Fractal and Multi-Fractal Dimensions For Farsi/Arabic Font Type and Size Recognition

Akram Alsadat Hajiannezhad
Electrical and Computer Engineering Department
Semnan University
Semnan, Iran
a_hajiannezhad@sun.semnan.ac.ir

Saeed Mozaffari
Electrical and Computer Engineering Department
Semnan University
Semnan, Iran
mozaffari@semnan.ac.ir

Abstract—**In this paper, a new method based on fractal geometry is proposed for Farsi/Arabic font recognition. The feature extraction does not depend on the document contents which considers font recognition problem as texture identification task. The main features are obtained by combining the BCD, DCD, and DLA techniques. Dataset includes 2000 samples of 10 typefaces, each containing four sizes. The average recognition rates obtained for these 10 fonts and 4 sizes (40 classes) using RBF and KNN classifiers are 96% and 91% respectively. The dimension of feature vectors extracted by the proposed fractal approach is very low. This property obviates the need for numerous training samples. Experimental results show that this algorithm is robust against skew. Simultaneously identifying type and size of the font is the most important innovation of this paper.**

Keywords-Optical Font Recognition (OFR), Fractal Dimensions (FD), Box Counting Dimension (BCD), Dilation Counting Dimension (DCD), Diffusion Limited Aggregates (DLA)

I. INTRODUCTION

Everyone who buys a scanner will probably receive an optical character recognition (OCR) product too. An OCR system converts scanned text images into machine-encoded form. Every OCR system is made of several parts. The most important and difficult module is character recognition block. Primary OCR systems didn't have the ability to work with different fonts and sizes. Prevalence of documents containing various fonts, sizes, and even different languages intensified the need of new OCR systems which are capable to deal with such variety.

Performance of OCR systems strongly depends on font recognition part. Although font recognition is becoming an integral part of every OCR system in other languages, but it is still undeveloped in Farsi OCR systems. This is mainly due to complexity of Farsi script. To get around this problem, some methods for Farsi font and size identification are presented in this paper.

II. POPULAR FONTS IN FARSI DOCUMENTS

Farsi documents like books, journals, official letters, and blogs contain widely used fonts such as: Lotus, Mitra, Nazanin, Traffic,Yaghut, Zar, Homa, Titr, Tahoma, and Times New Roman [1]. These fonts usually have sizes vary from 9 to 20. Since performance of an OCR system without the ability of font recognition decrease significantly, this paper devoted to Farsi optical font recognition (OFR). Experimental results are carried out on the mentioned 10 special fonts with 4 different sizes.

III. RELATED WORKS

There are several techniques have been proposed for font recognition in other languages. But the number of papers in Farsi and Arabic font identification are few. This section reviews shortly related works on Farsi and Arabic OFR systems. In [1], Sobel and Roberts gradients in 16 directions (called SRF) were used for Farsi font recognition.

Since fractal geometry is useful for modelling complex random experiments, it has been applied to various fields, such as mathematics, physics, biology and etc. Recently, researches demonstrate that fractal dimensions are useful to quantify the complexity of an image and constitute an excellent geometrical description. In [2] a new fractal based approach for the recognition of Arabic fonts was proposed, in which font recognition performs in block or paragraph level. The proposed methods are based on BCD and DCD features to extract visual aspect from image text.

In the following sections, we will explain our method which uses three techniques of fractal dimension (FD) to obtain features for font recognition. The novelty of this paper is the use of DLA feature besides previously proposed BCD and DCD features. Adding this new fractal feature increases the system's performance. The algorithm is in block level and 512×512 text blocks are used in this experiment.

IV. BINARIZATION

Among three fractal dimension methods proposed in this paper, two of them (BCD and DLA) can be applied only on binary images. So, a short description of image binarization is presented in this section.

In every binarization algorithm, first, a threshold is computed. Then, pixels with higher intensity than that threshold, are labeled as foreground (object), otherwise they are considered as background. According to [7], thresholding

978-1-4577-1535-8/11/$26.00 ©2011 IEEE

methods can be divided into two categories: global and local (adaptive) approaches. In global methods a unique threshold is computed and applied to the whole image. Among proposed global thresholding algorithms, Otsu's method is widely used because of easy implement and fast runtime. Fig. 1 shows a gray-level image and the results of different binarization methods. Fig. 1 (a), (b) and (c) are the results of global thresholding.

For binarizing images with complex background, global thresholding methods are not suitable. In this case, local thresholding algorithms are usually utilized. Among local algorithms, Niblack's method is simple and effective and best suits for character recognition. So, we used Niblack's method for text image binarization in this paper. Fig. 1 (e) shows the result of Niblack's, which outperforms the others.

