Character Representation and Recognition Using Quadtree-based Fractal Encoding Scheme

Saeed Mozaffari
Electrical Engineering Department, Amirkabir University of Technology, Tehran, Iran.
s_mozaffari@ aut.ac.ir

Karim Faez
Electrical Engineering Department, Amirkabir University of Technology, Tehran, Iran.
kfaez@ aut.ac.ir

Majid Ziaratban
Electrical Engineering Department, Amirkabir University of Technology, Tehran, Iran,
M_ziaratban@ aut.ac.ir

Abstract

In this paper we proposed a new method for isolated handwritten Farsi/Arabic numerals recognition using quadtree-based fractal representation and iterated function system. Fractal codes represent affine transformations which when iteratively applied to the range-domain pairs in an arbitrary initial image, the result is close to the given image. Each fractal code consists of six parameters such as the corresponding domain coordinates for each range block, brightness offset and an affine transformation. Based on fixed point theorem in iterated function system, we introduced fractal transformation classifier for optical character recognition. We also used Euclidean distance between fractal codes of a query image and fractal codes of all images in the database as a measure of distance for classification. Since fractal codes have different lengths, we applied PCA algorithm to normalize their lengths.

There are ten digits in Farsi/Arabic language and since two of them are not used in Iran postal codes, therefore 8 classes are needed for digits. By using fractal codes with nearest neighbor classifier and fractal transformation, the recognition rate of 92.6% is obtained on our numeral database which contains 480 samples per digit and was gathered from more than 200 people with different ages and different educational background.

1-Introduction

The recognition of handwritten numerals is a challenging problem in pattern recognition. This is due to large diversity of writing styles and image quality. English, Chinese and Kanji isolated handwritten character recognition have long been a focus of study, but a little researches have been done on Farsi and Arabic. Some previous works on recognition of isolated characters, words and scripts of Farsi and Arabic language have used structural features [1] [2] moment features [3] fractal feature [4] and wavelet feature [5,7]. Neural Networks [4] Hidden Markov Models (HMM) [6] and Support Vector Machines (SVM) [7,8] have been used as classifiers in these systems.

The fundamental principle of fractal theory consists of the representation of an image by a contractive transform of which the fixed point is close to the original image. Recently its potential in image recognition has been explored [9]. In [10] the input face image is transformed and its fractal model is then compared against the database of fractal models of known faces. Feed-forward neural networks have been utilized to implement the recognition part. Fractal transformation has been used in the application of face recognition in [11].

In [12] and [13] fractal feature and fractal neighbor distances or fractal dimensions have been used for face recognition. A writer identification system has been utilized based on fractal construction of a reference base in [14].

In this study we represent the image of input character by fractal codes which are obtained by encoding algorithm. Based on fixed point theorem in iterated function system, we introduced fractal transformation classifier for optical character recognition. We also used Euclidean distance between fractal codes of a query image and fractal codes of all images in the database as a measure of distance for classification. Since fractal codes are sensitive to translation, scaling and rotation, some preprocessing is required. For a given binary image containing a single numeral, two pre-processing tasks are needed to make the system invariant to scale and frame size changes. To remove any differences due to location of numeral within the image, the bounding rectangle box of each character or numeral is found. Then this bounding box is scaled to 64×64 pixel image, for scale normalization.

Proceedings of the 2005 Eight International Conference on Document Analysis and Recognition (ICDAR'05)
1520-5263/05 $20.00 © 2005 IEEE

2. Fractal Image Theory

With the advent of information age, the need for mass information storage and retrieval grows. Different image compression methods have been focused on for a long time to reduce this massive information, among them fractal image coding has drawn much attention recently. Fractal compression is a technique that has been developed by Y.Fisher [15]. The fundamental principle of fractal coding consists of the representation of any image by a contractive transform in which the fixed point is too close to the original image. A transformation w is said to be contractive if for any two points P_1, P_2, the distance

$$d(w(P_1), w(P_2)) < sd(P_1, P_2) \qquad (1)$$

for some $s < 1$.

Fractal encoding is based on the concepts and mathematical results of iterated function systems (IFS). The contractive mapping fixed point theorem says that if a map is contractive then when we apply it repeatedly starting with any initial point we converge to a unique fixed point. On the other word a fixed point $|w| \in X$ is a fixed point if satisfies

$$w(|w|) = |w| = \lim_{n\to\infty} w^n(x) \qquad (2)$$

Fractal compression became a practical reality with the introduction by Jacquin of the partitioned IFS (PIFS), which differs from an IFS in that each of the individual mappings operates on a subset of the image, rather than on the entire image [15].

