Pattern Recognition Letters 31 (2010) 361–371

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

New features using fractal multi-dimensions for generalized Arabic font recognition

Sami Ben Moussa [a,b,*], Abderrazak Zahour [b], Abdellatif Benabdelhafid [b], Adel M. Alimi [a]

[a] Research Group on Intelligent Machines (REGIM), National School of Engineers of Sfax, BP W, 3038 Sfax, Tunisia
[b] Le Havre University, Quai Frissard, BP 1137-76063, Havre, France

ARTICLE INFO

Article history:
Received 25 February 2008
Received in revised form 20 October 2009
Available online 28 October 2009

Communicated by C.L. Tan

Keywords:
Optical font recognition
Fractal features
Texture analysis
OCR
Arabic written

ABSTRACT

In this work, a new method is proposed to the widely neglected problem of Arabic font recognition, it uses global texture analysis. This method is based on fractal geometry, and the feature extraction does not depend on the document contents. In our method, we take the document as an image containing some specific textures and regard font recognition as texture identification.

We have combined both techniques BCD (box counting dimension) and DCD (dilation counting dimension) to obtain the main features. The first expresses texture distribution in 2-D image. The second makes possible to take on the human vision system aspect, since it makes it possible to differentiate one font from another. Both features are expressed in a parametric form; then four features were kept. Experiments are carried out by using 1000 samples of 10 typefaces (each typeface is combined with four sizes). The average recognition rates are of about 96.2% using KNN (K nearest neighbor) and 98% using RBF (radial basic function). Experimental results are also included in the robustness of the method against written size, skew, image degradation (e.g., Gaussian noise) and resolution, and compared with the existing methods. The main advantages of our method are that (1) the dimension of feature vector is very low; (2) the variation sizes of the studied blocks (which are not standardized) are robust; (3) less samples are needed to train the classifier; (4) finally and the most important, is the first attempt to apply and adapt fractal dimensions to font recognition.

1. Introduction

The character recognition is a field which is already the subject of several researches, although the recognition rate varies from one font to another (Amin, 1998; Khorsheed, 2007). A font is defined as a set of alphabets in the same family. Its typographical characteristics are standardized in printing industry. The most efficient Arabic optical character recognition system has a difficulty in recognizing some fonts and an error rate of about 4% in reading tasks in the absence of context (Amin, 2000).

Font recognition is a fundamental subject in the identification and document analysis (DA). However, automatic document processing (ADP) techniques have to take font type into account as one of the two main aspects.

First, the precision of an OCR (optical character recognition) system decreases considerably if the original text presents several fonts. It remains a classification problem which requires to be achieved by using appropriate conditions and a maximum of information about the pattern to be classified (Hamdani et al., 2006). Therefore, font classification allows the recognition of only one kind of font to reduce the number of alternative forms in recognition process (Aviles-Cruz et al., 2005).

Second, the retroconversion consists in finding the leading structure of the document (physical and logical fields). The reproduction of a digitized document requires not only the identification of the characters but also the fonts used in the original text. The identification of a word, a line, or a text font allows us to characterize the document structure. The font recognition is useful for the logical entities definition, the title of the document, of sections or of paragraphs (Belaïd et al., 2005; Ben Moussa et al., 2005; Ding et al., 2007).

In spite of the importance of the OFR in an OCR system, it often remains a neglected problem, and the studies in this field are few, particularly in the Arabic writing. The morphological and topological complexities of the Arabic writing make it difficult to develop an omni font OCR. According to Kahan and Pavlidis (1987) an omni font system is able to recognize any font, generally without training, while being based on the topological and morphological writing rules. Font identification in the Arabic writing is not an easy task because of the morphological complexity of this script (Alimi, 1995; Alimi, 1997). Most works used only the mono-font context (Shi and Paulidis, 1997; Zramdini, 1998; Khorsheed, 2007).

* Corresponding author. Address: Research Group on Intelligent Machines (REGIM), National School of Engineers of Sfax, BP W, 3038 Sfax, Tunisia. Fax: +216 74 275 595.

E-mail addresses: sami.benmoussa@ieee.org (S. Ben Moussa), abderrazak@univ-lehavre.fr (A. Zahour), benabdelhafid@univ-lehavre.fr (A. Benabdelhafid), Adel.Alimi@ieee.org (A.M. Alimi).

0167-8655/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2009.10.015

In this paper, we present a novel contribution to the a priori OFR approach based on the global texture analysis of document images. The key point is using texture analysis to extract global features to reduce the processing difficulties in a recognition system, and especially to make the Arabic multi-font recognition successful.

In Section 2, we present the related works in this field. In Section 3, we describe on the one hand, both techniques to extract features, based on fractal dimensions (FD) and on the other hand our system and the algorithm-connected parameters to obtain font features. In Section 4, we present the Arabic font database and experiment results. In Section 5, we show the robustness study of our approach on skew, written size, resolution and noise. In Section 6, we present the discussion and conclusions.

2. Related works

The Arabic font identification is not a simple task due to the morphological complexity of this script (Ball et al., 2006; Kherallah et al., 2009). Therefore, many works in Arabic recognition used multi-font contexts, without proceeding to font recognition (Kanoun et al., 2002; Khorsheed, 2007). Optical character recognition allows a considerable improvement in the system performance. The best results were achieved in the case of mono-font applications, or of a limited vocabulary (Essoukri Ben Amara and Bouslama, 2003).

There are several techniques that have been proposed for font recognition. We notice three font identification levels, which are the results of a segmentation step, namely block or paragraph level, line level and the word or letter level. (a) In (Yang et al., 2006) a proposed method to recognize Chinese fonts is based on empirical mode decomposition (EMD), by analyzing five basic strokes to characterize the stroke attributes of six fonts. Aviles-Cruz et al. (2005) proposed a method based on high-order central moments for 8 Spanish fonts. These methods are based on grey level analysis. However, the multi-channel Gabor filtering techniques applied by Zhu et al. (2001) were based on binary image analysis, which is considered as texture identification. (b) The statistical approach used by Zramdini and Ingold (1995) and Zramdini (1998) is based on the local typographical Latin font. (c) Ding et al. (2007) employed a 3-level wavelet transform on the normalized images of seven Chinese fonts. Each prototype is presented by 448 features. In (Sun, 2006) the proposed system is based on the stroke-stored template collection, to identify five fonts in both English and Chinese scripts. In the same context, Shi and Paulidis (1997) used heuristic hybrid approach which combines the priori and the posteriori approaches, based on histogram properties. The method of Cooperman (1997) is based on a set of local detectors to identify individual features of each font from word image, such as height, width, thickness, and base line. We can note that in spite of the importance of the AFR in an Arabic OCR system, it often remains a neglected problem, and the studies in this field are few. This is due to the morphological complexity of cursive Arabic script. We detected three types of problems of the Arabic OFR. First, a problem of calligraphical writing and its topology (Massoudy, 1998). Second, a problem of Arabic letters' variability; the characters of the same font can considerably change the morphology while changing position in the word. Finally, and especially the presence of diacritics, dots and zigzag in Arabic writing (Alimi, 2002).

The proposed methods in (Aviles-Cruz et al., 2005; Yang et al., 2006; Zhu et al., 2001) are based on global font identification. In these works, pre-processing steps to obtain uniform blocks are used. Aviles-Cruz et al. and Yang et al. regarded font recognition as grey level texture identification. However, Zhu et al. considered it as binary texture identification. In global approaches, font recognition is based on the feature extractions from uniform block; the robustness is depending on pre-processing step such as character size normalization, spacing between lines or successive characters and the existence of empty spaces. Local approach – line or word level – is based on heuristic and statistic analysis generally related to the studied scripts. Therefore, most of the proposed method can give good discrimination accuracy in prefix level (block or line); but it is not efficient at the same time for all levels and each method cannot be generalized on all image's size.

