A Set of Features Extraction Methods for the Recognition of the Isolated Handwritten Digits

S. Ouchtati, M. Redjimi, and M. Bedda

Abstract—In this paper we present an off line system for the recognition of the isolated handwritten digits. The study is based mainly on the evaluation of neural network performances, trained with the gradient back propagation algorithm and fed by several feature vectors. 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, sondes application, the Barr features, The central moments of image coding according to the directions of Freeman, and the centred moments of the different projections and profiles.

Index Terms—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]-[13] and handwritten numerals [14], [15]. 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.

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 [16]. 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 [17]-[19]. 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.

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 (see Fig. 1), this system is divided in three phases:

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

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 one thousand (1000) 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 Fig. 2 (see Fig. 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 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 object and white for the bottom). The Fig. 3 (see Fig. 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 Fig. 4).

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

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 Fig. 5), and every zone is of size 8x8 pixel 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.

2) Application of the Sondes

This method is based on measuring the normalized distance between the border of the image and the first object pixel encountered (see Fig. 6), from this point of view, we can define the following probes distance:
- Top vertical Sondes
- Left horizontal Sondes
- Right horizontal Sondes

Let’s note that each Sonde is applied every eight pixels which allows us to encode the digit image with thirty two (32) parameters.

3) Barr-features

The Barr features have been used with success in several works [20], 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 \([(h/3)x(w/2)]\) 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 \((r0, c0): r0=0, 2h/6, 3h/6, 4h/6 and c0=0, w/4, 2w/4\). The values in every zone of the parameters images are added and the sums are normalized, and the dimension of the features vector is 15x4=60. If we suppose \(f_1, f_2, f_3, f_4\) are the images parameters associated at a shape in entry and \(Z_i\) \((i=1, 2, ..., 15)\) is an rectangular zone of size \([(h/3)x(w/2)]\) with the top corner on the left is \(( r_0, c_0)\), the value \(P_{ak}\) of the parameter associated to the \(Z_i\) zone for the image parameter \(f_k\) \((k=1,2,3,4)\) is given like follows:

\[ P_{ak} = \frac{1}{N} \sum_{r=r_0}^{r_0+h} \sum_{c=c_0}^{c_0+b} f_k (r, c) \tag{2} \]

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

The Fig. 7 shows the images parameters of the digits three and eight and their Barr features.

![Fig. 5. The digits three and eight and their distribution sequences.](image)

![Fig. 6. Application of the Sondes on the handwritten digit three.](image)

![Fig. 7. The Images parameters of the digit three and its barr features.](image)
4) The central moments of image coding according to the directions of Freeman

This method consists to divide the digit image into four zones of equal size (size of each zone is 32×32 pixel) and to encode each zone according to the eight directions of Freeman (see Fig. 8). In other words, for a given zone, each parameter is the accumulated object pixel in one of the directions of Freeman. So each zone will be encoded by eight parameters, and the digit image will be encoded by a sequence of thirty-two parameters (see Fig. 9). The central moments of the sequence of Freeman are obtained by the formulas (3), (4) and (5).

\[
\mu_k = \sum_{i=1}^{M} (x_i - \bar{x})^k \cdot p(x_i)
\]

(3)

\[
p(x_i) = \text{prob} \{x = x_i\}
\]

(4)

\[
\bar{x} = \sum_{i=1}^{M} x_i \cdot p(x_i)
\]

(5)

With
- \(\mu_k\): 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 \(i^\text{'th}\) element \(x_i\) in the sequence projection.
- \(M\): is projection sequence size.

For our application, we chose the first six moments.

5) The centered moments of the different projections

In this case, the discrimination parameters are the centered moments extract from the follow projections (see Fig. 10):
- 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 (3), (4) and (5).

