FACE RECOGNITION BY FRACTAL TRANSFORMATIONS

T. Tan and H. Yan

School of Electrical and Information Engineering
University of Sydney
N.S.W. 2006, Australia

ABSTRACT

In this paper, we propose a new method for computerized human face recognition using fractal transformations. The popular use of fractal image coding has been for image compression. It is only recently that their uses for object recognition are being explored. We will show that by utilizing the intrinsic properties of block-wise self-similar transformations in fractal image coding we can use it to perform face recognition. The contractivity factor and the encoding scheme of the fractal encoder are shown to affect recognition rates. Using this method, an average error rate of 1.75% was obtained on the ORL face database.

1. INTRODUCTION

Humans can perform face recognition with relative ease under many varying conditions. It is desirable to have an automatic system that can perform that same task with similar speed and accuracy. As a result, automatic face recognition algorithms have been an active area of research for more than 20 years, although the performance has not reached that of humans.

Many algorithms have been proposed for face recognition, and one of the most popular and successful is based on the eigenface method [8][18][23][24]. This method uses principal components analysis (PCA) to find a reduced basis for the sample space spanned by a set of training faces. A set of weights is obtained by projecting faces onto this basis of lower dimension. These weights can be grouped into different classes for faces from different people. Comparing the weights of an unknown face with that of known faces using a distance measure then performs classification. This method has the advantage of rapid classification, but it is sensitive to variable lighting conditions and spatial mismatches between the input face and the face in the database.

Other methods of face recognition have been proposed that include the use of convolutional networks (CN) [12], hidden markov models (HMM) [20][21], self-organizing maps (SOM) [9], probabilistic decision-based neural networks (PDBNN) [13], a combination of several methods using discrete cosine transforms (DCT) with image synthesis and neural networks [22], and a combination of analytic and holistic approaches [11].

A new approach to face recognition using fractal image coding is presented in this paper. Fractal image coding for object recognition has been explored by Neil et al. [16][17] for use on binary images. It has also been used in conjunction with neural networks by Kouzani et al. [10]. This paper extends those works to include gray scale images of human faces and the further investigation of the properties of fractal codes to increase the recognition rate. In particular, an expression relating the contractivity factor of a fractal code to the recognition rate is derived. Explanations of the mechanisms of the block-wise transformations in a fractal code that gives it the property of limited invariance to illumination effects, translations, scaling, and rotations, will be given. Experiments were performed on the Olivetti Research Labs (ORL) database of human faces. This database consisted of 400 images of 40 different persons in varying pose and facial expressions.

2. FRACTAL IMAGE CODING

Fractal objects, in the mathematical sense, exhibit self-similarity at all scales of magnification [15] that can be represented by compact mathematical equations. Not all two dimensional images are mathematical fractal objects, but they can be approximated by using a collage of self-similar sub regions of the image [1][2][3][5][6][7][19]. Encoding an image into its fractal approximation then involves the search for self-similar sub-regions of the image. In particular, a larger domain block maps to a smaller range block through a geometric and affine transformation. The transformation considered in this paper is described in the following equation,

$$\tau_n(\mathbf{p}_i) = \alpha_n(\mathbf{D}_{n,r}\mathbf{p}_i^{(n_D)}) + \gamma_n \tag{1}$$

where the n th transformation τ_n operates on the input image $\mathbf{p}_i$ on the n th domain block $\mathbf{p}_i^{(n_D)}$, and $\mathbf{D}_{n,r}$ is a decimation by r operator on the n th domain block. The contrast scaling factor is α_n, given by Equation (2), and γ_n is the illuminance shift factor given by Equation (3):

$$\alpha_n = \frac{\left|\max(\mathbf{p}^{(n_R)}) - \min(\mathbf{p}^{(n_R)})\right|}{\left|\max(\mathbf{D}_{n,r}\mathbf{p}^{(n_D)}) - \min(\mathbf{D}_{n,r}\mathbf{p}^{(n_D)})\right|} \tag{2}$$

where $\mathbf{p}^{(n_R)}$ denotes the n th range block, and $\mathbf{p}^{(n_D)}$ denotes the n th domain block of the image $\mathbf{p}$;

$$\gamma_n = \frac{1}{BiSize \times BjSize}\left(\sum_{i=1}^{BiSize}\sum_{j=1}^{BjSize}\mathbf{p}^{(n_R)}(i,j) - \alpha_n\sum_{i=1}^{BiSize}\sum_{j=1}^{BjSize}(\mathbf{D}_{n,r}\mathbf{p}^{(n_D)})(i,j)\right) \tag{3}$$

where Bi and Bj are the height and width of the range block respectively.