The fractal dimension (FD) is a useful method to quantify the complexity of an image. It constitutes a geometrical description of an image (Peitgen et al., 1992; Moalla et al., 2006). In this paper, a generalized OFR based on fractal theory is proposed. The proposed method is based on global VHS features to extract visual aspect from image text. It does not depend on the level analysis.

Fractal geometry has been applied to various fields, such as mathematics, physics, and biology. It made possible to model complex random experiments. It is again used for texture image analysis (Kasparis et al., 1995; Jin, 1995; Ajay and Jibitesh, 2001), and in the pattern recognition (Neil and Curtis, 1997). In this latter, only one of the various fractal techniques is used for pattern classification. In the following sections, we will explain our choice to use two techniques for FD process to obtain features based on font recognition.

3. Fractal-based font features

The values of FD depend on the number of $N(r)$ intervals of length 'r' for covering an object 'E', whatever the way in which these intervals are selected. It is estimated theoretically as (Jin, 1995; Ajay and Jibitesh, 2001; Zhang et al., 2002):

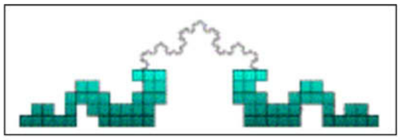
$$\mathrm{FD}(E) \equiv \frac{\log N(r)}{\log\left(\frac{1}{r}\right)}. \tag{1}$$

The concept of fractal dimension, as an estimator, is a topological notion. A number of fractal surface dimension estimators have been developed. However, different estimators lead to different results when applied to not strictly self-similar object (Zhou and Nina, 2005). The problem of creating the right notion of dimension is very complicated. Therefore, for the same fractal object, it is not true that the Hausdorff dimension and the Minkowski–Bouligand dimension are usually the same. Hausdorff dimension is obtained by using elementary dilation technique of image textures. However, Minkowski–Bouligand dimension is obtained by using box counting method, for details see (Peitgen et al., 1992). HD does not go beyond 2 or the MD can exceed the 2 value. The written image is not a perfect fractal curve. Therefore, those dimensions are not interpreted in the same way.

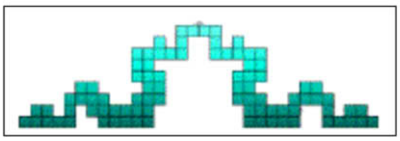
In this paper, we proposed a method that allows to explore two techniques to calculate fractal dimensions, such as Hausdorff dimension HD and Minkowski–Bouligand dimension MD (self-similarity dimension). Therefore fractal dimension estimators are studied using both (i) box counting dimension: BCD, related to HD and (ii) dilation counting dimension: DCD, related to MD.

Fig. 1 shows an example of BCD process of various steps of the Von Koch curve. It characterizes the texture repartition aspect in 2-D image, by varying the box sizes used as measurement unit. Fig. 2 constitutes an example of DCD process. It allows describing the vision aspect at varying scales for each dilation level. By increasing the dilation level, the details at vision of the curve decrease, this is related to increasing the distance vision between the object and the viewer (Ben Moussa et al., 2006).

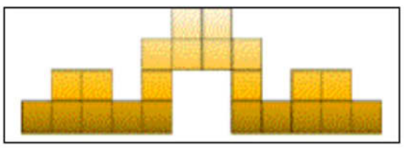
Our method is based on exploiting both the later aspects, on the one hand, on extracting visual aspect of visual human system through the image elementary dilation method; and on the other



(a): recovering process of Von koch curve

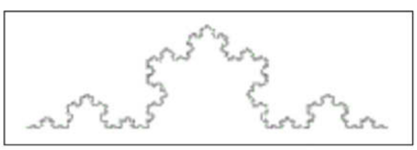


(b): Von Koch curve covered by boxes with r seize

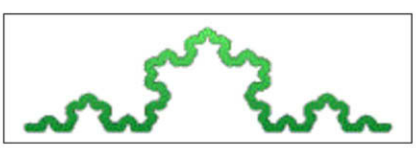


(c) : Von koch curve covered by boxes with n*r seize

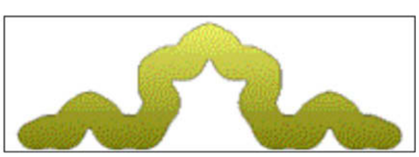
Fig. 1. General aspect of curve in BCD process.



(a) : Original image



(b) : Von Koch curve after dilation with r level



(c) : Von Koch curve after dilation with n*r level

Fig. 2. General aspect of curve in DCD process.

hand, on extracting the surface recovery aspect by using box unit. More details are exposed in the following sections.

3.1. Box counting dimension

Box counting dimension (BCD), related to the HD, is applied for any structure in a surface. It allows determining the recovering aspect in 2-D image for both invariant objects and with different scales, by covering an image by boxes of size 'r' and by determining the number of boxes of size 'r' noted $N(r)$, which are necessary to cover the image:

$BCD(r)$ = the box counting dimension of a set E, where r is the maximum size of the box.

$N(r)$ = the smallest number of boxes needed to cover the set E

$$\mathrm{BCD}(r) = \lim_{r \to 0} \frac{\log N(r)}{\log\left(\frac{1}{r}\right)}. \tag{2}$$

To estimate the box dimension, Euclidean space containing the image can be divided into a grid of boxes of size 'r', and counting boxes, which are not empty (Fig. 3). This process must be repeated every time. We fixed the initial value of the box size, and the advance step between two successive units of measurement, by the value one. A graph of $\log N(r)$ versus $\log(1/r)$ for each size can be produced. Then the correlation between $N(r)$ and $1/r$; BCD is adjusted by linear regression (Zengl et al., 2001).

3.2. Dilation counting dimension

Dilation counting dimension DCD is related to the definition of Minkowski–Bouligand (MD) for a curve as follows: to measure the obtained surface for each level of dilation centers, the levels of dilation are obtained by modifying the radius 'd' of the centers, for any curve E (Fig. 4). DCD depends on the radius of dilation. It is equal to (Cetera, 2001):

$$\mathrm{DCD}(d) = \lim_{d \to 0} \left(n - \frac{\log V(d)}{\log(d)} \right), \tag{3}$$

where

$\mathrm{DCD}(d)$ is the dilation counting dimension of a set E,
d is the maximum dilation radius,
n is the dimension of a space,
V is the dilation body of a set E which is defined as

$$V = \{x \in \Re^n : \exists a \in E / \mathrm{dis}(a, x) <= r\},$$

where

$\mathrm{dis}(a, x)$ is the distance between point a and x,and
$V(d)$ is the n-dimensional volume of the set V.

These procedures must be repeated until the radius dilation attains the minimal size. $\mathrm{DCD}(d)$ is obtained by a linear regression line of log uplet (the radius of dilation, the corresponding surfaces).

3.3. General scheme for FSAFR

Fractal-based system of Arabic font recognition: FSAFR is a topical system of pattern recognition. It is described in Fig. 5. It simplifies considerably the post processing in an OCR system. Moreover, it improves the performance of the system while profiting from various information relating to the available font. This work treated precisely the following steps after segmentation in OCR process and it can be used independently for other process in document analysis (clustering, indexing, etc.). The text images used in this works, as input, are provided by ALPH-REGIM database (Ben Moussa et al., 2008). Moreover, we consider that test basis is supplied by image blocks of writing, resulting from a segmentation step.

