ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 1, *January 2015*

Face Recognition System Using PCA and DCT in HMM

SamerKais Jameel

Lecturer, Computer Science, University of Raparin, Sulaimaniya, Iraq

Abstract: The speed of procedures became necessities during recent years; therefore using computers turn out to be the most important factors to increase the speed of implementation especially in security aspect such as recognition of people.There are a lot of waysto recognize the people face recognition is one of them. In this work the details of the face have been taken as blocks and Discrete Cosine Transform (DCT) is used, applied on face image's blocks. Then without doing inverse DCT Principal Component Analysis (PCA) is applied directly for dimensionality reduction this makes the system very fast. Olivetti Research Laboratory (ORL) database of faces had been used to obtain the results.Each face is considered as a numerical sequence (blocks) that can be easily modelled by HMM. On 400 face images of the (ORL) face database the system has been examined. The experiments showed a recognition rate of 95.211%, using half of the images for training.

Keywords: Face Recognition, Hidden Markov Model, Discrete Cosine Transform (DCT), Principal Component Analysis (PCA).

I. INTRODUCTION

Face recognition Plays an important role for research such as commercial and law enforcement applications [1-4], where it may be for identification of verification purposes.The algorithms developed for face recognition problems are generally grouped into two categories [5, 6] namely feature based and holistic based.The geometrical analysis of the facial features like eyes, nose and mouth are analyzed in feature based after facial feature detection, whereas faces are analyzed as two dimensional patterns in holistic approaches. One of important things for extract the effective features and also for reducing computational complexity in classification stage is Dimensionality reduction.Principal component analysis (PCA) [7], [8], Discrete cosine transform (DCT) [9], and Linear discriminate analysis (LDA) [10] are the main techniques used for data reduction and feature extraction in the appearance based approaches.The most efforts are given mainly on developing feature extraction methods and employing powerful classifiers such as Euclidean distance Classifier, Hidden Markov Models (HMMs) [11], and neural networks [12], [13].

This paper presents a new approach using one dimensional Discrete HMM as classifier and Principal component analysis (PCA) coefficients as features for face recognition afterapplying Discrete cosine transform (DCT).Using seven-states HMM to model face configuration. The approach has been examined on ORL database,gain more speed in training and testing by resized the 112×92 pgm formattedimages of database to 50% of its size to be 56 ×46 pgm formatted images.

II. HIDDEN MARKOV MODELS (HMMS)

Mathematical theory of Hidden Markov Models (HMMs) was originally describedduring the 1960's and early 1970's [14].(HMMs) are a technique applied in practical pattern recognition applications, more specifically in speech recognition problems [15]. Recently it has been used in vision: texture segmentation [16], face finding [17], object recognition [18] and face recognition [19]. face image represented as a sequence of states produced when the face is scanned from top to bottom, and HMM is made of states, where the probability to move from one state to another depends only on those two states and not any further history [20,21]. HMM can be represented as a triplet

$l=\{A, B, \pi\}$ (1)

- The number of states N, and the state at time t is given by

$qt, 1 \le t \le T.$ (2)

Where T is the length of the observation sequence.

- The initial state distribution:

$\pi = \{pi\}$, where $pi = p\{q1 = i\}, 1 \le i \le N$ *(3)*

- The state transition probability matrix

$A = \{aij\}$, where

$aij = p\{qt+1 = j \mid qt = i\}, 1 \le i, j \le N$, $0 \le aij \le 1$,and $\sum_{j=1}^{N} a_{ij} = 1$ (4)

- A probability distribution for each of the states,

$B = \{b_j(o_t)\}$ (5)

Usually probability density function is approximated by the weighted sum of M.Where

$b_j(o_t) = p[o_t = V_K | q_t = s_i], 1 < j < N, 1 < k < M$ (6)

M= |V| is the number of the different observationsymbols, where V={v1,v2,…,vM} is the setofall possible observation symbols.The observation symbol at time t is given by $o_i \in V$ S={s1,s2,…,sN} is the set of all possible states. The state of the model at time t is given by $q_t \in s$.HMMs generally work on sequences of symbols called observation vectors, that's why in this paper divided the face image into seven regions which each is assigned to a state in a left to right one dimensional HMM. Figure 1 shows the mentioned seven face regions.