V. THE PROPOSED FEATURE EXTRACTION

Feature extraction is a crucial step in any pattern recognition system. It is responsible for measuring the features of objects in an image.

As mentioned before, fractal dimension (FD) is proper for modelling a complex image [2]. Because of the complexity of Farsi font recognition, FD methods are used for feature extraction in this paper. Our work is based on the calculation of fractal and multi-fractal dimension for each text image.

First of all, fractal dimension is calculated by BCD method. Then, multi-fractal dimensions are measured by the DLA, and DCD methods, as will be described in the next sections. After extracting these features, they will be concatenated to construct the main feature vector.

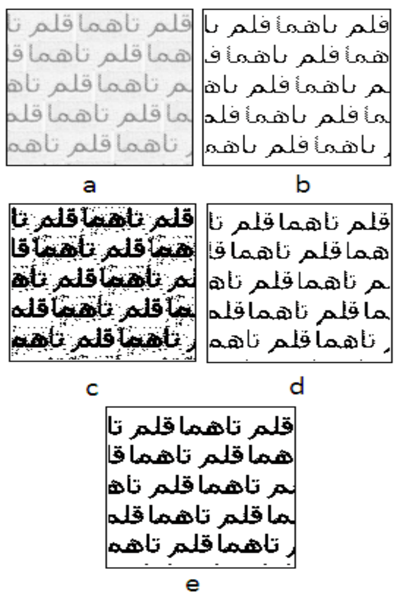

Figure 1. (a) gray-level document image, (b) and (c) binary images generated by MATLAB, (d) binary image obtained with Otsu's method, (e) binary image obtained with Niblack's method.

VI. BOX COUNTING DIMENSION (BCD)

Box counting dimension is one of the most widely used fractal dimensions. Its popularity is largely due to relatively simple mathematical calculation and estimation. To estimate the BCD, according to equation (1), the text block is divided into a grid of boxes of size ‘r’. Then, the numbers of boxes which are not empty are counted. In the next step, the two previous steps must be repeated for different amounts of ‘r’, and a graph of $\log N(r)$ versus $\log(1/r)$ for each size is produced. And finally, BCD is estimated by linear regression between $\log N(r)$ and $\log(1/r)$, through equation (1).

$$\mathrm{BCD}(r) = \lim_{r\to 0}\frac{\log N(r)}{\log(1/r)} \tag{1}$$

VII. MULTI-FRACTAL DIMENSIONS

There are many fractal objects in nature which fractal dimension is not precise enough to describe their complexity. In this situation, instead of using fractal dimension, multi-fractal dimension is usually used. There are a lot of multi-fractal dimension methods in fractal geometry. Among them, we chose two for font recognition which will be explained in the following sections.

A. *Diffusion Limited Aggregates (DLA)*

Calculation of DLA is very easy and the main steps of this algorithm are presented as follows [3]:

First of all, those pixels containing information should be individuated in (M_0) matrix. Second, about 100 pixels from the pixels containing information (M_0) should be selected randomly and some boxes of the radius Ri should be centered on them. Third, number of pixels containing information in every box of radius Ri is calculated. And then, While Ri < R_{max} calculate the average number of pixels containing information in all the boxes of the same radius Ri. After that, calculate the average number of pixels containing information in all boxes for all radiuses. If number of simulation<50 then go to the second step. Else, calculate the average for these 50 simulations to obtain the multi-fractal dimension. Finally, DLA can be estimated by linear regression between $\log[(\frac{M(R)}{M_0})^{q-1}]$ and $\frac{1}{q-1}\times\log\left[\frac{R}{L}\right]$, by equation (2).

$$\begin{cases} \text{If } q = 1 \text{ then} & D_1 \approx \dfrac{(D_{1-c} + D_{1+c})}{2} \\ \text{Else} & D_q = \dfrac{1}{q-1}\dfrac{\log[(\frac{M(R)}{M_0})^{q-1}]}{\log[\frac{R}{L}]} \end{cases} \tag{2}$$

B. *Dilation Counting Dimension (DCD)*

DCD is another multi-fractal dimensions, which is based on morphological dilation operation. For measuring this dimension through equation (3).

First, an edge detection algorithm is performed on the input image. Second, choose a number as maximum dilation radius and call it 'd'. Third, for every dilation radius smaller than 'd', dilate the edge detected image obtained from the first step. Then, for each dilated image use the BCD algorithm for r=1, in this algorithm V(d) and N(r) are the same. And finally, DCD can be estimated using a linear regression between log V(d) and log (d), the parameter 'n' shows the dimension of a space. We set n=2 in our algorithm because the input signal is an image with two dimensions.