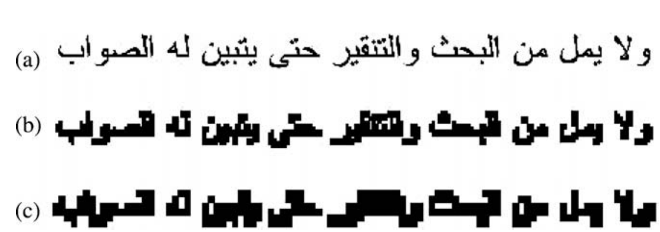


Fig. 3. Example of text image recovery by boxes: (a) original text image, (b) text image covered by boxes with r size, (c) text image covered by boxes with $2 * r$ size.

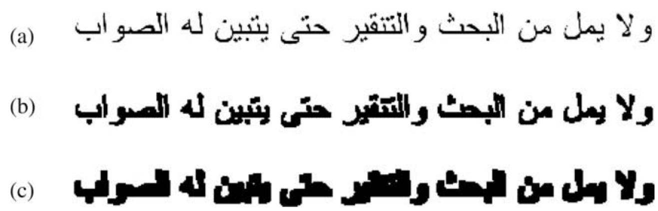


Fig. 4. Example of various dilation levels of text image: (a) original text image, (b) text image after dilation with r level, (c) text image after dilation with $2 * r$ level.

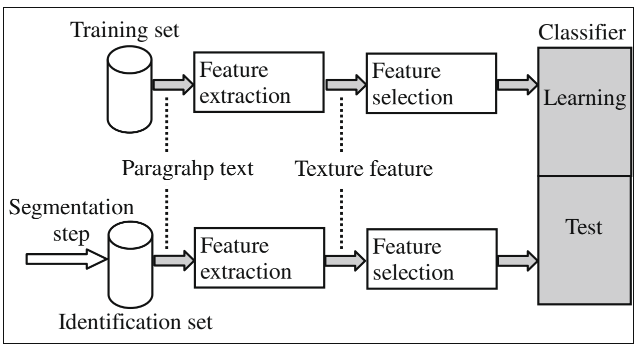


Fig. 5. The flow chart of FSAFR.

Therefore, the size of these blocks is not limited in their heights, and the maximum width of a box is that of the width of an A4 format page. BCD and DCD are used to extract features from text images.

The size of 'r' or 'd' cannot be fixed in advance, because of the diversity of sizes of images and writings. The choice of the sequence $\{r\}$ or $\{d\}$ can affect the precision of FD estimator because there is a powerful relation between the limits choice of 'r' or 'd' and the estimator. Some research tried to reduce the estimation error by choosing the "optimum" limit. This later showed better estimation accuracy for only some examples (Zengl et al., 2001). This is justified by the Zengl et al. works, which showed that the limits choice of 'r' or 'd' size depends on the abstraction level at the computing process of fractal dimensions and cannot be determined by a determinist function. BCD features depend on the limited size of selected boxes as measuring unit.

BCD dimension is expressed according to the several values of r, because of not normalized text image. The aim is to study the influence of these arguments on their impact on the discrimination of the writings. The extracted features are noted $B(r)$.

In practice, the algorithm for calculating $B(r)$ is presented below

```
Choose the maximum box size, noted 'r'.
For each image E
    For L = 2 to r
    Applied a grid G of L size to E
        For each unit of G in E
        Compute the number of Gi,j ∈ E
        Add the (N(L),L) uplet to SIZE_B set
        end
    end
Calculate linear regression of log-uplets in SIZE_B
Extract B(r)
end.
```

DCD is obtained by the texture surface measurement for each level of dilation, for each value of 'd' (radius limit of dilation), we obtained derivative function, noted $D(d)$. These features are characterized finally by a maximum dilation order. It depends on both parameters: the maximum dilation radius and the value of step between two successive dilations.

But in practice, it is estimated as follows:

```
Choose the maximum dilation noted 'd'.
For each image E
    For L = 1 to d
        For each pixel ai,j of E
        If ai,j ∈ E then dilatation (L)
        end if
        end
        Add the (V(L),L) uplet to SIZE_D set
    end
Calculate linear regression of log-uplets in SIZE_D
Extract D(d)
end.
```

BCD features depend on the size limit of the chosen box. DCD features depend on the maximum level of dilation. Both limits depend on the accuracy FD estimator and the text image sizes. An image is constituted of one or more lines. A line of text constitutes the smallest width of an image. We used a 1000-prototype test base, with 200 dpi resolution. Then we adopt several limits, which are evaluated in the experimentation.

In order to obtain an exact measurement containing the maximum of details, we fixed all the initial value of the box size, the initial dilation radius, and the step between two successive measurement units, equal to one. The limit sizes of the r, respectively, d, are equal to $\{15, 20\}$.

4. Experiments and results

There are mainly two combination strategies of the OFR with the OCR: the posteriori approach $(1-a,\ 1-b)$ and the priori approach $(2-a,\ 2-b)$ presented in Fig. 6. In this work, we opted the priori identification, which defines typefaces without recognizing characters.

In the following, we are interested in the description of our AFB: Arabic Font Base. Then we present the obtained results for the Arabic font recognition. Finally, we expose the details of the selected features BCD and DCD.

4.1. Arabic font basis: AFB

According to the related works presented in session 2, we can notice, with the exception of AFB, the absence of a standard database of the Arabic font. The IFN/ENIT-database is more interesting for handwritten Arabic words competition (Maergner et al., 2005). The AFB processes Arabic fonts. It is extracted from ALPH-database, which is constituted by Arabic and Latin/Printed and Handwritten scripts (Ben Moussa et al., 2008). AFB is composed by 1000 lines of

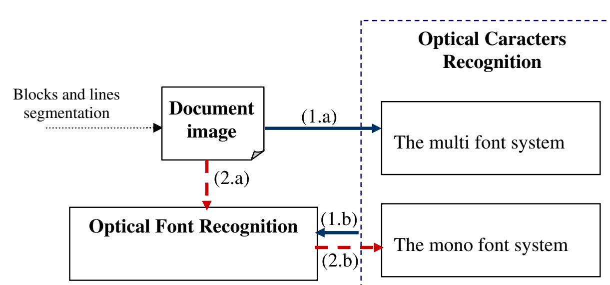


Fig. 6. OFR and OCR combination strategies' process: (1) a posteriori approach, (2) a priori approach (Shi and Paulidis, 1997).

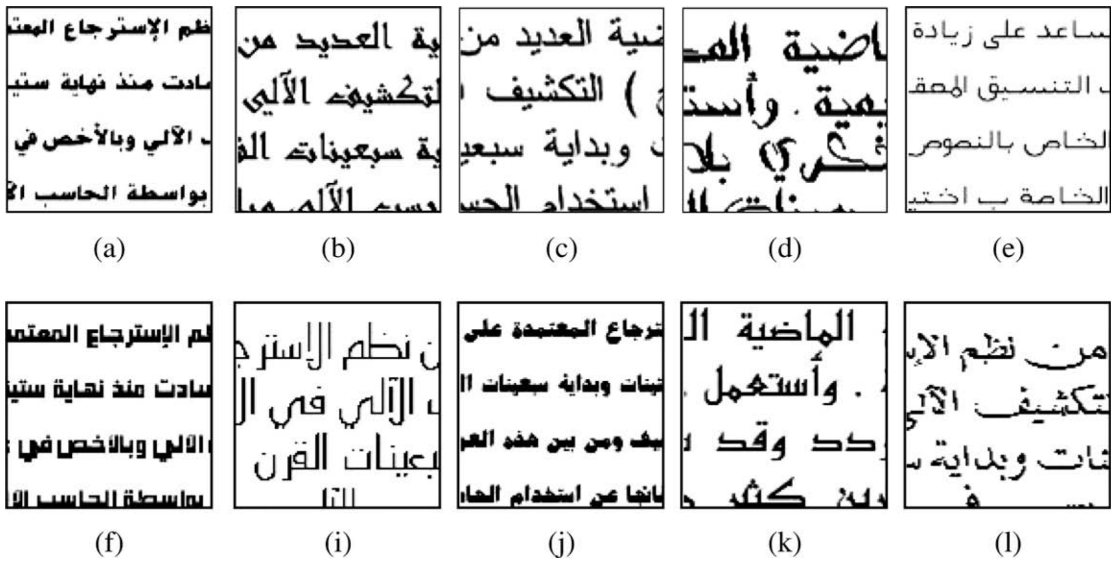


Fig. 7. Example of typeface with 16 points of size: (a) Ahsa, (b) Andalus, (c) Arabic_transparant, (d) Badr, (e) Buryidah, (f) Dammam, (i) Hada, (j) Kharj, (k) Koufi and (l) Naskh.