The following subclass of fractal encoders are considered in this paper:

- The input image is partitioned into non-overlapping square blocks, called range blocks.
- Each range block is the result of a mapping from a domain block within the image that is twice as large, which can overlap by half its size in the vertical and horizontal directions.
- Each transformation from the domain block to the range block involves a decimation by a factor of two using averaging followed by a contrast scale and illuminance shift, as shown in Equation (1).
- There are no domain block search restrictions.

The contractivity factor for this encoding scheme was derived. The method used is similar to the one used by Lundheim [14], but here it is extended to include overlapping domain blocks. Equation (4) is the result.

$$s = \sqrt{\max\left(\begin{array}{l} \frac{1}{r}\sum_{Q(n)=1}\left(\alpha^{(n)}\right)^2, \\ \frac{1}{r}\sum_{Q(n)=I_{dw}}\left(\alpha^{(n)}\right)^2, \\ \frac{1}{r}\sum_{Q(n)=I_{dh}I_{dw}-I_{dw}+1}\left(\alpha^{(n)}\right)^2, \frac{1}{r}\sum_{Q(n)=I_{dh}I_{dw}}\left(\alpha^{(n)}\right)^2, \\ \max_{2\le q\le I_{dw}-1}\left(\frac{1}{r}\sum_{Q(n)=q}\left(\alpha^{(n)}\right)^2+\frac{1}{r}\sum_{Q(n)=q-1}\left(\alpha^{(n)}\right)^2\right), \\ \max_{\substack{q=I_{dw}(1+n),\\ 1\le n\le I_{dh}-2}}\left(\frac{1}{r}\sum_{Q(n)=q-1}\left(\alpha^{(n)}\right)^2+\frac{1}{r}\sum_{Q(n)=q-I_{dw}-1}\left(\alpha^{(n)}\right)^2\right), \\ \max_{\substack{q=1+nI_{dw},\\ 1\le n\le I_{dh}-2}}\left(\frac{1}{r}\sum_{Q(n)=q}\left(\alpha^{(n)}\right)^2+\frac{1}{r}\sum_{Q(n)=q-I_{dw}}\left(\alpha^{(n)}\right)^2\right), \\ \max_{I_{dw}(I_{dh}-1)+2\le q\le I_{dh}I_{dw}-1}\left(\begin{array}{l}\frac{1}{r}\sum_{Q(n)=q-I_{dw}}\left(\alpha^{(n)}\right)^2+ \\ \frac{1}{r}\sum_{Q(n)=q-I_{dw}-1}\left(\alpha^{(n)}\right)^2\end{array}\right), \\ \max_{\substack{q=mI_{dw}+n,\\ 2\le n\le I_{dw}-1,\\ 1\le m\le I_{dh}-2}}\left(\begin{array}{l}\frac{1}{r}\sum_{Q(n)=q}\left(\alpha^{(n)}\right)^2+\frac{1}{r}\sum_{Q(n)=q-1}\left(\alpha^{(n)}\right)^2+ \\ \frac{1}{r}\sum_{Q(n)=q-I_{dw}}\left(\alpha^{(n)}\right)^2+\frac{1}{r}\sum_{Q(n)=q-I_{dw}-1}\left(\alpha^{(n)}\right)^2\end{array}\right)\end{array}\right)} \quad (4)$$

Where I_{dh} and I_{dw} are the number of range blocks in the vertical and horizontal directions respectively, $Q(n)$ is the domain block index of the n th range block. The range blocks are indexed sequentially row-wise from the top left to the bottom right corner. Equation (4) implies that the contractivity factor s is obtained by taking the square root of the largest sum of squares of $\alpha^{(n)}$ each time the domain block with index $Q(n)$ is used, divided by the decimation factor r . For simplicity we set

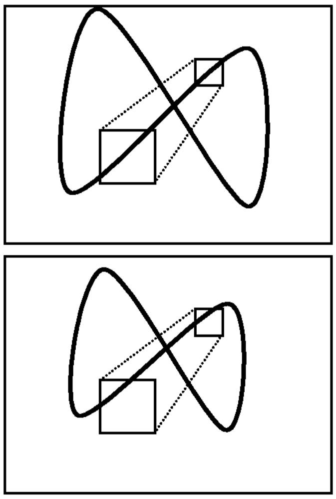

Figure 1. The top illustration shows an arbitrary image with a domain to range block transformation. The bottom illustration shows that same image after some rotation, scaling and shifting, with the same domain to range block transformation. In both cases, that same transformation captures the self-similarities in the image.

