Fuzzy nearest neighbor system: An application to Handwritten Arabic literal amount words recognition.

Nadir FARAH, Mokhtar SELLAMI
Laboratoire de recherche en informatique
Université Badji Mokhtar
BP 12 Annaba
ALGERIE

Abstract: In recent years, fuzzy logic has been increasingly used to improve conventional methods especially in pattern recognition fields. The aim of this paper is Arabic literal words amount recognition using a fuzzy classifier. We introduce briefly the technique for processing handwritten words, which begins with the extraction of features, then their classification. The purpose of the classifier is to allocate a class to the test word on a basis of a training set. The fuzzification is introduced in two stages, firstly to reclassify the obtained K nearest neighbors by a classical K nearest neighbors approach. Secondly in the classification of the tested word to a class among its K neighbors. The proposed system was tested with a wide range of test images and an interesting success rate of classification was obtained.

Key Words: Word recognition, fuzzy nearest neighbors, membership value.

1. Introduction

Handwritten word recognition is among the most widely studied fields. It supports not only statistical, structural information and semantic ones, but also some physiological and psychological state of the writer.

This characteristic makes handwritten word recognition of consent especially in bank checks area.

Word recognition has become during the later decades almost universal. From that, many automatic systems have been developed and implemented. Most existing systems deal with some constraints and the results are interesting. For example, a limited lexicon or restricted writer number.

However, the handwritten deal with variability of the script and the noises generated by scanner.

To recognize a word, a fuzzy K nearest neighbor [1] is implemented in the Arabic handwritten literal amount recognition system described in this paper.
The proposed system we deal with consists of five parts, among them: data acquisition, preprocessing, feature extraction, recognition and post classification.
In data acquisition a handwritten literal amount are captured by a scanner, after which preprocessing techniques are used to prepare the image of words for feature extraction.

The preprocessing stage begins by dividing the literal amount into words,

using vertical histogram and a heuristic (space between words is of 1.5 times greater than the spaces between sub word). Then, binarisation is done on the obtained words; this consists of having a bimodal image from a multigray-level one [2], then a smoothing is used to filter noises [3].
The third part of our system is features extraction; this part is used to reduce the input vector image by measuring (expressing) it, using certain properties or features of the word image.
The features used by our system are the holistic ones, which are ascenders, descenders, loops, etc.

These features are quantitatively extracted from the image and used to recognize words.
For the recognition we use a fuzzy classifier to classify words. After feature extraction we use the vector obtained to compare it against a training set of feature vectors. The classifier tries to match these features to one of the 48 class's vectors.

The classifier generates candidate words with maximum proximity, which will be used by a syntactic analyzer to make decision about the word which satisfies the grammatical rules designed for this problem.
Ideally the words using the structural features should be well classified. But this is not the case due to the poor features extracted and variability of the script. This is always a certain amount of overlap between classes in the feature space.

In the proposed system a fuzzy nearest neighbor posses advantages of both nearest neighbor and fuzzy systems and are particularly powerful in handling complex, non linear and imprecise problems [4] such as handwritten word recognition. Two membership functions are used, the first one is to reclassify the generated K nearest neighbors and second one is to classify the test word according to the K nearest neighbors.

2. Feature extraction

We have been inspired by the human recognition that, considers the global high level words shape [5] [6]. For holistic paradigm there is a wide range of methods to words recognition. They can be basically classified in two categories:

- Statistical
- Structural

The statistical method is expressed in terms of partitioning the word feature space. The features are statistics based such as spatial distribution of black pixels, number of black pixels etc..

The structural method is expressed as a composition of structural units, and a word is recognized by matching its structural representation with that of a reference words.

احد	تسعة	ستون	اربعمائة	ألفا	مليار ان
اثنان	عشر	سبعون	خمسمائة	الفان	ملايير
ثلاثة	عشرة	ثمانون	ستمائة	مليون	سنتيم
اربعة	اثنا	تسعون	سبعمائة	ملايين	و
خمسة	عشرون	مائة	ثمانمائة	مليونا	دينار
ستة	ثلاثون	مائتا	تسعمائة	مليونان	دنانير
سبعة	اربعون	مائتان	ألف	مليار	سنتيمات
ثمانية	خمسون	ثلاثمائة	الاف	مليارا	جزائري

Table 1: Vocabulary of Arabic literal amounts.

The feature extraction step is carried out to determine words structures which may be used for recognition. These features are the observables, where the observation provides a value for each of the set of properties.

The main concept is to calculate the number of ascenders, descenders, loops, etc.
Base line detection [3] is the most important information that permits us to situate diacritical point's position, and the main part of the word.