Arabic texts of size 16 points, produced by image processing software and scanning process, with a resolution of 200 dpi. The 500 first lines are adopted in A4 format; for the remaining lines, we adopted other formats. We retained different margins on the right and on the left in order to not normalize the sizes of text lines or paragraphs, and by admitting that the paragraphs do not have the same number of lines. Then we reproduced the pages obtained from the preceding process in 10 fonts. The selected fonts are AH: Ahsa, AN: Andalus, AT: Arabic_transparant, BD: Badr, BU: Buryidah, DA: Dammam, HA: Hada, KH: Kharj, KF: Koufi et NS: Naskh, which are frequently used in the Arabic publications (Massoudy, 1998; Alimi, 2002; Ben Amara and Gazzah, 2004; Ben Moussa et al., 2006; Zaghden et al., 2006); an example of each font is represented in Fig. 7.

Then we segmented the pages of each font in several paragraphs (a paragraph can consist of one or several lines); the latter are marked at the beginning of this process; we retained the 100 first paragraphs in each font. Finally, for each font, we applied a modification of the writing sizes in 50 random paragraphs. The sizes applied are 14, 18 and 20. The same paragraph cannot have the same writing size in the whole of 10 fonts on the one hand; on the other hand, the writings belonging precisely to a font cannot have the same size.

The AFB is constituted of 100 prototypes by font. In all 1000 font images, the originality of base choice is expressed by the reconstitution of the same block (paragraph or line) for different selected font.

4.2. Recognition of types faces

We used both algorithms BCD and DCD with several alternatives of ‘r’ and ‘d’; the extracted features are $B(15)$, $B(20)$, $D(15)$ and $D(20)$. We studied graphically the feature discrimination capacity by using two features among four. Fig. 8 represents the localization of the font zones according to $B(15)$ and $D(15)$. The presented figure shows the discriminative aspect of our features.

We can notice a good localization of three fonts such as BD: Badr, BU: Buryidah and DA: Dammam on the one hand; on the other hand the overlapping between the classes, due to the similar fonts. The overlapping problem can be treated by the combined use of several features. Finally, we can announce that BCD and DCB can be considered as Arabic font features. Fig. 9 shows an example of discriminative capacity of only one characteristic $D(15)$. The choice of an adequate combination among the four features will be carried out in the following experiments. Table 1 shows some examples of the results obtained by the use of only one feature and by the association of others.

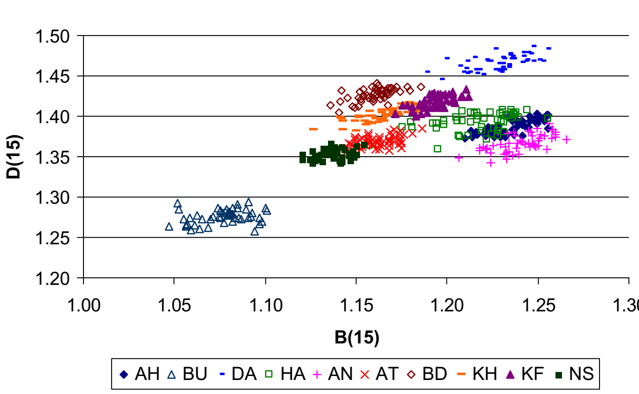


Fig. 8. Graphic localization of 10 fonts according to $B(15)$ and $D(15)$ features.

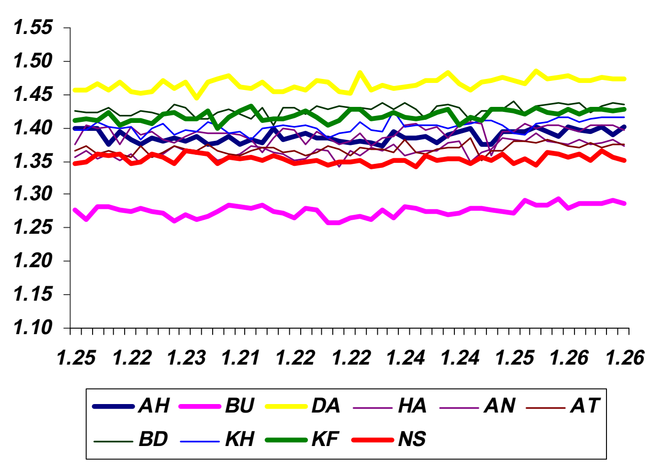


Fig. 9. Discriminative aspect by using $D(15)$ feature.

Table 1
Example of the obtained results according to the selected features.

Features list		Recognition rate (%)
B(15)		42.80
B(20)		50.20
D(15)		53.20
D(20)		45.80
B(15)	*D*(15)	88.20
	D(20)	81.60
B(20)	*D*(15)	82.00
	D(20)	77.60

Table 2
The confusion matrix of 10 fonts.

Font	AH	AN	AT	BR	BU	DA	HA	KH	KF	NS
AH	94%	4%	0	0	0	0	2%	0	0	0
AN	0	94%	0	0	0	0	4%	0	2%	0
AT	0	0	92%	0	0	0	0	0	4%	4%
BR	0	0	0	100%	0	0	0	0	0	0
BU	0	0	0	0	100%	0	0	0	0	0
DA	0	0	0	0	0	100%	0	0	0	0
HA	0	0	0	0	0	0	100%	0	0	0
KH	0	0	0	4%	0	0	0	88%	8%	0
KF	0	0	0	0	0	2%	0	0	98%	0
NS	0	0	2%	0	0	0	0	0	0	98%

4.3. Obtained results

The experimental results are carried out by using KNN classifier to classify the 10 fonts. Each font is presented by 100 samples. For each font, we used 50 samples for training process and 50 samples for test process. Each sample is expressed by a vector of four features. In total, the number of prototypes is equal to 500 for each one of the training and test database.

The average discrimination rate of the 10 fonts is about 96.6%. The confusion matrix is presented in Table 2. We noted a good discrimination rate for some fonts such as Badr, Buryidah, Dammam, and Hada. Although the errors of classification are produced by confusion between some prototypes in the classes: AN, BR, HA, KF and NS. An increase in the number of prototypes in the base of training cannot resolve this problem. Table 3 represents the obtained results by a various alternative of a training and test prototype number. These results obtained due to increase/decrease of prototypes' number in the base of training prove that our method requires only some significant prototypes for the training. A better rate of classification is not related to several prototypes in the training basis, but depends on the choice of the component of training.