2.1. Fractal Encoding

The procedure for finding a fractal model for a given image is called encoding, compression, or searching for a fractal image representation. An image to be encoded is partitioned into some collection of Ranges R_i; then for each R_i, seek for some collection of image pieces a D_i (Domain block) that has a low rms error when mapped to R_i. If we know R_i and D_i, then we can determine a_i, b_i, c_i, d_i, e_i and f_i in Equation (3).

$$w_i \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a_i & b_i & 0 \\ c_i & d_i & 0 \\ 0 & 0 & s_i \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} e_i \\ f_i \\ o_i \end{bmatrix} \qquad (3)$$

The parameters (x,y) indicate the coordinates of a image pixel and z its gray level. The coefficients a_i, b_i, c_i and d_i determine the geometrical spatial transformation, e_i and f_i determine the translation, o_i the contrast and s_i the luminosity. Fractal encoding process replaces the image by the system of transformations each defined by 8 parameters and position of associated R_i and D_i.

2.2. Partitioning

Various schemes of fractal image compression were proposed, which differ in portioning method, composition of domain blocks pool, class of transformation or type of search used in locating suitable domain blocks. Fig.1 shows some partitioning schemes.

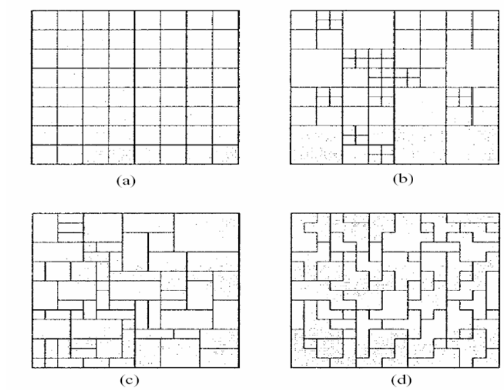

Fig.1. Different partitioning schemes
(a) Fixed block size. (b) Quadtree. (c) Horizontal-vertical. (d) Irregular partition.

In this paper we used quadtree partitioning method for finding range blocks. A quadtree partition is a representation of an image as a tree in which each node, corresponding to a square portion of the image, contains four subnodes, corresponding to the four quadrants of the square. The root of the tree is the initial image. See Fig.2.

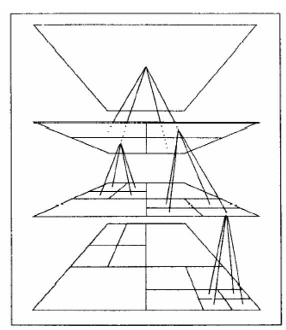

Fig.2. Representation of an image as a tree in quadtree partitioning

A weakness of fixed block size partitioning is the use of fixed-size R_i, since there are regions of the image that are difficult to cover well this way.

Proceedings of the 2005 Eight International Conference on Document Analysis and Recognition (ICDAR'05)
1520-5263/05 $20.00 © 2005 **IEEE**

2.2.1. Partitioning Algorithm

The ranges are selected as follows:

1-After some initial number of quadtree partitions are made (corresponding to a minimum tree depth), the squares at the nodes are compared with domains from the domain library D, which are twice the range size.

2-The pixels in the domain are averaged in groups of four so that the domain is reduced to the size of range.

3-The affine transformation of the pixel values is found that minimizes the rms difference between the transformed domain pixel values and the range pixel values.

4- All the potential domains are compared with a range.

5-If the resulting optimal rms value is above a preselected threshold and if the depth of the quadtree is less than a preselected maximum depth, then the range square is subdivided into four quadrants.

6- The process is repeated.

Fig.3 shows the effect of minimum and maximum depth (minimum and maximum number of times that the largest square sub-image will be recursively quadtree partitioned) parameters in the partitioning scheme.

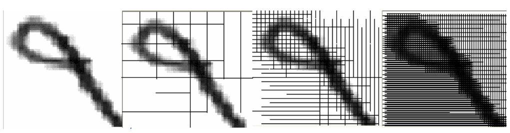

Fig.3. Effect of minimum and maximum depth parameters in the partitioning scheme.