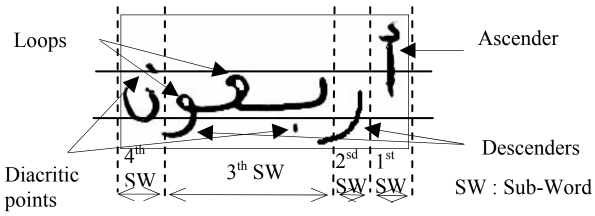

Fig. 2. Word's structural features (forty).

The boundaries follow up of the word's image, permits to have different parts like: sub words, loops, ascenders, descenders and diacritical points [7].
The structural features used in our approach are high level ones, which are numbers (table 2) of:

- Descenders,
- Ascenders,
- Loops,
- One dot above,
- Two dots above,
- Three dots above,
- One dot below,
- Two dots below,
- Sub words.

The features extracted are corresponding to 9 structural (Fig. 1) ones according to their possible occurrence numbers in the lexicon's word:

3 for ascenders,
2 for descenders,
2 for a one dot above,
2 for two dots above,
2 for three dots above,
1 for one dot below,
2 for two dots below,
3 for loops,
4 for sub words.

Arabic words	A	D	OD A	DD A	TD A	OD B	DD B	L	S B
خمسة			1	1				2	1
ستة				2				1	1
سبعة				1		1		2	1
تسعة				2				2	1
احد	1								2
ثلاثة	2			1	2			1	2
ثمانية	1		1	1	1		1	2	2
اثنان	2		2		1				3
اربعة	1	1		1		1		2	3
عشر		1			1				1
اثنا	2		1		1				2
عشرة		1		1	1			1	2
خمسون		1	2					2	2
ستون		1	1	1				1	2

A : Ascender, D : Descender, ODA : One Dot Above, DDA : Double Dot Above, TDA : Triple Dot Above, ODB : One Dot Below, DDB : Double Dot Above, L : Loop, SB : Sub-Word.

Table 2. A part of lexicon's word within their structural features.

Example: For the word ثلاثمائة (Thee hundred), we have: 3 ascenders, 1 double dot above, 2 triple dots above, 2 loops, 3 sub words.

3. Fuzzy K Nearest neighbor Classifier: (Fuzzy K-NN)

The classifier used in our system is a Fuzzy K-NN, which consist on proximity measures. It has been suggested by Pal & Majumder [8].

Fuzzy nearest neighbor classifiers are ideally suited for modeling the non parametric distribution on handwritten word recognition data.
For the purpose of our system the data were divided as of training and test type.

For a given word X, the fuzzy classifier computes the membership X in different classes $C_1,..,C_j,..C_m$. The membership of X in class C_j can be expressed as $\mu_j(X)$. The test word is allocated to a class for which the membership function yields the maximum value.

After having generated the K nearest neighbors for a test word, the fuzzification principle is used in two stages. Firstly, it is used in reclassification of the K nearest neighbors obtained by the classical K-NN. This reassignation tries to redefine class boundaries. Formally we express it by: looking for memberships (by distance calculation) of each neighbor (noted y_j) with training classes (noted i class), for every training class we have p_i prototypes noted Z_p, this memrbership function is given in (1):

$$\mu_i(y_j) = \left[1 + \left(\max_{p = 1..p_i} d\left(y_j, Z_p\right) / F_d \right)^{Fe} \right]^{-1} \qquad (1)$$

This function permits to introduce fuzziness, which permits to reclassify y_j in classes where it presents the highest membership value. When neighbor's membership value has been tested with the training set, we compute the membership of test word X noted $\mu_i(X)$ calculated for each of the K nearest neighbor classes, using formula (2):

$$\mu_i(X) = \left\{ \mu_i(y_j) * \exp\left(-a * d(X, y_j) / d_m\right) \right\} \qquad (2)$$

dm represents the average distance between words of the same class in the reference set. *a*, *Fe*, *Fd* are constants that determine the degree of fuzziness in membership space, which has been fixed experimentally to the following values : *a*=0.45, *Fd*=1, *Fe*=1. We have used a threshold S that has been fixed to 0,7.

Since the value of $\mu_i(X)$ increases while the distance value (d) decreases, therefore, for a tested word using a threshold S, a decision rule is stated as follows: let N be the number of classes where the membership function is greater than S, then:

- If N =0, X is rejected, membership function too low.
- if N=1 or N>1 and $\mu_i(X)$is unique, X is recognized
- If N>1 and $\mu_i(X)$ is not unique, there is ambiguousness.