$\alpha^{(n)}=\alpha$, a constant value. Equation (4) then becomes Equation (5) .

$$s \le \sqrt{\frac{t\alpha^2}{r}} \quad (5)$$

where t is the maximum number of times a domain sub-block is used. In this paper t was set to 6.

3. RECOGNITION

In this paper, the following metric was used in the encoding schemes,

$$d(\mathbf{p},\mathbf{q}) = \|\mathbf{p}-\mathbf{q}\|_2 = \sqrt{\sum_{k=0}^{I_h}\sum_{l=0}^{I_w}\left(\mathbf{p}_{(k,l)}-\mathbf{q}_{(k,l)}\right)^2} \quad (6)$$

where I_h and I_w are the height and width of the images $\mathbf{p}$ and $\mathbf{q}$. This metric is also known as the Euclidean or RMS metric. For recognition the value of $d\left(f_j(\mathbf{p}_i),\mathbf{p}_i\right)$ is minimized, where f_j is the j th fractal code in the database, and $\mathbf{p}_i$ is the i th input image, then it is said that the j th face in the database is the best match of the i th input face. Using $d\left(f_j(\mathbf{p}_i),\mathbf{p}_i\right)$ for recognition results in limited invariance to translations, scaling, and rotations.

Figure 1 illustrates why a fractal block transformation can be invariant to rotation, scaling and shifting. The illustration shows that even after those distortions the self-similar sub-features of the image are still captured by the same domain to range transformation. Thus by measuring the distortion between an input image and the image after one decoding iteration, the

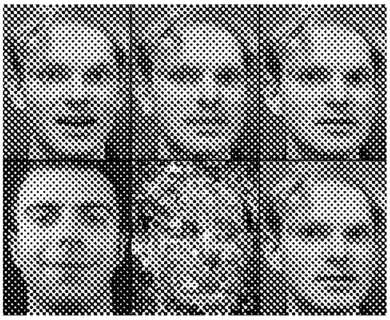

Figure 2. Left images are input faces. Right images are the attractor of the fractal code. The middle images are the result of applying that fractal code on the input face.

feature similarity between the input image and the attractor of the fractal code can be quantified. This value is given by $d(f_j(\mathbf{p}_i),\mathbf{p}_i)$. To further improve the recognition rate, the input image can be shifted in four extra directions by an amount equal to the size of the range block before decoding. This results in a four fold increase in computation, which can be counteracted by decreasing the size of the images used by a factor of 2. Thus the best matching face in the database to an input image is given by,

$$j_{best} = \arg\min_{1\le j\le W}\begin{pmatrix} d\left(f_j\left(\mathbf{p}_i^{left}\right),\mathbf{p}_i^{left}\right), d\left(f_j\left(\mathbf{p}_i^{up}\right),\mathbf{p}_i^{up}\right), \\ d\left(f_j\left(\mathbf{p}_i^{right}\right),\mathbf{p}_i^{right}\right), d\left(f_j\left(\mathbf{p}_i^{down}\right),\mathbf{p}_i^{down}\right), \\ d\left(f_j(\mathbf{p}_i),\mathbf{p}_i\right) \end{pmatrix} \quad (7)$$

where W is the number of images in the database, $\mathbf{p}_i^{<direction>}$ denotes an input face to be classified that has been shifted in the direction $<direction>$ by the size of a range block in that direction, and j_{best} is the index of the best matching face in the database.

The contractivity factor was found to affect the recognition rate in the manner described by Equation (8).

$$(1-s) \le \frac{d\left(f_j(\mathbf{p}_i),\mathbf{p}_i\right)}{d\left(\mathbf{p}_i,\tilde{\mathbf{p}}_j\right)} \le (1+s) \quad (8)$$

where $\tilde{\mathbf{p}}_j$ is the attractor of f_j. Increasing the contractivity factor results in slower convergence to the attractor of the fractal code, and thus decreases its sensitivity to distortions. This is desirable when the input faces vary significantly in pose and expression to the faces in the database. Decreasing the contractivity factor results in faster convergence, and thus increases its sensitivity to image differences and similarities. This is desirable when the input faces do not vary significantly in terms of pose and expression to the ones in the database.