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

6) 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 Fig. 11):
- 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.
7) Used features vector

It is the features vector used to characterize the digit image, and with which, we will nourish the recognition module. For our application we have tried to test six vectors (V1, V2, V3, V4, V5, V6) the Table I shows the composition of each vector.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Nbr</th>
<th>Nbr</th>
<th>Nbr</th>
<th>Nbr</th>
<th>Nbr</th>
<th>Nbr</th>
</tr>
</thead>
<tbody>
<tr>
<td>D_S</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>0</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td>S_P</td>
<td>0</td>
<td>0</td>
<td>32</td>
<td>32</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>B_F</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td>0</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>CM_FC</td>
<td>0</td>
<td>24</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>CM_PR</td>
<td>0</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>CM_PRF</td>
<td>0</td>
<td>24</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>24</td>
</tr>
</tbody>
</table>

With:
- D_S: Distribution Sequence
- S_P: Sondes Parameters
- B_F: Barre Features
- CM_FC: Centred Moments of Freeman Coding
- CM_PR: Centred Moments of different Projections
- CM_PRF: Centred Moments of the 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 ten thousand (10000) 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 one thousand (1000) 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 structure of the used neural network vary depending of the used vector parameters, the input layer nodes number (N_IL) is equal to the size of the used features vector, the output layer nodes number is equal to the classes number to recognize (N_OL=10), for the hidden layers, we used a double hidden layer for all the feature vectors, except for the fourth feature vector where we use one hidden layer, and the number of nodes (N_HL) for each hidden layer is fixed by groping. The initial connection weights are in the range [-1, 1]. The Table II shows us the different structures of the used neural network.

<table>
<thead>
<tr>
<th>Used vector</th>
<th>N_IL</th>
<th>N_HL1</th>
<th>N_HL2</th>
<th>N_OL</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>154</td>
<td>95</td>
<td>45</td>
<td>10</td>
</tr>
<tr>
<td>V2</td>
<td>172</td>
<td>85</td>
<td>50</td>
<td>10</td>
</tr>
<tr>
<td>V3</td>
<td>126</td>
<td>80</td>
<td>50</td>
<td>10</td>
</tr>
<tr>
<td>V4</td>
<td>98</td>
<td>45</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>V5</td>
<td>126</td>
<td>85</td>
<td>50</td>
<td>10</td>
</tr>
<tr>
<td>V6</td>
<td>210</td>
<td>125</td>
<td>97</td>
<td>10</td>
</tr>
</tbody>
</table>

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 \( \eta \) and momentum \( \mu \). 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: \( \eta =0.35 \) and \( \mu =0.9 \). During the learning phase the neural network learns by example and the connection weights are updated in an iterative manner. 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 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 III:

<table>
<thead>
<tr>
<th></th>
<th>R_R(%)</th>
<th>A_R(%)</th>
<th>J_R(%)</th>
<th>NR_R(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>93.06</td>
<td>2.24</td>
<td>2.5</td>
<td>2.2</td>
</tr>
<tr>
<td>V2</td>
<td>94.30</td>
<td>2.46</td>
<td>1.3</td>
<td>1.94</td>
</tr>
<tr>
<td>V3</td>
<td>90.50</td>
<td>4.30</td>
<td>2.4</td>
<td>2.8</td>
</tr>
<tr>
<td>V4</td>
<td>82.24</td>
<td>2.68</td>
<td>4.40</td>
<td>10.68</td>
</tr>
<tr>
<td>V5</td>
<td>89.90</td>
<td>2.5</td>
<td>3.1</td>
<td>4.50</td>
</tr>
<tr>
<td>V6</td>
<td>98.10</td>
<td>0.30</td>
<td>0.70</td>
<td>0.90</td>
</tr>
</tbody>
</table>
With:
R-R: Recognizer rate; A-R: Ambiguity rate.
J-R: Reject rate; NR-R: Non recognizer rate.

IV. CONCLUSION AND PERSPECTIVES

The recognition of the handwritten isolated digit is a problem for which a model of recognition must necessarily take in account an important number of variabilities and constraints due to 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. et al.). In our work, we presented an off line system for the recognition of isolated digits recognition, the study is based mainly on the evaluation of neural network performances, trained with the gradient back propagation algorithm and fed by several feature vectors. 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, spondees application, the Barr features, the central moments of image coding according to the directions of Freeman, and the central moments of the different projections and profiles.

