opted for an MLP (Multi-Layers Perceptron) which is the most widely studied and used neural network classier. Moreover, MLPs are efficient tools for learning large databases. The used MLP in our work is trained with the back propagation with momentum training algorithm. The transfer function employed is the familiar sigmoid function.

1) The input data

The database consists of 1000 binary images. These images represent all classes possible of the handwritten digits (0,1,2,3,4,5,6,7,8,9) with variable sizes and variable thickness, and with 100 samples for every class. This database is divided to two sets, 70% for training the neural network and 30 % for testing it.

2) Neural network parameters

The input layer nodes number is equal to the size of the

used features vector (64 parameters of the distribution sequence, 60 parameters obtained by the Barr features method, the 24 centered moments of the different projections and the 24 centered moments of the different profiles, in other words N_IL=172), the output layer nodes number is equal to the classes number to recognize (N_OL=10), for the hidden layers, we used a single hidden layer with 87 nodes fixed by groping (N_HL=87). The initial connection weights are in the range [-1, 1]. The figure 9 shows us the structure of the used neural network.

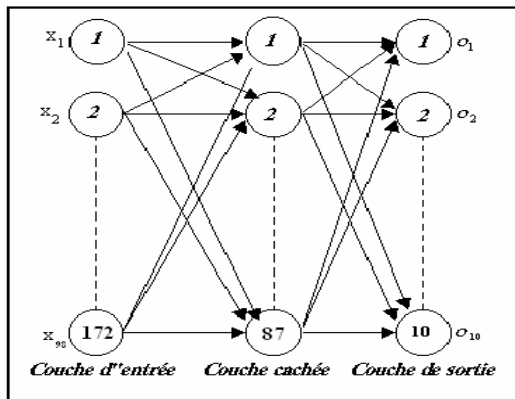Fig. 9 The structure of the used neural network.

3) The training process

For training the neural network, back propagation with momentum training method is followed. This method was selected because of its simplicity and because it has been previously used on a number of pattern recognition problems. The method works on the principle of gradient descent. The algorithm uses two parameters which are experimentally set, the learning rate η and momentum μ. These parameters allow the algorithm to converge more easily if they are properly set by the experimenter. For our case, we have opted for the following values: η =0.35 and μ =0.9. During the learning phase the neural network learns by example and the connection weights are updated in an iterative manner (Figure 8 and figure 9). The training process for the network is stopped only when the sum of squared error falls below 0.001.

4) The experimental results

The neural network performances are measured on the entire database (training or learning set and testing set). During this phase, we present the digit image to recognize to the system entry, and we collect at the exit its affectation to one of the possible classes.

The results can be:

- Recognized digit: the system arrives to associate one and only one prototype to the digit to recognize.
- Ambiguous digit: the system proposes several
- prototypes to the digit to recognize.
- Rejected digit: the system doesn't take any decision of classification.
- Non recognized digit: the system arrives to take a decision for the presented digit, but it is not the good decision.

The results and the different rates are regrouped in the table 1:

TABLE I
RESULTS AND DIFFERENT RATES

	R_R(%)	A_R(%)	J_R(%)	NR_R(%)
Training set	94.28	1.74	2.57	1.42
Testing set	93.00	3.66	2.00	1.33

With:

R-R: Recognizer rate   A-R: Ambiguity rate.
J-R: Reject rate.   NR-R: Non recognizer rate.

IV. HANDWRITTEN NUMERIC CHAINS RECOGNITION

In this part, we tried to widen our system not only to the recognition of the isolated handwritten digits, but also to the recognition of the handwritten numeric chains constituted of a variable number of digits. The questions that arise are how to exploit the realized system and what are the steps of processing that we must add to attain this objective. In this setting, we proposed an improved version of the previous system, but before the presentation of the new system, let's note that we are not interested at this moment at the segmentation problems; this last will be treated in our next work.

A. Preprocessing

After presentation of the handwritten numeric chain, the new system starts with cleaning it by applying an operation of filtering. In our case, some numeric chains are marked by the noise of type "peppers and salt", the application of the filter median on the chain permitted us to eliminate easily this type of noise. Let's note that for reasons of calculation we reversed the gray levels of the numeric chain image (black for the bottom and white for the object). The Figure 10 shows us the filtering and inversion operation of the gray levels of some handwritten numeric chains.
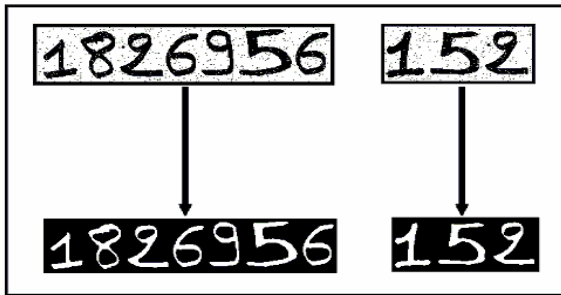
Fig. 10 Filtering and inversion of the gray levels of some handwritten numeric chains.