To evaluate our discriminative aspect features, we used a neuronal technique of self-organization map (Versanto et al., 1999). Each prototype uses four features. The network of Kohonen is composed of two layers; the first constitutes the input of the network by using four features for each prototype and the second is the output – the number of neurons in the output layer is equal to 30 * 30. The map's *X* and *Y* axes presented according to the number of neurons. This technique allows a representation in 2-D to facilitate the observation of the distributed zones, which is not the case for a 4-D representation. The construction of the distribution map is based on two steps: (1) for each initial prototype among the 10 classes, the values at the map may even be selected at random; (2) each map unit *i*, collects almost similar sample belonging to the neighborhood set. The second step is competitive and unsupervised. The obtained result is represented in Fig. 10 according to the number of neurons. It shows, on the one hand, the efficiency of our features for the font classification, and on the other hand, the source of overlapping between the classes. The number of semantic classes in Fig. 10 is higher than the number of the real classes. The use of a neuronal technique for font classification can improve the system performance.

4.4. Neuron network classification

We used the RBF neuron network to define architecture (one class one neural network). We defined an OCONN system including as much neuron networks as classes (10 classes). Each neuron network represents a class among 10 fonts; the adapted strategy is one versus other. Finally, we obtain dedicated neuron network systems, each one for one class. All units constitute a global recognition system. The number of input is equal to four (number of features); the number of neural networks is equal to 10 (a number of definite font classes). Each neuron network is a recognition system making it possible to recognize only one class among all classes. Each RBF neuron network is able to recognize only one class among all recognized by the system. The results are presented in Table 4.

In Table 4, we have decreased the training sample number then we have tested the accuracy results; in spite of training base composition, which did not include all models, the obtained accuracy results did not decrease.

4.5. Script independent

The proposed method does not involve a study of local topography. Furthermore, no explicit local analysis is needed in the method. In our method, we take font recognition as texture identification, so text image is considered as an image containing some specific textures. This method is script independent as well as the method proposed by Zhu et al. (2001) and Aviles-Cruz

Table 3
Example of the obtained results by increase/decrease of a number of training and test prototypes.

Training prototype number		800	700	600	500	400	300	200
Test prototype number		200	300	400	500	600	700	800
Font recognition rate (%)	AH	100	93.33	95	94	95	95.71	95
	AN	100	93.33	95	94	95	97.14	97.5
	AT	100	100	97.50	92	88.33	88.57	86.25
	BR	100	100	100	100	100	100	100
	BU	100	100	100	100	100	100	100
	DA	100	100	100	100	100	100	96.25
	HA	100	100	100	100	100	98.57	90
	KH	95	93.33	87.50	88	90	91.42	96.25
	KF	95	96.66	97.50	98	98.33	97.14	95
	NS	100	100	95	98	96.66	95.71	97.50

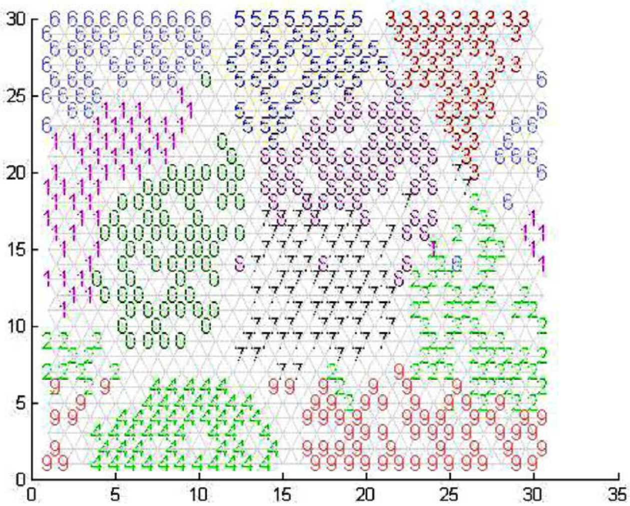


Fig. 10. Graphic map distribution of 10 font according to neurons number (1: Ahsa, 2: Andalus, 3: Arabic_transparant, 4: Badr, 5: Buryidah, 6: Dammam, 7: Hada, 8: Kharj, 9: Koufi and 10: Naskh) by four features, with final quantification error as 0.009 and final topographic error as 0.057.

et al., 2005, and contrary to the method proposed by Zramdini (1998) and Yang et al. (2006). It is script independent and can be used on Latin font discrimination.

To test our method for Latin case, experiments have been conducted on eight kinds of fonts the same ones used on Aviles-Cruz et al., 2005 and Zhu et al. (2001). The images text used in the experimentation are extracted from ALPH-database. All the samples are binary text images with different forms (one or more lines), and with a resolution of 200 pixels/in. Each font is presented by 80 samples. Fig. 11 shows an example of each Latin font used.

The graphic font localization is carried out by using SOM. The main advantages of this technique are explained in Section 4.3. Fig. 12a shows the similar font classes, which can produce overlapping zone, such as (0): Arial, (6): Modern, (7): Times New Roman. The rest of the font classes can be detected easily (Fig. 12b). By using RBF classifier, the obtained results are listed in Table 5. It shows that the accuracy result does not decrease when training samples decrease. The main advantage of our method is that it is able to process heterogeneous images size without pre-processing steps.

5. Robustness test

The study was carried out on two features DCD and BCD to test the robustness of the four factors: size, skew, resolution and noise, not only in the font discrimination context, but also in a more general one.

We established a reference base of 50 texts; these prototypes are reproduced for each font to have the same text (paragraph) in inter font. In this base we retained 16 points as size writings, and 200 dpi as resolution. In total, the reference base consists of 500 prototypes. It will be useful for the development of the transformations carried out for the four factors to study.

We calculated both features presented before with four alternatives $B(15)$, $B(20)$, $D(15)$ and $D(20)$. The study is based on the comparison of the feature values of each image, after applying every

Table 4
Some recognition rates by using RBF classifier.

Training prototype number		800	700	600	500	400	300	200
Test prototype number		200	300	400	500	600	700	800
Font recognition percent rate	AH	98.75	98.57	98.33	98.00	97.50	96.67	95.00
	AN	100	100	98.33	98.00	97.50	96.66	100
	AT	98.75	98.57	100	98.00	97.50	97.50	95.00
	BR	98.75	98.57	98.33	98.00	97.50	96.67	95.00
	BU	98.75	98.57	98.33	98.00	97.50	96.66	95.00
	DA	98.75	98.57	98.33	98.00	97.50	96.67	95.00
	HA	98.75	98.57	98.33	98.00	97.50	100	95.00
	KH	98.75	98.57	98.33	98.00	97.50	96.66	95.00
	KF	98.75	98.57	98.33	98.00	97.50	96.66	95.00
	NS	98.75	98.57	98.33	98.00	97.50	96.66	95.00

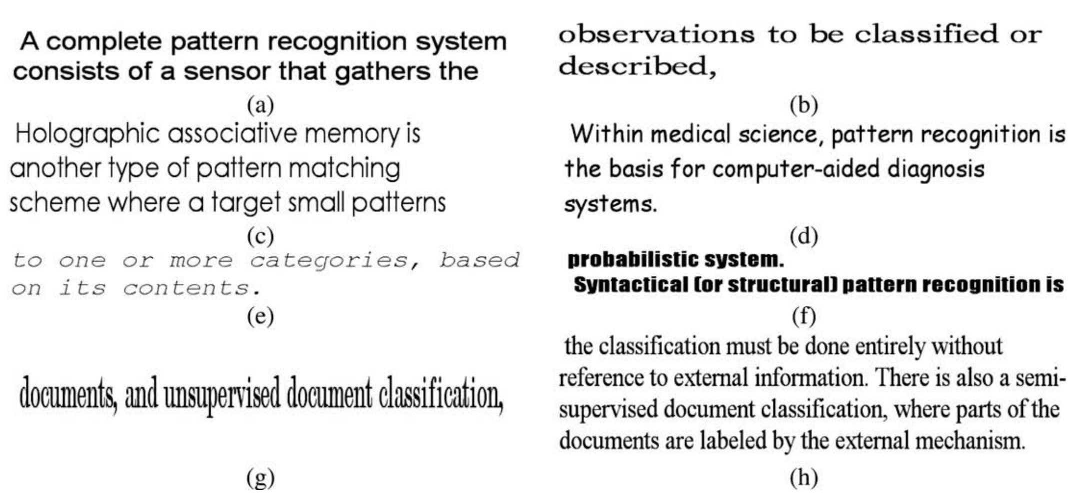


Fig. 11. Example of text images, with various sizes, which can be produced by segmentation process: (a) Arial font, (b) Bookman font, (c) Century font, (d) Comic font, (e) Courier font, (f) Impact font, (g) Modern font and (h) Times New Roman font.