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 1, *January 2015*

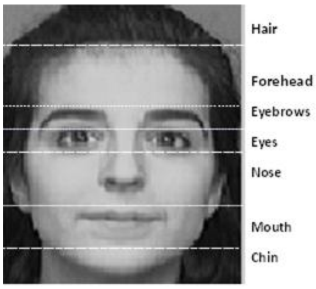

Fig. 1: Seven regions of face coming from top to down in natural order.

A simple structure and small number of parameters is used to build the model as shown in figure 2.

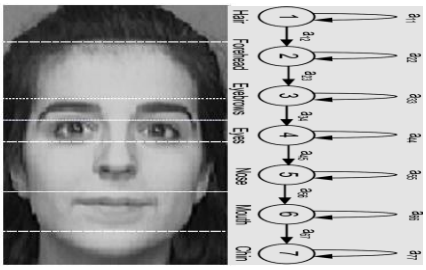

Fig. 2: A one dimensional HMM model with sevenstates for face image with seven regions.

III.PRINCIPAL COMPONENT ANALYSIS (PCA)

PCA aims to maximize between-class data separation [22]. It reduces the dimensionality of the description by projecting the points onto the principal axes, where orthonormal set of points are in the direction of maximum covariance of the data. PCA is an optimal compression scheme that minimizes the mean squared error between the original images and their reconstructions for any given level of compression [23,24]. works by finding a new coordinate system for a set of data, where the axes (or principal components) are ordered by the variance contained within the training data [25].The approach for face recognition aims is decompose face images into small set of characteristic feature images called eigenfaces which used to represent both existing and new faces. The training database consists of M images which is same size. The images are normalized by converting each image matrix to equivalent image vector I.
The training set matrix is the set of image vectors with

Training set$I = [I_1, I_2, I_3 \dots \dots I_M]$ (7)

The mean face (ψ) is the arithmetic average vector as given by:

$$\psi = \frac{1}{M}\sum_{I=1}^{M} I_i \qquad (8)$$

The deviation vector for each image Φi is given by:

$\Phi = I_i - \psi \qquad i = 1,2, \dots \dots, m$ (9)

Consider a difference matrix A= [Φ1, Φ2, …….ΦM] which keeps only the distinguishing features for face images and removes the common features. Then eigenfaces are calculated by find the Covariance matrix C of the training image vectors by:

$$C = A.A^T \qquad (10)$$

Due to large dimension of matrix C, consider matrix L of size (Mt X Mt) which gives the same effect with reduces dimension. The eigenvectors of C (Matrix U) can be obtained by using the eigenvectors of L (Matrix V) as given by:

$Ui = AV_i$ (11)

The eifenfasecs are:

eigenface= $[U_1, U_2, U_3, \dots \dots U_M]$ (12)

Instead of using M eigenfaces, the highest m' <= M is chosen as the eigenspace. Then the weight of each eigenvector ωi to represent the image in the eigenface space, as given by:

$\omega i = = U_i^T(I - \psi),\ i = 1, 2, \dots \dots, m'$ (13)

Weight matrix $\Omega = [\omega_1, \omega_2 \ \dots \omega_m]\ T$ (14)

Average class projection $\Omega_\psi = \frac{1}{x_i}\sum_{i=1}^{x_i} \Omega_i$ (15)

IV.DIMENSIONAL DISCRETE COSINE TRANSFORM (DCT)