2.3. Fractal Features

We used a set of geometrical parameters $(a_i, b_i, c_i, d_i, e_i, f_i)$, contrast (o_i) and luminescence (s_i) in a feature vector. However the size of each vector varies from one image to another depends on the size of image , image complexity and minimum and maximum depth. In order to normalize the size of each vector, we used Principal Component Analysis (PCA). Maximum and minimum of obtained fractal feature vector was 1752 and 728 respectively.

By using of PCA algorithm, all fractal feature vectors are normalized to the length of 240.

In [13] domain numbers for each range block are used in a feature vector. They also used the quadtree partitioning geometry and applied each feature vector value at its geometrical position to normalize the size of each vector. Based on fractal theory, our method is robust to scaling, translation and rotation changes.

The original image is scaled, rotated and translated. Then the obtained fractal codes from each of above changes applied to the initial image. Fig.4 shows these results after one iteration.

3. Classification

For classification we used two different methods. Fractal codes with nearest neighbor classifier and fractal transformation classifier which will be described bellow have been used in this study.

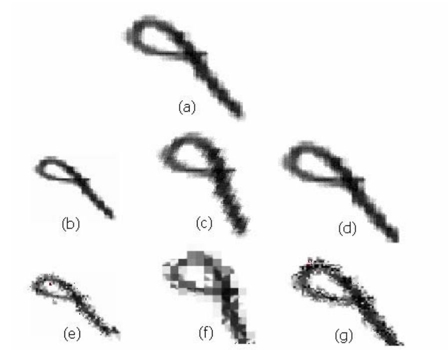

Fig.4. Robustness of the system to scale, rotation and translation changes.
(a) Original image. (b) scaled image .
(c) Rotated image. (d) Translated image.
(e) Decoded image in (b) after 1 iteration.
(f) Decoded image in (c) after 1 iteration.
(g) Decoded image in (d) after 1 iteration.

3.1. Fractal nearest Neighbor Classifier

For this classifier we used the Peak Signal-to-Noise Ratio (PSNR) between fixed-size fractal codes of the query image and normalized fractal codes of all images in the database as a measure of distance, and a nearest neighbor classifier.

3.2. Fractal Transformation

Fractal Transformation classifier is directly based on the concepts and mathematical results of iterated function systems (IFS). The contractive mapping fixed point theorem says that if a transformation is contractive, then when applied repeatedly starting on any initial point, it converges to a unique fixed point. Fig.5. shows fixed point theorem results. The initial image was white.

We applied fractal codes of all images in the database which are obtained by coding algorithm on an input image. After one iteration, by measuring minimum Euclidean distance between input and decoded images, our system can classify the query image. Distortion between input and decoded images, after one iteration, varies by the minimum and maximum depth, and it is shown in Fig.6.

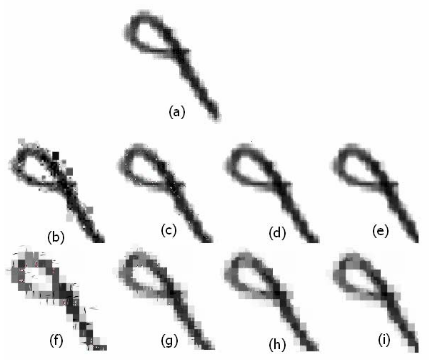

Fig.5. Fixed point theorem results.
First row: input image.
Second row: decoded images after one, second, third and tenth iteration with maximum depth=6 and minimum depth=4.
Third row: decoded images after one, second, third and tenth iteration with maximum depth=4 and minimum depth=1.

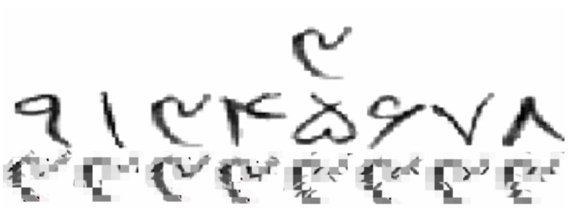

Fig.6. . First row: input image.
Second row eight samples in the training set.
Third row: results of applying fractal codes of input image on second row images after one iteration.

Fig.6 shows that the original image has the minimum distortion among all after decoding process.

4. Experimental Results

In Farsi language, there are ten digits that are shown in Fig.7. Because of similarity between “۵” and “۰” and also between “۳” and “۲”, especially in the handwritten texts, 8 numerals are used in postal codes in Iran. Thus, we have 8 different classes for digits.