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

VIII. Robustness Against Skew

Document skew is a distortion that may encounter in every OCR and OFR system. Since techniques used in these systems only work on aligned input images, skew correction became an inseparable part of these systems.

One of the most important advantages of our algorithm is its robustness against skew. Because rotating an object in the space, does not change its dimension. This is true for all fractal dimension measurements and we won't need skew correction part in our algorithm.

As shown in Fig. 2, the fractal dimension of text images in parts (a), (b) ,(c), and (d) are the same. Using fractal methods for feature extraction, not only makes the system robust against skew but also obviating the need for skew correction module.

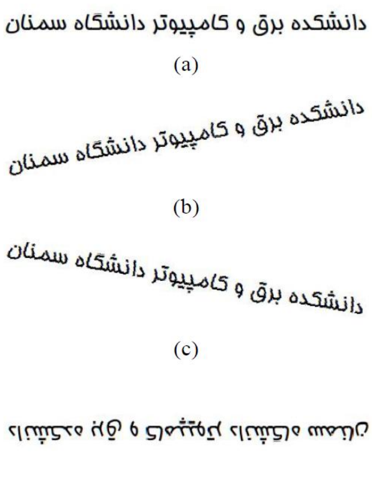

Figure 2. Rotated images, (a) 0°, (b) 20°, (c) -20°, (d) 180° deviation.

IX. Experimental Results

The experimental results were carried out by RBF and KNN classifiers to classify the 10 Farsi fonts with 4 sizes which constitute 40 classes. We performed our proposed algorithm on a 2000 samples of Farsi dataset, that each font presented by 200 samples. We used 150 samples for training and 50 samples for testing. In our algorithm, each sample is expressed by a 6D feature vector. As mentioned before we used three different FD algorithms in our work, namely BCD, DCD, and DLA, with several amount of 'r', 'd', and 'q'. The utilized features for Farsi font recognition are BCD(8) , BCD(32) , DCD(4), DCD(6), DLA(10) in this paper.

Since the sizes of text blocks used in this experiment are 512×512, in most cases each text block image contains most of the Farsi characters. In DLA algorithm, the amount of D_q for all $-20 < q < 20$ is the same. So, we chose q=10 in this experiment. For calculating DLA(10), since a constant number of q was chosen, for different radius Ri=2, 4, 6, ..., 32 we only need calculating linear regression between $\log[(\frac{M(R)}{M_0})^{q-1}]$ and $\log[\frac{R}{L}]$. In fact, we used the slope and intercept of this linear regression for two DLA features.

The confusion matrix for font size 16 and the RBF classifier is presented in TABLE I, As it can be seen, good discrimination rate is achieved and few errors occur among prototypes in Nazanin, Zar, Lotus and Mitra classes.

A graphical figure for showing the discrimination capacity of our algorithm was shown in Fig. 3. Since 6 features were extracted from each prototype and the figure is two dimensional, some overlaps are presented in Fig. 3.

The average discrimination rates of the 40 classes (the 10 fonts and 4 sizes) are presented in TABLE II. The average discrimination rates are about 96% using RBF, and 91% for KNN classifier. Using all six extracted features, the discrimination rates will be increased to 98% using RBF, and 94% using KNN classifiers

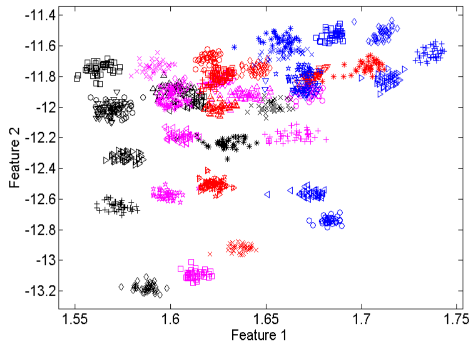

Figure 3. Graphic localization of 10 fonts with 4 sizes using two features out of six features.

TABLE I. The Confusion Matrix of 10 Fonts with Size 16 by RBF

True / Estimated	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*Total*
1	47	0	0	0	0	2	0	0	0	0	49
2	0	50	1	0	0	0	0	0	0	0	51
3	2	0	47	0	0	2	0	0	0	0	51
4	0	0	0	50	0	0	0	0	0	0	50
5	0	0	0	0	49	0	0	0	0	0	49
6	1	0	2	0	0	46	0	0	0	0	49
7	0	0	0	0	0	0	49	0	0	0	49
8	0	0	0	0	0	0	0	50	0	0	50
9	0	0	0	0	0	0	1	0	50	0	51
10	0	0	0	0	1	0	0	0	0	50	51
Total	50	50	50	50	50	50	50	50	50	50	500

To explore the effect of each fractal dimension, we trained these two classifiers with different feature sets. TABLE II, Fig. 4 and Fig. 5 show these results. According to them, multi-fractal features are the most important features for Farsi/Arabic font recognition.