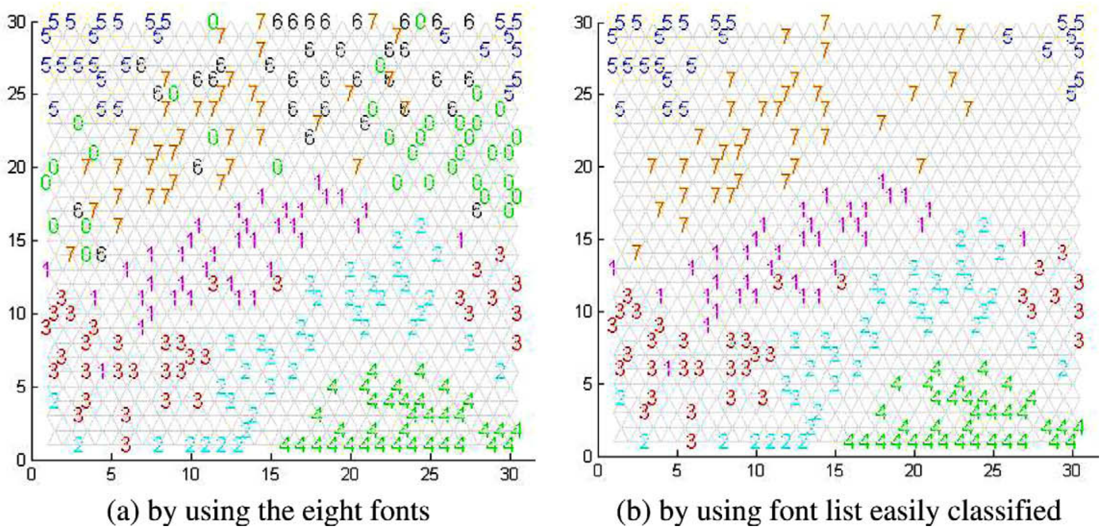


Fig. 12. Graphic map distribution of Latin fonts according the number of neuron in x and y axes: (0) Arial, (1) Bookman, (2) Century, (3) Comic, (4) Courier, (5) Impact, (6) Modern, (7) Times New Roman.

effect, with the corresponding feature values of original images – the reference base. The goal is to study the independence of our features for several effects.

5.1. Size influence

Several works highlighted the writing size influence (Zramdini, 1998; Zramdini and Ingold, 1995). The reference base is reproduced with different sizes of writing, for each font. The selected sizes are 14, 18 and 20.

We can easily detect that the writing size is proportional to the choice of the font. The sizes of the writings are not proportional in inter font. The font sizes higher than 20 points are generally used in ergonomic aspects. The writings of size lower than 14 points are generally used in specific contexts; they are interpreted with difficulty in the majority of the Arabic font. Table 6 shows the variation of our features, according to each font in front of the factor size. We can notice that this residual error depends on the font used.

Table 7 shows the global error rates of our features according to the writing size in the 10 fonts. The presented results show that our features B and D are robust in front of the writing sizes. Fig. 13 shows the various global error rates for each feature on the one hand; and on the other hand, the robustness of BCD compared with that of DCD. The average global rate of residual error of our features is about 0.02%.

5.2. Skew influence

Preliminary experiments consist of studying the influence of the skew of our features. Fig. 12 shows some applied effects on the same image text.

Initially, we applied the five transformations represented by Fig. 14 for each image of the reference base. The original images will be useful to study the skew effect according to the transformed images. Table 8 shows the average error rates of the features for the 10 fonts according to the skew level. The global error rates represented in Table 8 are carried out for the first four

Table 5
Example of the obtained results by increasing/decreasing the number of training and test samples.

Training prototypes number	160	320	240	320	480
Test prototypes number	480	400	320	240	160
Font recognition rate (%)	99.52	99.42	99.28	99.04	98.57

Table 6
Residual error rates for each feature and font in front of size factor.

Font	B(15)	D(15)	B(20)	D(20)
AH (%)	1.44	2.26	1.68	3.01
AN (%)	1.45	0.90	1.17	1.79
AT (%)	0.69	0.56	1.06	1.07
BR (%)	1.87	1.62	2.21	2.47
BU (%)	1.96	1.98	2.56	1.53
DA (%)	1.27	1.55	0.84	0.52
HA (%)	4.55	4.83	3.73	4.15
KH (%)	0.65	1.81	1.71	1.27
KF (%)	1.02	0.82	0.56	0.35
NS (%)	1.18	0.93	0.30	0.94

Table 7
Global residual error for each studied size, for the 10 fonts.

Size	B(15) (%)	D(15) (%)	B(20)(%)	D(20) (%)
14 points	1.49	1.34	1.77	0.93
18 points	1.30	1.32	1.45	1.99
20 points	2.03	2.52	1.52	2.21

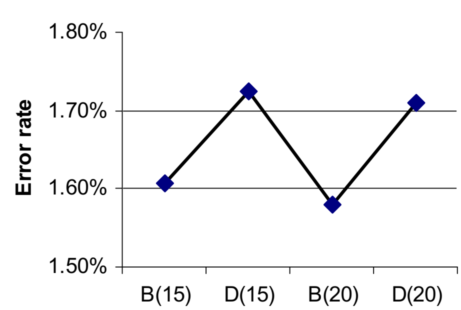


Fig. 13. Residual error for each feature in front of written size.

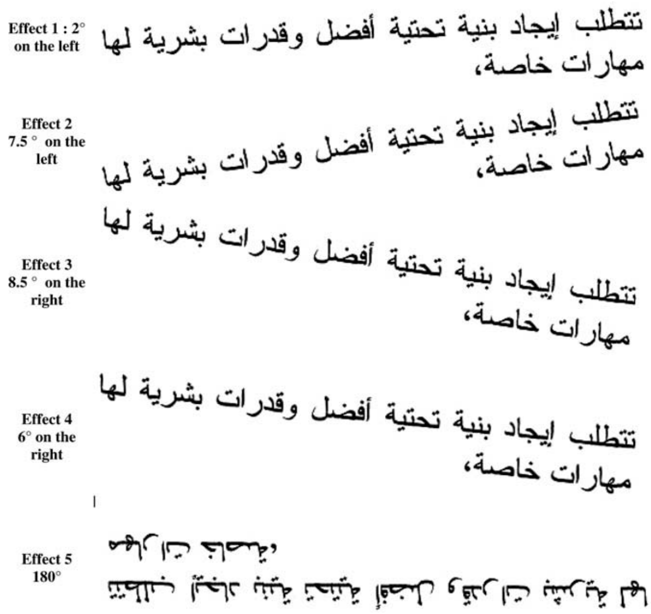


Fig. 14. Some skew and rotation effects applied on the text image.

effects, the fifth effect constitutes the most extreme case of the skew level; it is used to better evaluate the robustness of our features.

We noticed that the obtained errors are practically very weak; so our features are very slightly related to the skew effects. The average residual error is about 0.86%.

Table 8
The average residual errors of each feature according to the skew level.