In Discrete Cosine Transform (DCT), a series of finitely several data points are expressed in terms of a sum of cosine functions oscillating at diverse frequencies [26, 27]. It can help to extract the feature of face image by apply two dimensional Discrete Cosine Transform(DCT) because the coefficients of most upper region and most left region in DCT transform represent edge information [28, 29]. So from the upper most left coefficients, average is extracted.DCT II is mostly used in signal processing and is often named as "the DCT". A 2D M X N DCT is defined as follows:

$$C(u,V) = a(u)a(v)\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)$$

$$X \cos\left[\frac{\pi(2x+1)u}{2M}\right]\cos\left[\frac{\pi(2y+1)v}{2N}\right] \qquad (16)$$

Where,

$$a(u) = \frac{1}{\sqrt{N}}, if\ u = 0$$

$$a(u) = \sqrt{\frac{2}{M}}, if\ u = 1, 2, \quad , M-1$$

$$a(v) = \frac{1}{\sqrt{N}} \quad if\ v = 0$$

$$a(v) = \sqrt{\frac{2}{N}}\ if\ v = 1, 2, \quad N-1 \qquad (17)$$

IV .PROPOSED SYSTEM

A. *Order-statistic filter*

The highlights in the subject's eyes affected the classification accuracy [30].This highlights made of flash effecting, so the first step in this system is use a specific filter which compensates the flash effect. Filter also reduces salt noise as a result of the min operation, and directly affects the speed and recognition rate of the algorithm; this filter is "Order-statistic", that used a two dimensional order statistic filter, which replaces the centered element of a 3×3 window with the minimum element in the window,Figure 3 show An example of

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 1, *January 2015*

operation the order static filter. Order static filter can simply be represented by the following equation.

$$\hat{f}(x,y) = \min_{(S,t)\in S_{xy}}\{g(s,t)\} \qquad (18)$$

$g(s,t)$ is the grey level of pixel (s, t) and S_{xy} is the mentioned window.

a)

b)

Fig.3: An example of operation of the order static filter1
a) Before filteringb) After filtering.

Figure 4 shows a simple example demonstrating how minimum order-static filter works using a 3×3 window operates on a 6×6 region of first image.

190	191	188	187	184	180
193	194	189	188	185	179
194	194	189	188	185	179
88	188	187	184	183	178
179	180	181	178	178	176
182	175	176	181	178	179

0	0	0	0	0	0
0	188	187	184	179	0
0	187	184	183	178	0
0	179	178	178	176	0
0	175	175	176	176	0
0	0	0	0	0	0

Fig. 4: An example of the operation of minimum order static filter.
a) Before filtering b) After filtering

B. *Feature Extraction*

EveryHMM is associated with non-observable (hidden) statesand an observable sequence generated by the hidden stats individually.Extract the feature must start after construct the Observation Vectors to be use it in HMM that require a one-dimensional observation sequence, so the images should be interpreted as a one dimensional sequence.The Observation Vectors can obtain by divided face image into overlapping blocks that is mean each face image with width W and height H must converted into number of blocks of height L and width Ware given by:

$$T = \frac{H-L}{L-P} + 1 \qquad (19)$$

$P$is overlap size of two consecutive blocks.

Reduce the computational complexity and memory consumption by resize ORL face database from 112×92 into 96×46, and use a sampling window of six pixels height and 46 pixels width, (L=6, P=5), so the observation vector is large number of L×W (L=6 and w=46) blocks each containing 276 pixels.So there are now 51 observations that correspond to the numberof blocks for each face image.One major improvement is use PCA coefficients as features instead of gray values of the pixels in the sampling windows (blocks). The process of select the feature is start with applies DCT on the blocks of the image and processes each block individually to eliminatethe redundancies in an image and extract from them the most significant elements (i.e. coefficients), The best performance is found for most effective for recognition is to divided the block into subblocks of size (23x23). The next step is apply PCA without doing the inverse DCT to reduce the computational complexity and to select a subset of size m that leads to the smallest classification error and smallest computational cost. The PCA is contains three matrixes (COEFF, SCORE, and latent).Obviously the first two values of the "Latent" vector are very big and from the "SCORE" the second one of the first column is bigger than the others, so as expected, the two biggest values along with "Latent" vector have the best classification rate. Which is why in this system use two first coefficients of matrix latent and first coefficient of matrix SCORE as three features (latent11, latent21, and SCORE11) associating each block. Thus each block of size 276(=6×46) pixels, is represented by 3 values.