TABLE II. THE AVERAGE DISCRIMINATION RATES OF THE 10 FONTS AND THEIR 4 DIFFERENT SIZES (12-18).

Classifiers / Features	*RBF*	*KNN*
BCD,DLA,DCD	%96	%91
BCD,DLA	%80	%75
DLA,DCD	%90	%84
BCD,DCD	%82	%78

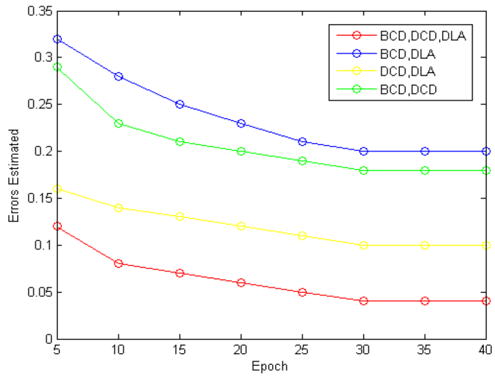

Figure 4. Error estimation using RBF classifier.

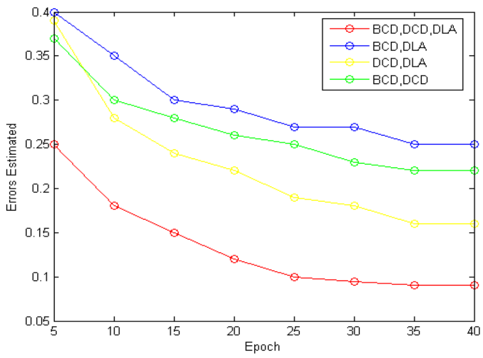

Figure 5. Error estimation using KNN classifier.

X. CONCLUSION

In this paper, we proposed a new algorithm for font recognition in Farsi document images. Our proposed features are based on combination of one fractal dimension (BCD), and two multi-fractal dimensions (DCD and DLA). With these features, each document image is converted into a 6D feature vector. The advantages of FD methods over previous approaches are: low dimensional feature vector, low computational complexity, and robust against skew. The proposed method not only identifies the font type but also recognizes its size.

REFERENCES

[1] Hossein Khosravi, and Ehsanollah Kabir, “Farsi font recognition based on Sobel–Roberts features,” Elsevier, Pattern Recognition Letters 31 (2010) 75–82.

[2] Sami Ben Moussa, and Abderrazak Zahour, and Abdellatif Benabdelhafid, and Adel M. Alimi “New features using fractal multi-dimensions for generalized Arabic font recognition,” Elsevier, Pattern Recognition Letters 31 (2010) 361–371.

[3] Aymen Chaabouni, and Houcine Boubaker, and Monji Kherallah, and Adel M. Alimi and Haikal El Abed, “Fractal and Multi-Fractal for Arabic Offline Writer Identification,” IEEE Trans, 2010 International Conference on Pattern Recognition.

[4] R. Lopes, and N. Betrouni, “Fractal and multifractal analysis: A review”, Elsevier. , Medical Image Analysis 13 (2009) 634–649

[5] Edith Perrier, and Ana M. Tarquis, and Annette Dathe, “A program for fractal and multifractal analysis of two-dimensional binary images: Computer algorithms versus mathematical theory,” Elsevier ,Geoderma 134 (2006) 284–294

[6] Ajay, B.K., and Jibitesh, M, “On calculation of fractal dimension of images”, 2001.. Pattern Recognition Lett. 22, 631–637.

[7] Morteza Valizadeh · Ehsanollah Kabir, "Binarization of degraded document image based on feature space partitioning and classification", IJDAR, DOI 10.1007/s10032-010-0142-4, Springer-Verlag 2010.

[8] Herbert F. Jelinek, Eduardo Fernandez, "Neurons and fractals: how reliable and useful are calculations of fractal dimensions?", Journal of Neuroscience Methods 81 (1998) 9–18, Elsevier.

[9] Yong Xia, Student Member, IEEE, (David) Dagan Feng, Fellow, IEEE, and Rongchun Zhao, "Morphology-Based Multifractal Estimation for Texture Segmentation", IEEE Transactions on Image Processing, Vol. 15, No. 3, March2006.