Figure 2 shows the results of one decoding iteration using a fractal code from the database with different faces as input.

4. RESULTS

The ORL face database was used in our experiments. The experiments performed involve the use of four different selections of training and testing samples. One to five faces per person were used for training, and the ones that were not used for training were used for testing. The four different training/testing sets were selected at random except one, where the faces used for training for each person were taken in sequential order from the database. The range blocks were chosen to be 4×4 pixels in size, and the domain blocks had a size of 8×8 pixels.

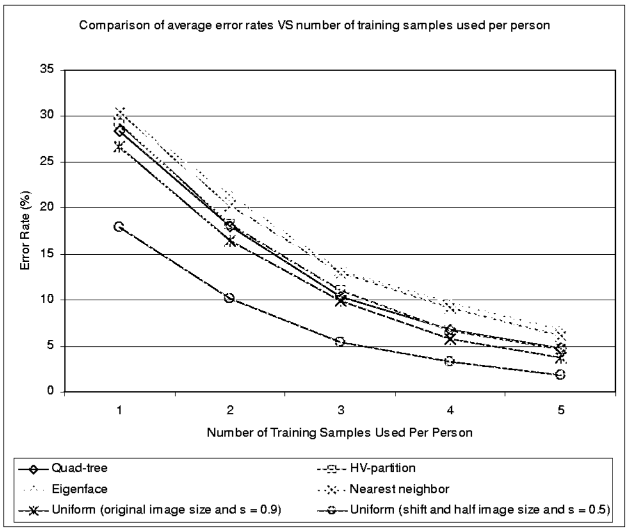

Figure 3. Comparison of the error rates obtained by averaging the results from four different training/testing sets.

In addition to the method described in this paper, experiments were also performed on a variant of this method that doesn't involve the extra shifting of the input image, using different fractal encoding schemes such as quad tree encoding [4] and HV partition encoding [4], eigenfaces, and the nearest neighbor classifier. The results of the experiments performed in this paper are summarized in Figure 3.

The best results were obtained with uniform block partitions with input image shifting, with an error rate of 1.75 %. When one sample per person was used for training, our system achieved an error rate of 17.8 % compared to the eigenface method which gave an error rate of 30.3 %.

5. DISCUSSION

Figure 3 shows that the method described in this paper outperforms all the other methods investigated. The best results were obtained when five samples per person were used for training, and when the contractivity factor was at 0.5. It is interesting to note that the optimum contractivity factors for uniform block partition encoding are different when input image shifting is used. This is explained by the fact that shifting the input image reduces the effects of distortions due to translations, scaling, and rotations. Thus using a smaller contractivity factor results in a better discrimination of faces, because the distortions are then mainly due to feature differences. Using quad tree encoding and HV partition encoding also gave inferior results. The main explanation for this is that those methods involve the use of non-uniform range blocks. This implies that a spatial structure is imposed on the input image to be decoded, further restricting the invariant effects of the block transformations.

Although those encoding schemes gave better approximations to the original image, they were not as suited for recognition tasks.

Table 1 shows a comparison of our results with others quoted in the literature. We can see that the method described in this paper outperformed the others in terms of error rate. One draw back of our method is that the complexity is linear to the size of the database, which is not as much the case for neural networks, but neural networks have the disadvantage of significant performance degradation when a large number of training samples are used. It is yet to be investigated whether our method has this similar drawback. Our method also has the smallest training time and the advantage of incremental training where the addition or removal of a sample from the database does not require the re-training of the whole database, as is required in methods using neural networks or eigenfaces.

System	Error Rate	Classification Time	Training Time
This paper	1.75%	3.23 seconds	61 seconds
2DCT+NN	2.4%	0.1 seconds	2.95 minutes
SOM+CN	3.8%	0.5 seconds	4 hours
PDBNN	4%	<0.1 seconds	20 minutes
2D-HMM	5%	240 seconds	n/a
HMM	13%	n/a	n/a

Table 1. Comparison of results with others quoted from the literature. Other results were obtained from Lawrence et al. [12] Lin et al. [13], and Satonaka et al. [22]. The experiments in this paper were performed on an Intel Pentium II 400 MHz system.

6. CONCLUSIONS

In this paper, we have presented a new face recognition algorithm using fractal image coding. The inherent properties of fractal codes based on the iterated function system can be exploited to perform face recognition. The effects of the contractivity factor and the fractal encoding scheme on the recognition rate were explained. It was shown that the method described in this paper achieved an error rate of 1.75 % when five samples per person were used for training, in an average recognition time of 3.23 seconds. For future work, we aim to explore the effects of performance degradation for extremely large databases.