The gotten results are very encouraging and promoters; however we foresee the following evolution possibilities:
- To widen the database by taking in account a bigger number of writers and writing instruments.
- To consider other classification methods.
- 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.
- To use of other features extraction methods.
- To use of the post-processing techniques to improve the system performances.

ACKNOWLEDGMENT

This work was supported in part by the Electronic Research Laboratory of Skikda, University of August 20, 1955 – Algeria- and Electrical Engineering Department, Al-jouf University – Arabie Saoudite.

REFERENCES


Salim Ouchtati was born on August 7, 1970 in Azzabba W Skikda Algeria. He received his B.Eng. and M.Sc. degrees in electronic from Annaba University, Algeria in 1994 and 1999 respectively. He obtained his doctorate diploma in 2007 in automatic and his HDR diploma (habilitation to direct the research) in electronic in 2010 from Annaba University. He is currently an associated professor at Skikda University, member of the Skikda Electronic Laboratory, responsible course of industrial control, responsible of the research project with the title: Selection and weighting of discrimination parameters in a recognizing handwritten digits system and member of the project: Realization of a system for reading of the Algerians postal checks. He has a several scientific works published in many international journals such as: “Segmentation and Recognition of Handwritten Numeric Chains” (Journal of Computer Science), “A New off Line System for Handwritten Digits Recognition” (Asian Journal of Information Technology), “An off line System for the Handwritten Numeric Chains Recognition” (International Journal of Soft Computing), His researches area focused mainly on the handwritten recognition artificial intelligence and image processing. Dr. Salim Ouchtati was a member of the scientific committee of the Electrical Engineering Department in Skikda University from 2005-2010, member of the scientific committee of the Technology Faculty in the Skikda University from 2005-2008, he was a reviewer in several international conferences such as: First International Conference on 'Networked Digital Technologies' (NDT)
Mohammed Redjimi was born on March 6, 1956 in Rejattas, Skikda, he received the diploma of docteur-ingénieur in computer science from the University of Lille 1, France in 1984, and his HDR diploma (Habilitation to direct the research) in computer science in 2007 from Annaba University. He is currently an associate professor of computer science at Skikda University, responsible of the team research “Modeling and simulation of complex processes” at Skikda Computer Science and Communication Laboratory, responsible of the research project with the title: “Cooperative approaches for images segmentation,” responsible of the research project with the title: “Bayesian approach in computer vision.” He has several scientific works published in many international journals such as: “An adaptive multi-agent system approach for image segmentation,” (International Journal of Computer Applications), “An image encryption approach using stream ciphers based on nonlinear filter generator,” (Theoretical and Applied Information Technology). His main research interests include modeling and simulation systems mainly by using multi-agents systems, image recognition and computing hardware and software systems. Dr. Mohammed Redjimi was a reviewer and member of the Scientific Committee in several international conferences such as: First Conference on Theoretical and Applicative Aspects of Computer Science – Skikda University--; Fractional Order Systems and Applications (SOFA 2010) – Skikda University

Mouldi Bedda was born on March 10, 1956 in El_Oued Algeria. He received bachelor’s degree from the University of Haouari Bouchaïd El_Arbi, Algeria in 1981, the M.Sc degree from the University of Languedoc Montpellier, French, in 1982, and the PhD degree in electrical engineering from the University Nancy 2, Nancy, French, in 1985. He was an assistant professor from 1985-1990, associate professor from 1990-2004 and professor from 2004-2006 at the University of Annaba Algeria, from 2006 to date full professor at the College of Engineering at Aljouf University KSA. He has several scientific works published in many international journals. Professor Mouldi Bedda was the director of Automatic and Signals Laboratory from 2001 -2006. His researches interests are: DSP; speech processing, OCR, artificial intelligence, and biomedical.