B. *The segmentation of the numeric chain*

The vertical projection for a given image is the number of the object pixels in every column of the image; it is obtained by the following expression:

$$V(j) = \sum_{i=1}^{n} IM(i, j) \qquad (5)$$

With:

- $V(j)$: is the value of j order of the projection sequence.
- $IM(i, j)$: The value of the pixel whose coordinates are i and j in the image of the numeric chain.
- n: is the number of the lines of the numeric chain image.

The segmentation of the numeric chain in digits is achieved by exploiting the transition black/white (or white/black) when we sweep for the last line the image columns of the vertical projection. (Figure 11).
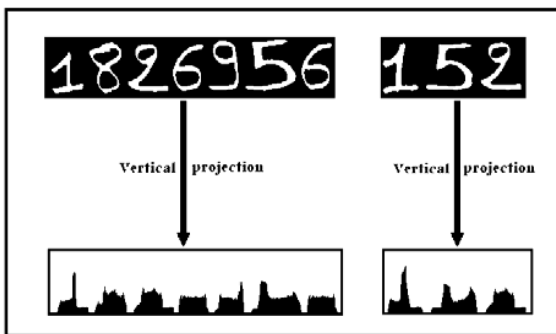

Fig. 11 The vertical projections of some handwritten numeric chains.

C. *Recognition of the numeric chain*

Every obtained digit by segmentation will be presented separately to the entry of the system achieved in the first part (recognition system of the isolated handwritten digits) and will undergo the following processing: normalization, features extraction and finally the recognition. Finally and after the presentation of all segments, the system displays the result of the recognition of the presented handwritten numeric chain. A detailed schema of the developed system is given by the Figure 12.
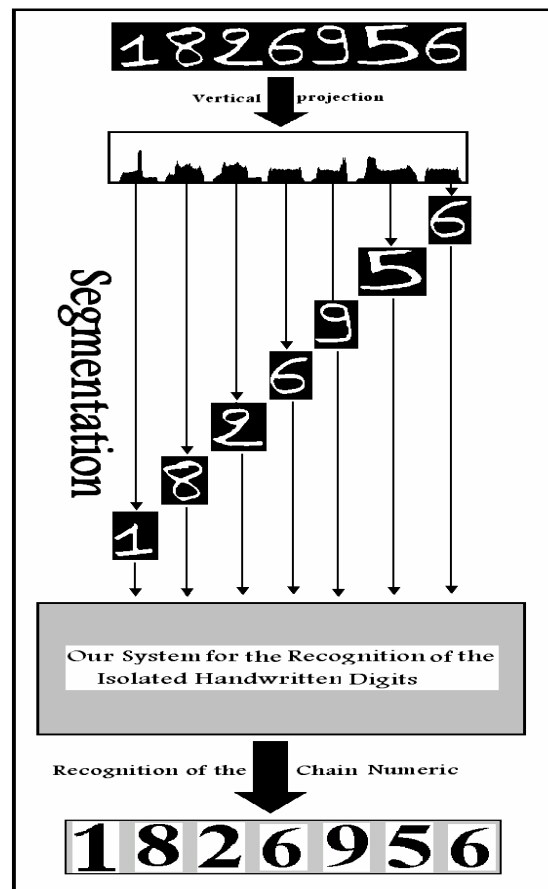

Fig. 12 Detailed schema of the developed system

V. CONCLUSION AND PERSPECTIVES

The recognition of the handwritten numeric chains is a problem for which a model of recognition must necessarily take in account an important number of variabilities and constraints due at the variation of the digit shape of the same class (variation of the writing styles, use of different writing instruments, variation of writing of a writer to another.. etc) and to the problems of the chain segmentation at isolated digits (problem of overlap for example). In our work, we presented an off line system for the recognition of the handwritten numeric chains. This work is divided in two big parts:

1. The realization of a system of isolated digits recognition: in this case the study is based mainly on the evaluation of neural network performances, trained with the gradient back propagation algorithm. The used parameters to form the input vector of the neural network are extracted on the binary images of the digits by two methods: the moments

of the distributions sequences and the Barr features.

2.The extension of our system for the reading of the handwritten numeric chains (constituted of a variable number of digits) by using the vertical projection to segment the numeric chain at the isolated digits and every digit will be presented separately to the entry of the system achieved in the first part (recognition system of the isolated handwritten digits).

The gotten results are very encouraging and promoters; however we foresee the following evolution possibilities:

1. To widen the database by taking in account a bigger number of writers and writing instruments.
2. To consider other classification methods.
3. To use of the algorithms capable to control the ambiguity, reject and non recognizer rates by adjusting the reject and ambiguity rates by use of suitable doorsteps.
4. To use of other features extraction methods.
5. Use of the post-processing techniques to improve the system performances.
6. To improve our system by Taking in account the segmentation problems and especially the problem of the digits overlap in the numeric chain.