۰	۱	۲	۳	۴	۵	۶	۷	۸	۹
0	1	2	3	4	5	6	7	8	9

Fig.7. Digits in Farsi and English

Training and test sets, were gathered from more than 200 people with different ages and different educational background. Our database contains 480 samples per digit (total of 3840). We used 280 samples of each digit for training and the rest (200) for test. Fig.8. shows typical digits in our database. Poorly written numerals in the test set are also included.

Fig.8. Some of samples in our database which were difficult to interpret.

Table.1. Experimental results of different methods

Classifier	Data set	Recognition Rate
Fractal Nearest Neighbor	Train set	100%
	Test set	86.3%
Fractal Transformation	Train set	96.2%
	Test set	90.6%

5. Conclusion

In this paper, we have presented a new handwritten zip code recognition system using fractal feature. We used the inherent properties of fractal codes based on iterated function systems in fractal transformation classifier to perform digit recognition. Because quadtree portioning can be applied to an image of arbitrary size, this makes the proposed method robust to size and scale changes.

6. References

[1]- M.Dehghan, K.Faez, M.Ahmadi and M.Shridhar” Off-Line unconstrained Farsi Handwritten Recognition Using Fuzzy Vector Quantization and hidden Markov Word Models”. Proceeding of 15th International Conference on Pattern Recognition , Vol.2, 2000, pp.351-354.

[2]- I.S.I.Abuhiba, S.A.Mahmoud and R.G.Green” Recognition of handwritten cursive Arabic characters”. IEEE Trans on Pattern Analysis and Machine Intelligence.Vol.16,No.6, June 1994,PP.664-672.

[3]-M.Dehghan and K.Faez,”Farsi Handwritten character Recognition with Moment Invariants”, Proceeding of 13th International Conference on Digital Signal Processing, Vol.2, 1997,pp.507-510.

[4]- S.Mozaffari, K.Faez and H.Rashidy Kanan”: Recognition of Isolated Handwritten Farsi/Arabic Alphanumeric Using Fractal Codes” IEEE Proceeding of Southwest Symposium on Image Analysis and Interpretation. 2004,pp.104-108.

[5]- A. Mowlaei, K. Faez and A. T. Haghighat ,”Feature Extraction with Wavelet Transform for Recognition of Isolated Handwritten Farsi/Arabic Characters and Numerals”, Proceedings of 14th International Conference on Digital Signal Processing, Vol. 2, July 2002, pp. 923-926.

[6]- M.Dehghan, K.Faez, and M.Ahmadi ,”A hybrid handwritten word recognition using self-organizing feature map, discrete HMM and evolutionary programming”. Proceedings of the IEEE-INNS ENNS International joint Conference on Neural Networks. VOL.5 ,pp.515-520.,2000.

[7]- S.Mozaffari, K.Faez and H.Rashidy Kanan” Feature Comparison between Fractal codes and Wavelet Transform in Handwritten Alphanumeric Recognition Using SVM Classifier” International Conference on Patter Recognition, pp. 331-334,2004.

[8]- A. Mowlaei and K. faez,” Recognition of Isolated Handwritten Persian/Arabic Characters and Numerals Using Support Vector Machines”, IEEE Int. Workshop on Neural Networks for Signal Processing (NNSP), 2003, P.P. 547-554.

[9]- G.Neil and K.M.Curtis,”Scale and rotationally invariant object recognition using fractal transformations”, IEEE ICASP, 6,pp.3458-3461,1996.

[10]-A.Z.Kouzani,F.He and K.Smmut,” Fractal face representation and recognition,” IEEE International Confe. On Systems, Man and Cybernetics, Vol.2,pp.1609-1613,1997.

[11]-T.Tan and H.Yan, “ Face recognition by fractal transformations,” In Proceeding of ICASSP, Vol.3,pp.3537-3540,1999.

[12]- H.Ebrahimpour,V.Chandran and S.Sridharan,” Face Recognition Using Fractal Codes” ,Proceeding of WoSPA,Australia,PP.58-61,2000 .

[13]- H.Ebrahimpour,V.Chandran and S.Sridharan,” Robustness to expression variations in fractal-based face recognition”, In the Proceeding of ISSPA, Malasia,pp.359-362, 2001

[14]- A.Seropian, M.Grimaldi and N.Vincent,”Writer Identification based on fractal construction of a reference base”, Preceeding of ICDAR,2003.

[15]-Y.Fisher,” Fractal Image Compression: Theory and Application”, Springer-Verlag,New York,1995.