Robust Face Recognition Method Based on Statistical Technique

Vinayadatt V. Kohir
P. D. A. College of Engineering
GULBARGA 585 101 INDIA
E-mail:vvkohir@ee.iitb.ernet.in

U. B. Desai
Electrical Engineering Department
IIT, MUMBAI 400 076 INDIA
E-mail: ubdesai@ee.iitb.ernet.in

Technical Areas: Vision, others

Introduction

In the last couple of decades many researchers [2, 3, 4, 6, 7, 8] have investigated the problem on face recognition. Though lot of effort has gone in to this filed, the problem has substantial scope to investigate into. A detailed survey can be found in [1].

The paper presents face recognition based on statistical methods. The method is based on the application of HMMs (Hidden Markov Models) in a DCT (Discrete Cosine Transform) domain. This method fuses the best features of DCT and HMM, makes it immune to noise. The proposed method is tested on the public domain ORL (Olivetti Research Laboratory) face database database. Recognition rate of 99.5%[1] is obtained with the original images. The recognition rate for noisy images ranges from 92% to 100% depending on the type of noise.

The space limitation prohibits inclusion of relevant mathematical formula.

DCT-HMM Based Face Recognition

HMMs are basically 1D data models. For face recognition application the face images must be represented in 1D format without loosing any vital information. So, the approach adopted has two parts - first one takes care of converting 2D face images to 1D data (data transformation), and the next part applies HMMs to transformed 1D sequence.

Data Transformation

The 2D face images are converted to 1D to enable use of HMMs. Input face image is sampled to obtain a sub-image sequence by sliding a square window (of size 16×16) with some over lap (75%) (see Fig.1). The overlap extracts the correlation between the adjacent image pixels.

Feature Extraction: Each of the sub-image obtained earlier is passed through the DCT transformation. The DCT coefficients are scanned in zig-zag fashion and are arranged as a column vector (as in Fig. 2. For every sub-image sequence a DCT transformed vector sequence is obtained. The DCT transformation decorrelates the sub-images and enables to retain few terms (energy compaction) reducing computation time and space.

HMM Training

Ergodic HMMs are trained using segmental k-means algorithm ([9, 10]). The probability distribution of DCT coefficients is assumed to be Gaussian. The the initial segmentation of segmental k-means algorithm does, to some extent determines the recognition rate.

For every subject to be recognized, a HMM is trained. HMM is trained using the 1D sequence obtained in the data transformation step. To incorporate variations in facial expressions (like smiling, frowning) and accessories (like stubble, spectacles), in stead on one typical face, a set of face images/subject are used for the training.

[1] 5 Training images and 5 Test images per subject

Recognition

The input face image is transformed to 1D sequence by generating a sub-image sequence followed by DCT transformation and zig-zag scan ordering of DCT coefficients. This sequence is applied to HMMs and viterbi score (state optimized probability function) is computed. The HMM with the highest score gives the recognition result.

Results

The ORL database is used to evaluate the recognition performance of the proposed method. ORL data base has 40 subjects, each having 10 different poses. All the images are in grey scale with 92×112 size.

Experiments are carried out with sub-image of size 16×16 and 75% sub-image overlap. Significant 10 DCT coefficients are chosen to form the observation vector. HMM is trained using 5 images per subject, and remaining 5 images of the subject are used for recognition, test set.

The recognition rate of original clean test set in 99.5%. Train set and test set of a typical subject are in Fig. 3 and Fig. 4 respectively.

To test the noise immunity of the approach. Gaussian noise is added to the test set and recognition is performed. Recognition rate varies from 96% to 100% depending on the SNR of test images. Further, binary (salt and pepper) noise is added at 400 random points of test images, and used to evaluate recognition performance. The recognition rate is 94%. The results are in Table 1. Noisy faces of the subject are in Fig. 5.

The test images are synthetically rotated, and recognition rate is: 92% for 5^o, 85.5% for 10^o and 70.5% for 15^o rotation. Last row of Fig. 6 represents the rotated images of the subject.

References

[1] R. Chellappa, C. L. Wilson and S. A. Sirohy, "Human and Machine Recognition of Faces: A Survey" ,*IEEE Proceedings*, Vol.83, No.3, pp.704-740, May, 1995.

[2] S. Lawarance, C. L. Giles, A. C. Tsoi and A. D. Back, "Face Recognition : A Convolutional Neural-Network Approach", *IEEE Trans. on Neural Networks*, Vol.8, No.1, pp.98-113, Jan.,1997.

[3] S. H. Lin, S. Y. Kung and L. J. Lin, "Face Recognition/Detection by Probabilistic Decision Based Neural Network", *IEEE Trans. on Neural Networks*, Vol.8, No.1, pp.114-132, Jan., 1997.

[4] S. M. Lucas, "Face Recognition with the Continuous n-tuple Classifier", *BMVC97*, 1997.