REFERENCES

[1] J. Mantas, "An overview of character recognition methodologies", *Pattern Recognition*, Vol.19, No.6, 1986, pp.425–430.
[2] S. Mori, C. Y. Suen and K. Yamamoto, "historical review of OCR research and development", *proceedings of the IEEE*, Vol. 80, No. 7, 1992, pp. 1029-1058.
[3] A.L. Koerich, R. Sabourin, C.Y. Suen & A. El-Yacoubi, "A Syntax Directed Level Building Algorithm for Large Vocabulary Handwritten Word Recognition", *In 4th International Workshop on Document Analysis Systems (DAS 2000)* , Rio de Janeiro, Brazil, December 2000.
[4] L.S. Oliveira, R. Sabourin, F. Bortolozzi and C.Y. Suen, "A Modular System to Recognize Numerical Amounts on Brazilian Bank checks", *6th International Conference on Document Analysis and Recognition (ICDAR 2001)*, Seattle-USA, IEEE Computer Society Press, September 10-13, 2001. pp 389-394.
[5] A. Filatov, N. Nikitin, A. Volgunin, and P. Zelinsky, "The Address Script TM recognition system for handwritten envelopes", *In International Association for Pattern Recognition Workshop on Document Analysis Systems (DAS'98)*, Nagano, Japan, November 4-6 1998, pp. 157–171.
[6] A. El-Yacoubi, "Modélisation Markovienne de L''écriture Manuscrite Application `a la Reconnaissance des Adresses Postales", *PhD thesis, Université de Rennes 1, Rennes, France*, 1996.
[7] J. Hu, M. K. Brown and W Turin, "HMM based on-line handwriting recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, October . 1996, pp. 1039-1045.
[8] G. Kim and V. Govindaraju, "A lexicon driven approach to handwritten word recognition for real-time applications", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, April 1997, pp. 366-379.
[9] R. Buse, Z-Q Liu, and T. Caelli, "A structural and relational approach to handwritten word recognition", *IEEE Trans. Systems, Man and Cybernetics, Part-B*, Vol. 27, October. 1997. pp. 847-861.
[10] K. Liu, Y. S. Huang and C. Y. Suen, "Identification of fork points on the skeletons of handwritten Chinese characters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, October. 1999, pp. 1095-1100.
[11] A. Amin, "Off-line Arabic character recognition – the state of the art [review]", *Pattern Recognition*, vol. 31, No. 5, 1998, pp. 517-530.
[12] J. Cai and Z-Q Liu, "Integration of structural and statistical information for unconstrained handwritten numeral recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, March 1999, pp. 263-270.
[13] Oivid Due Trier, Anil K. Jain, & Torfinn Taxt, "Feature extraction methods for character recognition – a survey", *Pattern recognition*, Vol. 29, No. 4, 1996, pp.641-662.
[14] M. Bedda, M. Ramdani et S. Ouchtati., "Sur le choix d'une représentation des caractères manuscrits arabes", *Proceeding du 2ème Conférence Internationale Signaux, Systèmes, et Automatique SSA2'99*, Université de Blida, Algérie, 10-12 Mai 1999, pp73-84.
[15] F. Grandidier, "Un nouvel algorithme de sélection de caractéristiques application à la lecture automatique de l'écriture manuscrite", *thèse de doctorat en génie PH.D*, école de technologie supérieure, université du Québec Canada, Janvier 2003.
[16] N. Benahmed, "Optimisation des Réseaux de Neurones Pour la Reconnaissance des Chiffres Manuscrits Isolés, Sélection et Pondération des Primitives par Algorithmes Génétiques", *Thèse pour l'obtention de la Maîtrise en Génie de la Production Automatisée*, Montréal le 01 Mars 2002.
[17] T. K. HO, J. J. Hull and S. N. Srihari. "Decision Combination in Multiple Classifiers System", *IEEE Transactions on pattern Analysis and Machine Intelligence*, Vol. 16, No. 1, 1994, pp. 66-75.
[18] Y. S. Huang and C. Y. Suen.: "A Method of Combining Multiple Experts for the Recognition of Unconstrained Handwritten Numerals", *IEEE Transactions on pattern Analysis and Machine Intelligence*, Vol. 17, No. 1, 1995, pp.90-94.
[19] S. Ouchtati M. Ramdani et M. Bedda, "Un Réseau de Neurones Multicouches Pour la Reconnaissance Hors-Ligne des Caractères Manuscrits Arabes", *Revue Sciences et Technologie Université de Constantine*, No. 17, Juin 2002, pp. 99-105.