Effect	$B(15)$ (%)	$D(15)$ (%)	$B(20)$(%)	$D(20)$ (%)
Effect 1	0.12	1.07	0.77	1.13
Effect 2	0.95	0.02	0.35	0.40
Effect 3	1.58	0.41	0.24	0.31
Effect 4	1.88	0.49	0.51	0.39
Effect 5	2.07	1.03	1.45	0.80
Average error	1.13	0.50	0.47	0.56

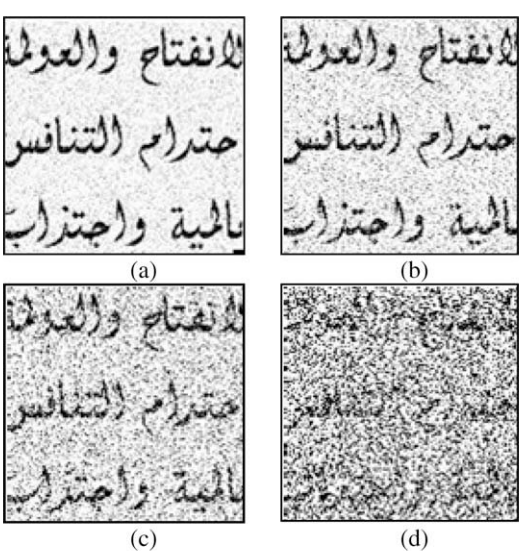


Fig. 15. Example of noising images with various values of SNR: (a) SNR = 50, (b) SNR = 40, (c) SNR = 30 and (d) SNR = 10.

Table 9
Global residual error of features for various SNR values, and for the 10 fonts.

SNR	$B(15)$ (%)	$D(15)$ (%)	$B(20)$(%)	$D(20)$ (%)
50	4.35	2.14	2.66	5.52
40	12.72	6.17	8.55	6.16
30	5.39	32.48	37.34	18.40
20	32.75	56.56	56.80	46.74
10	42.06	59.13	59.28	54.01

5.3. Noise effect

The majority of the proposed methods examined the robustness by comparing the accuracy identifications (Zhu et al., 2001; Aviles-Cruz et al., 2005; Ben Moussa et al., 2006). Consequently, we made a noised image basis of document in order to examine the robustness of B and D features.

Firstly, for each image of the reference basis, we applied a Gaussian noise with five values of SNR (signal-to-noise ratio); an example of the noised images is represented in Fig. 15. Secondly, we calculated the difference between feature values of each noised image (after applying noise effect) and the corresponding of feature values of original images (without noise). Finally, this process is repeated with different SNR ratios for each feature. The result is given in Table 9.

The experimental results show that BCD and DCD depend on the noise level. We noted that feature degradation depends not only on the noise level but also on the font used. An example of the average deviations of the values resulting from B and D for

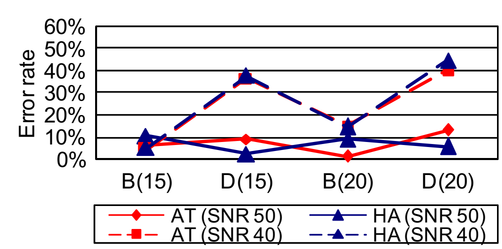


Fig. 16. Example of residual error of B and D features, in intra and inter fonts, according to SNR values.

Table 10
Residual errors of each feature, and for various resolution levels, compared with 200 dpi resolution.

dpi resolution	$B(15)$ (%)	$D(15)$ (%)	$B(20)$(%)	$D(20)$ (%)
150	4.39	5.03	4.07	5.53
100	6.28	8.06	5.84	10.05
80	7.18	10.98	7.79	13.25
60	8.33	15.07	10.36	17.22

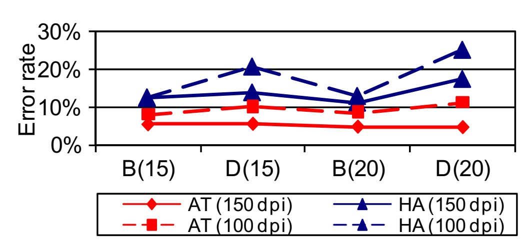


Fig. 17. Example of residual error of B and D features, in intra and inter fonts, according to resolution choice.

Table 11
Comparative study of recent proposed methods.

Technique	Average recognition rate (%)	The number of used features	The number of used typefaces	Binarization, pre-processing	Style, size
In (Zramdini, 1998)	96.91	10	8 Latin	No, No	Yes, Yes
In (Zhu et al., 2001)	98.58	16	8 Latin 6 Chinese	Yes, Yes	Yes, No
In (Aviles-Cruz et al., 2005)	99.9	64	8 Espagnole	No, Yes	Yes, No
In (Yang et al., 2006)	97.2	10	6 Chinese	No, Yes	Yes, No
Ours	96.6	4	10 Arab	Yes, No	Yes, Yes
	99.28	4	8 Latin	Yes, No	Yes, Yes

the two Arabic_Transparent and Hada fonts is presented in Fig. 16. It shows the variation of *B* and *D* in intra and inter fonts, according to SNR.

5.4. The resolution effect

Firstly, we regenerated each paragraph or line of reference text in five levels of resolution (200, 150, 100, 80 and 60). Secondly, we calculated the difference between the obtained feature values for each image in each level of resolution and those of the obtained feature values for a 200 dpi resolution. The degradation rates of four features are represented in Table 10.

We noted that the precision of BCD and DCD depends not only on the used resolution, but also on the studied font. An example of resolution effect in intra and inter fonts is presented in Fig. 17.

6. Discussion and conclusion

We noticed a clear discrimination in certain fonts such as Buryidah, Koufi, Dammam, and Hada and we noticed also that the classification errors are brought by a fusion of similar fonts. The increase of training basis cannot resolve this problem. The obtained results, by training basis increase, show that our method requires some adequate prototypes for the training. Then a better rate of classification is not related to the number of prototypes in the training basis, but it depends on the choice of the training components.

There are few works on Arabic font discrimination. The comparison of our techniques with other existing methods is presented in Table 11. The font recognition accuracy is satisfying since it is obtained from not standardized blocks or lines, but under the natural conditions of an original document, contrary to the other methods in the block level, in which several pre-processing steps are used, such as the works of Aviles-Cruz et al. (2005), Yang et al. (2006) and Zhu et al. (2001).

A new automatic method of font identification is proposed and experimented. This method is based on the global texture analysis, and does not depend on the content studied. The average font discrimination accuracy is about 98% for Arabic script and 99.28 Latin script. According to this method, lines or block texts are not normalized. It does not require a complex processing. Studies of the size, skew, resolution and noise influence of the writings are tested.

We noted that our features are more robust in translation and skew effects than the resolution and noising one. The proposed method can be more efficient for uniformed blocks. The AFB can make easy the comparison between the works of Arabic recognition font.

The main advantages of our system FSAFR are listed below:

- It does not need standardization of blocks, the use of heterogeneous data base (line, paragraph).
- It does not need most prototypes for training base.
- It is script independent, which makes it easier to adapt for other application.
- The feature dimensions are very low.
- Finally and most important is the use of two fractal techniques for covering aspects in 2-D images, and for covering the vision aspect, and their application on font recognition.

Acknowledgments

This work was partially supported by the "Agence Universitaire de la Francophonie" (AUF), within the framework of the "SAuvegarde du Patrimoine Culturel de Civilisation Ancienne Arabe" project. The authors would also like to acknowledge the financial support of this work by grants from the General Direction of Scientific Research and Technological Renovation (DGRSTRT), Tunisia, under the ARUB 01/UR/11/02 program.