4. Syntax based Post classification

The classification phase has generated a list of candidate's words pondered by confidence values, which is the membership value. We will consider from this point that a candidate is a couple of information, the word class and its confidence value.

```
<Hundreds> ::=
        <Hund>+و+<Less_Hund> |
        <Hund> |
        <Less_Ten>+ مائة + و+<less_Hund> |
        <Less_Ten>+ مائة
<Hund > ::=
     مائتان | ثلاثمائة | اربعمائة | خمسمائة
     | ستمائة | سبعمائة | ثمانمائة | تسعمائة
< less_Hund > ::=
        < Less_Ten > |
        <Comp_Nbr>
< Less_Ten > ::=
        احد |
        اثنان |
        ثلاثة | اربعة | خمسة | ستة
              | سبعة | ثمانية | تسعة
```

Table 3. A part of the grammatical rules used.

When obtaining the list of candidate words by the recognition stage (Classification), we first sort it by word's confidence value, and then we can consider two cases:

- If on one hand, there is a word which confidence value is greater than the other, and if this word succeeds the syntactic analysis, the word is kept, and will be part of the resulted literal amount. If on the other hand, the word

doesn't match the syntax, it is rejected and the next word of the list will be analyzed.
- If at the head of the list, two words have the same confidence value and satisfy the syntactic analysis, we consider this case like an ambiguousness, which can be raised with the use of high level information, the courtesy (numeric) amount for example.

5. Results

For the purpose of the fuzzy K-NN we have constructed four (04) reference sets of different sizes (Table 4), in order to determine the best value of the K parameter and recognition rates.

Reference set	set 1	set 2	Set 3	set 4
Number of tested words	1200	1200	1200	1200
Number of reference set words	96	144	240	480

Table 4. Reference sets used.

Recognition rates gotten for this classifier, according to reference sets and the value of the parameter K, are represented in table 5.

	Recognition rate		
K	1	3	8
Set 1	85,00	85,00	87,86
Set 2	91,20	92,10	82,10
Set 3	92,30	93,10	90,13
Set 4	92,60	93,80	89,47

Table 5. Word recognition rates.

From these results the value of the parameter K has been fixed to 3, which represents the K classes with highest membership values.

6. Conclusion

The Arabic literal amount considered in our case are composed with 48 words. In this paper we have used structural features to perform recognition, and we have tested our system on a basis of 1200 words (the 48 words of the lexicon written by 25 different writers). The word shapes are analyzed with a fuzzy classifier obtaining an average recognition rate of 93,80 %. Comparing this work with the work on a basis of Arabic literal amount using a neuro-symbolic system described in [9], where a recognition rate of 80 % was highlighted, we have raised the recognition rate by about 13,80 % which is a significant improvement.

We conclude that this performance is very interesting and represents a promising platform, on which more investigation and/or improvements may be done.

References

[1] S. Singh, A. Amin, ‘Fuzzy Recognition of Chinese Characters’, Proc. Irish Machine Vision and Image Processing Conference (IMVIP’99), Dublin, 8-9 September, 1999.

[2] T. Pavlidis, ‘Algorithms for Graphic and Image Processing’, Rockville, MD: Computer science press, 1982.

[3] A. Belaid, Y. Belaid, ’Reconnaissance des formes: Méthodes et applications’, InterEditions, 1992.

[4] S. K. Pal, ‘Fuzzy Sets in Image Processing and Recognition’, IEEE Technology Update Series, 1994.

[5] S. Madhvanath, V. Govindaraju, 'The Role of Holistic Paradigms in Handwritten word Recognition', IEEE Trans. On Pattern Analysis and Machine Intelligence, vol. 23 no.2, February 2001.

[6] T. Steinherz, E. Rivlin, N. Intrator, 'Off-line cursive script word recognition: A survey', International Journal on Document analysis and Recognition, IJDAR, Vol 2, pp: 90-110, 1999

[7] A. Ameur, K. Romeo-Pakker, H. Miled, M. Cheriet, "Approche globale pour la reconnaissance de mots manuscrits Arabes", CNED'94, 3th National Colloque on writing and document, pp: 151-156, July 1994.

[8] S. K. PAL, D. D. Majumder,'Fuzy mathematical approach to pattern recognition', John Wiley, New York, 1986.

[9] L. Souici-Meslati, M. Sellami, 'Reconnaissance de montants littéraux arabes par une approche hybride neuro-symbolique', RFIA'2002, 11th francophone Congress AFRIF-AFIA, for Pattern recognition and Artificial intelligence, Angers, January 2002.