[5] L. R. Rabiner, "A tutorial on Hidden Markov Models and Selected Applications in Speech Recognition", *IEEE Proceedings*, Vol.77, No.2, pp.257-285, 1989. .F. S. Samaria and S. Young, "HMM-based Architecture for Face Identification", .*Image and Vision Computing*, Vol.12, No.8, pp.537-543, Oct., 1994.

[6] F. S. Samaria, "Face Recognition Using Hidden Markov Models", *Ph.D. Thesis*, University of Cambridge, UK, 1994.

[7] Theme Section - Face and Gesture Recognition, *IEEE Pattern Recognition and Machine Intelligence*, Vol.19, No.5, July, 1997.

[8] M. Turk and A. Pentland, "Eigenfaces for Recognition", *Journal of Neuroscience*, Vol.3, No.1, MIT, 1991.

[9] Vinayadatt V. Kohir and U. B. Desai, "Face Recognition using a DCT-HMM Approach", *WACV98*, Princeton, Oct., 1998.

[10] Yang He and Amalan Kundu, "Unsupervised Texture Segmentation Using Multichannel Decomposition and Hidden Markov Models", *IEEE Image Processing*, Vol. 4, No. 5, pp.603-619, May, 1995.

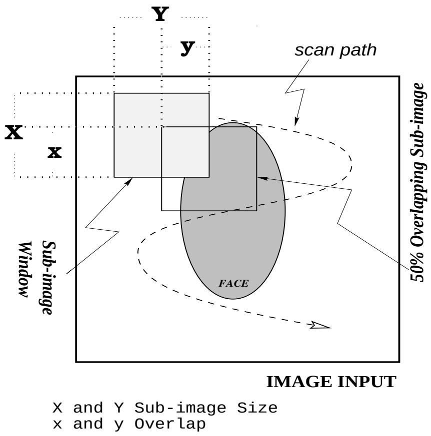

Figure 1: Sampling scheme to generate sub-image sequence

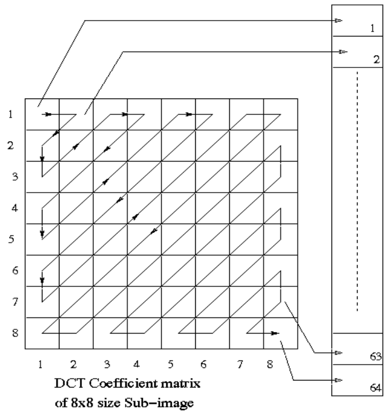

Figure 2: Zig-zag scanning of DCT coefficients to construct vector from a sub-image

Figure 3: Train set of a typical ORL subject used to build the HMM representing subject's face model.

Figure 4: Test set of a typical ORL subject used to evaluate the recognition performance of the proposed method.

Figure 5: Noisy images of the subject. First three images have Gaussian noise with SNR of 20dB, 10dB and 5dB respectively. The last image is corrupted by binary noise (SNR approx. 10dB).

Figure 6: Synthetic rotation of image. The degree of rotation, left to right, is $5^0, 10^0, 20^0, -5^0, -10^0, -15^0$.

Recognition rate in percentage								
Clean	Gaussian Noise			Binary Noise		Rotation in Deg.		
images	20dB SNR	10dB SNR	5dB SNR	Case 1	Case 2	5	10	15
99.5	99.5	98.5	94.5	100	92	92	85.5	70.5

Table 1: Recognition results.

AUTHOR BIOGRAPHY

Vinayadatt V. Kohir: obtained B.E. from P.D.A. College of Engineering, Gulbarga, India, in 1984 and M.Tech. from IIT-Kharagpur, India, in 1989, all in Electronics and Communication Engineering. Currently he is a Doctoral candidate at IIT-Bombay.

He is Senior Lecturer in Electronics and Communication Engineering Department of P. D. A. college of engineering, Gulbarga, India, since 1985. The areas of interest include ANN, Pattern recognition, Wavelets, Computer vision and Image processing. He is Life member of ISTE.

U. B. Desai: received B. Tech degree from IIT-Kanpur, India in 1974, the M.S. degree from the State University of New York, Buffalo,in 1976, and the Ph.D. degree from Johns Hopkins University, Baltimore, 1979, all in Electrical Engineering.

From 1979 t0 1984 he was Assistant Professor in Electrical Engineering Department at Washington State University, Pullman, WA, USA, and as Associate Professor at the same place from 1984 to 1987. Since 1987 he has been a Professor in the Electrical Engineering Department at IIT-Bombay. He has held Visiting Associate Professor's position at Arizona State University, Purdue University and Stanford University.

His research interests are in the the areas of image processing, computer vision, wavelets, adaptive signal processing and artificial neural networks. He is the Editor of the book Modeling and Applications of Stochastic Processes (Kluwer Academic Press, Boston, USA 1986). He is senior member of IEEE.