References

Ajay, B.K., Jibitesh, M., 2001. On calculation of fractal dimension of images. Pattern Recognition Lett. 22, 631–637.
Alimi, Adel M., 1995. Evolutionary neuro-fuzzy approach to recognize on-line Arabic handwriting. In: ICDAR, vol. 1, pp. 382–386.
Alimi, Adel M., 1997. A neuro-fuzzy approach to recognize Arabic handwritten characters. In: Proc. IEEE Internat. Conf. on Neural Networks, vol. 3, pp. 1397–1400.
Alimi, A.M., 2002. Evolutionary computation for the recognition of on-line cursive handwriting. IETE J. Res. 48 (5), 385–396.
Amin, A., 1998. Off-line Arabic character recognition: The stat of the art. Pattern Recognition 31 (5), 517–530.
Amin, A., 2000. Recognition of printed Arabic text based on global features and decision tree learning techniques. Pattern Recognition 33, 1309–1323.
Aviles-Cruz, C., Rangel-Kuoppa, R., Reyes-Ayala, M., Andrade-Gonzalez, A., Escarela-Perez, R., 2005. High-order statistical texture analysis font recognition applied. Pattern Recognition Lett. 26, 135–145.
Ball, G., Srihari, S.N., Srinivasan, H., 2006. Segmentation-free and segmentation-dependent approaches to Arabic word spotting. In: Proc. IWFHR-10, La Baule, France.
Belaïd, A., Rangoni, Y., Alusse, A., Rangoni, Y., Cecotti, H., Farah, F., Gagean, N., Fiala, D., Rousselot, F., Vigne, H., 2005. Document retro-conversion for personalized electronic reedition. In: Proc. Internat. Workshop on Document Analysis, Allied Publishers Pvt. Ltd., 2006, Kolkata, India.
Ben Amara, N., Gazzah, S., 2004. Une approche d'identification des fontes arabes. In: Conférence Internationale Francophone sur l'Ecrit et le Document, CIFED'04, La Rochelle, France, Juin 2004.
Ben Moussa, S., Zahour, A., Alimi, M.A., Benabdelhafid, A., 2005. Can fractal dimension be used on font classification. In: 8th ICDAR, Seoul, Korea, vol. 1, pp. 146–150.
Ben Moussa, S., Zahour, A., Kherallah, M., Benabdelhafid, A., Alimi, M.A., 2006. Utilisation de nouveaux paramètres à base de fractale pour la discrimination des fontes Arabes. In: CIFED'06, Suisse, Fribourg, 18–22 September, 2006, pp. 282–287.
Ben Moussa, S., Zahour, A., Benabdelhafid, A., Alimi, M.A., 2008. Fractal-based system for Arabic/Latin, printed/handwritten script identification. In: 19th Internat. Conf. on Pattern Recognition, ICPR'08, Tampa, Florida, USA, 8–11 December, 2008.
Cetera, A., 2001. The Minkowski dimension and critical effects in fractal evolution of defects. Chaos, Solitons Fract. 12, 475–482.
Cooperman, R., 1997. Producing good font attribute determination using error-prone information. Internat. Soc. Opt. Eng. J. 3027, 50–57.
Ding, X., Chen, L., Wu, T., 2007. Independent font recognition on a single chinese character. IEEE Trans. Pattern Anal. Machine Intell. 29 (2), 195–204.
Essoukri Ben Amara, N., Bouslama, F., 2003. Classification of Arabic script using multiple sources of information: State of the art and perspectives. Internat. J. Document Anal. Recognition 2, 195–212.

Hamdani, T.M., Alimi, A.M., Karray, F., 2006. Distributed genetic algorithm with bi-coded chromosomes and a new evaluation function for feature selection. In: IEEE Cong. on Evolutionary Computation, pp. 581–588 (Article number 1688362).
Jin, X.-C., 1995. A practical method for estimating fractal dimension. Pattern Recognition Lett. 16, 457–464.
Kahan, S., Pavlidis, T., 1987. On the recognition of printed characters of any font or size. IEEE Trans. Pattern Anal Machine Intell. 9 (2), 274–288.
Kanoun, S., Ennaji, A., Lecourtier, Y., Alimi, A.M., 2002. Script and nature differentiation for Arabic and Latin text images. In: 8th Internat. Workshop on Frontiers in Handwriting Recognition, IWFHR'02, Ontario, Canada, August 2002, pp. 309–313.
Kasparis, T., Tzannes, N.S., Bassiouni, M., Chen, Q., 1995. Texture description using fractal and energy features. Computers Elect. Engng. 21 (1), 21–32.
Kherallah, M., Bouri, F., Alimi, M.A., 2009. On-line arabic handwriting recognition system based on visual encoding and genetic algorithm. Eng. Appl. Artif. Intell. 22, 153–170.
Khorsheed, M.S., 2007. Offline recognition of omnifont Arabic text using the HMM ToolKit (HTK). Pattern Recognition Lett. 28 (12), 1563–1571.
Maergner, V., Pechwitz, M., El-Abed, H., 2005. Arabic handwriting recognition competition. In: Proc. of 8th Internat. Conf. on Document Analysis and Recognition, Seoul, Korea, August 29th–September 1st, pp. 70–74.
Massoudy, H., 1998. Calligraphie arabe vivante. Flammarion, Paris.
Moalla, I., LeBourgeois, F., Emptoz, H., Alimi, A.M., 2006. Contribution to the discrimination of the medieval manuscript texts: Application in the palaeography. Lect. Notes Comput. Sci. – LNCS 3872, 25–37.
Neil, G., Curtis, K.M., 1997. Shape recognition using fractal geometry. Pattern Recognition 30 (12), 1957–1969.
Peitgen, H., Jûrgens, H., Saupe, D., 1992. Chaos and Fractals New Frontiers of Science. Springer-Verlag, New York, Berlin, Heidelberg.
Shi, H., Paulidis, T., 1997. Font recognition and contextual processing for more accurate text recognition. In: Proc. ICDAR'97, Ulm-Germany, pp. 39–44.
Sun, H.M., 2006. Multi-linguistic optical font recognition using stroke templates. In: ICPR'06, pp. 889–892.
Versanto, J., Himber, J., Alhoniemi, E., Parhankangas, J., 1999. Self organizing map in matlab: The Somtoolbox. In: Proc. of the Matlab DSP Conference, Finland, pp. 35–40.
Yang, Z., Yang, L., Qi, D., Suen, C.Y., 2006. An EMD-based recognition method for Chinese fonts and styles. Pattern Recognition Lett. 27, 1692–1701.
Zaghden, N., Ben Moussa, S., Alimi, A.M., 2006. Reconnaissance des fontes arabes par l'utilisation des dimensions fractales et des ondelettes. In: Conférence Internationale Francophone sur l'Ecrit et les Documents, CIFED'06, Suisse, Fribourg, 18–22 Septembre, 2006, pp. 277–282.
Zengl, X., Koehl, L., Vasseur, C., 2001. Design and implementation of an estimator of fractal dimensions using fuzzy technique. Pattern Recognition 34, 151–169.
Zhang, P., Bui, T.D., Suen, C.Y., 2002. Recognition of similar objects using 2-D wavelet-fractal feature extraction. In: Proc. ICPR'02, vol. 2, pp. 316–319.
Zhou, G., Nina, S.-N.L., 2005. A comparison of fractal dimension estimators based on multiple surface generation algorithms. Comput. Geosci. 31 (10), 1260–1269.
Zhu, Y., Tan, T., Yunhong, W., 2001. Font recognition based on global texture analysis. Trans Pattern Recognition Machine Intell. 23 (10), pp 1192–1200.
Zramdini, A., 1998. Optical font recognition using typographical features. Trans. Pattern Anal Machine Intell. 20 (8), 877–882.
Zramdini, A., Ingold, R., 1995. A study of document image degradation effects on font recognition. In: Proc ICDAR'95, vol. 2, August 1995, pp. 240–743.