Now to make all blocks of an image mapped to a sequence of integer numbers that is considered as an observation vector will go to "quantization" the images will be so interpreted can simply use HMM for classification. Consider a vector $X = (x1, x2, \dots, xn)$ with continuous components. Suppose xi is to be quantized into Di distinct levels. So the difference between two successive quantized values will be as equation.

$$\Delta_i = (x_{imax} - x_{imin})/D_i (20)$$

x_{imax} and x_{imin} are the maximum and minimum values that xi gets in all possible observation vectors respectively. Now xi replaced with its quantized value computed as below:

$$x_{iq} = \left[\frac{x_i - x_{imin}}{\Delta_i}\right] \qquad (21)$$

Where x_{iq} is quantized value, At last each quantized vector is associated with a label that here is an integer number. So each block of image will mapped to an integer. Consequently the image is mapped to a sequence of integer numbers that is used as an observation vectorwith size 1x51, and finally uses it in HMM for classification.In this paper first feature is quantized the (latent11,) into 10, the second feature (latent21) 12 and the third one (SCORE11) into 10 levels, leaving 1200 possible distinct vectors for each block. These quantization levels have been found based on experimental results.

C. Training Process

Five images of the same face (one person) are used to train the related HMM and the remaining five are used for testing. The Baum-Welch algorithm [31] used to train a HMM for each person in the database, at the first step $\lambda = (A, B, \pi)$is initialized. The initial values for A and π are as follows:

$$a_{i,i} = a_{i,i+1} = 0.5 \qquad 1 < i < 6$$
$$a_{7,7} = 1$$

DOI 10.17148/IJARCCE.2015.4103

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 1, *January 2015*

$\pi_0 = 1$

Initial estimates of the observation probability matrix B are obtained as following:

$$B = \frac{1}{M}\, Ones(N, M) \qquad (22)$$

Where M is the number of all possible observation symbols obtained from quantization procedure and N is the number ofstates(in the proposed model N equals to 7). Figure 6 shows the estimation process related to one learning image. This process is iterated for all training images of a person. The iterations stop, when variation of the probability of the observation vector in two consecutive iterations is smaller than a specified threshold or the number of iterations reaches to an upper bound.The estimated parameters of each training image are used as initial parameters of next training image as show in figure 6. The estimated HMM of the last training image of a class is considered as its final HMM.

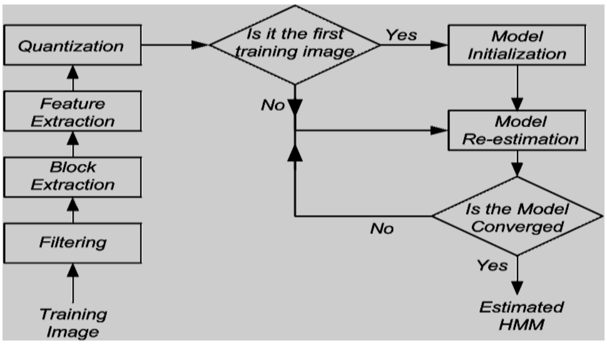

Fig. 6: The training process of a training image.