7. REFERENCES

[1] Barnsley M. F. *Fractals Everywhere*. Academic Press, 1993.

[2] Barnsley M. F. and Hurd L. P. *Fractal Image Compressions*. AK Peters Ltd., 1993.

[3] Barnsley M. F. and Jacquin A. E. "Application of Recurrent Iterated Function Systems to Images". *Proceedings of SPIE*, 1001:122-131, 1988.

[4] Fisher Y., editor. *Fractal Image Compression: Theory and Application*. Springer-Verlag Inc., 1995.

[5] Jacquin A. E. "Fractal Image Coding: A Review". *Proc. of the IEEE*, 81(10):1451-1465, Oct 1993.

[6] Jacquin A. E. "A Novel Fractal Block Coding Technique for Digital Images". *IEEE ICASSP*, 4:2225-2228, 1990.

[7] Jacquin A. E. "Image Coding Based on a Fractal Theory of Iterated Contractive Image Transformations". *IEEE Trans. on Image Processing*, 1(1):18-30, Jan 1992.

[8] Kirby M. and Sirovich L. "Application of the Karhunen-Loeve procedure for the characterization of human faces". *IEEE Trans. PAMI*, 12(1):103-108, 1990.

[9] Kohonen T. "The Self-organizing Map". *Proceedings of the IEEE*, 78(9):1464-1480, 1990.

[10] Kouzani A. Z., He F., and Sammut K. "Fractal Face Representation and Recognition". *IEEE Int'l Conf. on Systems, Man and Cybernetics*, 2:1609-1613, 1997.

[11] Lam K. M. and Yan H. "An Analytic-to-Holistic Approach for Face Recognition Based on a Single Frontal View". *IEEE PAMI*, 20(7):673-686, July 1998.

[12] Lawrence S., Giles C. L., Tsoi A. C., and Back A. D. "Face Recognition: A Convolutional Neural Network Approach". *IEEE Trans. on Neural Networks, Special Issue on Neural Networks and Pattern Recognition*, 8(1):98-113, 1997.

[13] Lin S. H., Kung S. Y., and Lin L. J. "Face Recognition/Detection by Probabilistic Decision-Based Neural Network". *IEEE Trans. on Neural Networks*, 8(1):114-132, Jan 1997.

[14] Lundheim L. "A Discrete Framework for Fractal Signal Modeling". *Fractal Image Compression: Theory and Application*, Fisher Y., editor. Springer-Verlag Inc., 1995, chapter 7, pages 137-152.

[15] Mandelbrot B. *The Fractal Geometry of Nature*. Freeman, 1982.

[16] Neil G. and Curtis K. M. "Scale and Rotationally Invariant Object Recognition Using Fractal Transformations". *IEEE ICASSP*, 6:3458-3461, 1996.

[17] Neil G., Curtis K. M., and Craven M. P. "Shape Recognition Using a Novel Fractal Technique". *Proc. of the IEEE Int'l Conf. on Electronics, Circuits and Systems*, 2:724-727, 1996.

[18] Pentland A., Moghaddam B., Starner T., and Turk M. "View-based and Modular Eigenspaces for Face Recognition". *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1994.

[19] Popescu D. and Yan H. "A non-linear model for fractal image coding". *IEEE Trans. on Image Processing*, 6(3):373-382, 1997.

[20] Samaria F. S. "Face Recognition Using Hidden Markov Models". *Ph.D. thesis*, Trinity College, University of Cambridge, Cambridge, 1994.

[21] Samaria F. S. and Harter A. C. "Parametrisation of a Stochastic Model for Human Face Identification". *Proceedings of the second IEEE Workshop on Applications of Computer Vision*, Sarasota, Florida, 1994.

[22] Satonaka T., Baba T., Otsuki T., Chikamura T., and Meng T. H. "Object Recognition with Luminance, Rotation and Location Invariance". *IEEE Int'l Conf. on Image Processing*, 3:336-339, 1997.

[23] Turk M. and Pentland A. "Eigenfaces for Recognition". *Journal of Cognitive Neuroscience*, 3(1), 1991.

[24] Turk M. and Pentland A. "Face Recognition Using Eigenfaces". *Proc. IEEE Conf. On Computer Vision and Pattern Recognition*, 1991, pages 586-591.