V. Result and Discussion

For an incoming face image, it is simply calculate the probability of the observation vector (current test image) given each HMM face model. A face image z is recognized as face y if:

$$p(o^{(z)} / \lambda_y) = max\_n\, p(o^z | \lambda_n) \qquad (23)$$

The database(ORL) contains ten different face images per person of 40 people with the resolution of 112×92 pixels. by applying the proposed system the recognition rate is 95.122%, Table 1 represents a comparison among different face recognition techniques and the proposed system on the ORL face database, notice that all different face recognition techniques in Table 1 use 112×92 resolution of ORL face database, wherethe proposed system use 56×46 image size which refers to reducing the processing time.

The classification process was repeated using each of these features individually. Figure 7 shows the obtained results for some important features, for different values of the number of symbols the recognition rate of the system has been calculated.By changing the number of quantization levels the number of symbol will simply change.the propose system quantized the first feature ($SCORE_{21}$) into ten, the second feature ($latent_{11}$) into twelve and the third feature ($latent_{21}$) into ten levels and obtained 95.122% recognition rate with1200 symbols.To illustrate the relation between number of symbols and recognition rate varied the number of symbols from 56 to 4096. The recognition rate is illustrated in figure 8. Increasing the number of symbols to achieve greater recognition rate leads to more time consumption for training and testing procedure. To prevent this event can use low number of symbols.

TABLE I

COMPARATIVE RESULT ON ORL DATABASE. "PCA +DCT " REPRESENTS THE PROPOSED METHOD WHICH TESTED ON IMAGES SIZE 56X46 OF ORL DATABASE WHERE THE OTHER METHODS USE IMAGES SIZE 112X96

Method	Error	Ref.
Elastic matching	20.0%	[32]
Point-matching and Correlation	16%	[33]
Top-down HMM + DCT coef.	16%	[4]
Independent Component Analysis	15%	[34]
Top-down HMM + gray tone features	13%	[35]
Markov Random Fields	13%	[36]
Eigenface	9.5%	[7]
Gabor filters + rank correlation	8.5%	[37]
Pseudo 2D HMM + gray tone features	5%	[2,19]
PDNN	4%	[38]
SVM + PCA coef.	3%	[39]
Pseudo 2D HMM + Wavelet	0%	[40]
1DHMM + Wavelet	0%	[41]
PCA+DCT (The proposed system)	**4.8%**	**This paper**

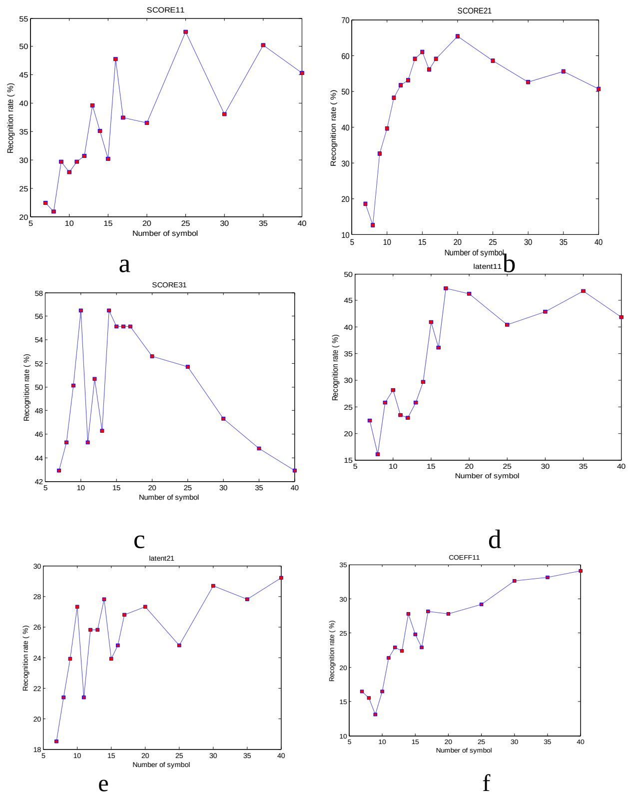

Fig.7: show the effect of each feature to recognition rate. A) SCORE11 classification ability. Maximum value is 52.6% for 25 symbols. B) SCORE21 classification ability. Maximum value is 65.3% for 20 symbols. C) SCORE31 classification ability. Maximum value is 56.5% for10 symbols. D) latent11 classification ability. Maximum value is 47.3% for 17 symbols E) latent21 classification ability. Maximum value is 29.2% for 40 symbols F) coeff11 classification ability. Maximum value is 34% for 40 symbols.

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 1, *January 2015*

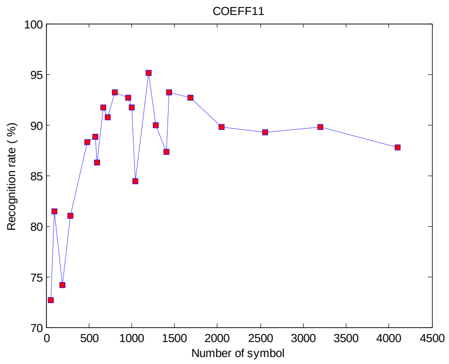

Fig. 8: showing the relation between number of symbols and recognition rate. Maximum value is 95.122% and encounter on 1200 symbols.

Because resized images database into low dimension (56x46), the system operates very fast with 1200 symbols and 95.122% recognition rate. Also resized face images to 28×23 resolution and obtain 90.7% recognition rate. Table II shows the comparison results on 56×46 and 28×23 face images.

TABLE II
COMPARING RESULT OF DIFFERENT IMAGE SIZE OF ORL DATABASE

Image size	56x46	28x23
# of training images	5	5
Training time per image (seconds)	0.69	0.34
Recognition time per image (seconds)	0.29	0.16
# of symbols	1200	640
Recognition rate(%)	95.122	90.7

And table 3 shows a comparison of the different face recognition techniques on the ORL face database which reported their computational cost.

TABLE III
COMPARATIVE COMPUTATIONAL COSTS AND RECOGNITION RESULTS OF SOME OF THE OTHER METHODS ASREPORTED BY THE RESPECTIVE AUTHORS ON ORL FACE DATABASE.

Method	Recognition %	Training time per image	Recognition time per image
PDBNN [42]	96	20min	$\leq$ 0.1 sec
n-tuple [43]	86	0.9 sec.	0.025 sec.
1DHMM + Wavelet [41]	100	1.13 sec.	0.3 sec.
Pseudo-2D HMM [19]	95	n/a	240 sec.
DCT-HMM[44]	99.5	23.5 sec.	3.5 sec.
(The proposed system)	95.122	0.6	0.2 sec.

From Table 3 the proposed system has a recognition rate of 95.122% and a low computational cost. Besides these advantages, the system has low memory space consumption because of resizing the face images.

VI. CONCLUSION

A fast and efficient system was presented. Images of each face were converted to a sequence of blocks. Each block was processed by DCT and featured by a few number of its PCA parameters. Each class has been associated to hidden Markov model as its classifier. The evaluations and comparisons were performed on the well-known face image data base; ORL. The system was very fast. This was achieved by resizing the images to smaller size and using a small number of features describing blocks.

ACKNOWLEDGMENT

Thanks to God who helped me to complete this research, and special thanks to my parents and to my family.

REFERENCES

[1] FERDINANDO SAMARIA & FRANK FALLSIDE (1993) FACE IDENTIFICATION AND FEATURE EXTRACTION USING HIDDEN MARKOV MODELS. IMAGE PROCESSING: THEORY AND APPLICATIONS, 1, EDITED BY G. VERNAZZA, ELSEVIER.
[2] F. SAMARIA, FACE RECOGNITION USING HIDDEN MARKOV MODELS, PHD THESIS, UNIVERSITY OF CAMBRIDGE, 1994.
[3] A. V. NEFIAN AND M. H. HAYES, "HIDDEN MARKOV MODELS FOR FACE RECOGNITION", PROC. INTERNATIONAL CONF. ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, SEATTLE, 1998, VOL. 5, PP. 2721-2724.
[4] A. V. NEFIAN AND M. H. HAYES, "AN EMBEDDED HMM-BASED APPROACH FOR FACE DETECTION AND RECOGNITION", ICASSP'99.
[5] W. ZHAO, R. CHELLAPPA, P.J. PHILLIPS, A. ROSENFELD, FACE RECOGNITION: A LITERATURE SURVEY, ACM COMPUTING SURVEYS 35 (2003) 399–458 C, 1963, PP. 271–350.
[6] R.CHAPPELA, C. L. WILSON, AND S. SIROHEY, " HUMAN AND MACHINE RECOG -NITION OF FACES: A SURVEY," PROC. IEEE, VOL. 83,NO. 5, PP. 705-740, MAY 1995.
[7] M. A. TURK AND A. P. PENTLAND, "EIGENFACES FOR RECOGNITION," J. COGN. NEUROSCI., VOL. 3, PP. 71–86, 1991.
[8] P. N. BELHUMEUR, J. P. HESPANHA, AND D. J. KRIEGMAN, "EIGENFACES VERSUS FISHERFACES: RECOGNITION USING CLASS SPECIFIC LINEAR PROJECTION," IEEE TRANS. PATTERN ANAL. MACH. INTELL., VOL. 19, NO. 7, PP. 711–720, JUL. 1997.
[9] Z. M. HAFED AND M. D. LEVINE, "FACE RECOGNITION USING THE DISCRETE COSINE TRANSFORM," INT. J. COMPUT. VIS., VOL. 43, NO. 3, PP. 167–188, 2001.
[10] M. KIRBY AND L. SIRVOICH, "APPLICATION OF THE KARHUNEN-LOEVEPROCE-DURE FOR THE CHARACTERIZATION OF HUMAN FACES," IEEE TRANS. PATTERN ANAL. MACH. INTELL., VOL. 12, NO. 1, PP. 103–108, JAN. 1990.
[11] H. OTHMAN, AND T. ABOULNASR, "A SEPARABLE LOW COMPLEXITY 2D HMM WITH APPLICATION TO FACE RECOGNITION", IEEE TRANS. PATTERN. ANAL.MACHIEINELL., VOL. 25, NO.10, PP. 1229-1238, 2003..
[12] M. ER, S. WU, J. LU, AND L.H.TOH, "FACE RECOGNITION WITH RADIAL BASIS FUNCTION (RBF) NEURAL NETWORKS", IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 13, NO. 3, PP. 697-710, 2003.
[13] M.J.ER, W.CHEN, AND S.WU, "HIGH SPEED FACE RECOGNITION BASED ON DISCRETE COSINE TRANSFORM AND RBF NEURAL NETWORK", IEEE TRANS. ON NEURAL NETWORK, VOL. 16, NO.3, PP. 679-691, 2005.
[14] BAUM, L. E. & PETRIE, T. (1966). STATISTICAL INFERENCE FOR PROBABILISTIC FUNCTIONS OF FINITE STATEMARKOV CHAINS.ANNALS OF MATHEMATICAL STATISTICS, VOL. 37, 1966.
[15] JELINEK, F.; BAHL, L.R. & MERCER, R.L. (1975). DESIGN OF A LINGUISTIC STATISTICAL DECODER FOR THERECOGNITION OF CONTINUOUS SPEECH.IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 21, NO 3, PP. 250 – 256, 1975.
[16] YANG HE AND AMALANKUNDU, "UNSUPERVISED TEXTURE SEGMENTATION USING MULTI-CHANNEL DECOMPOSITION AND HIDDEN MARKOV MODELS," IEEE IMAGE PROCESSING, VOL. 4, NO. 5, PP. 603-619, MAY 1995.

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 1, *January 2015*

[17] Rajagopalan A.N., Sunil Kumar K., JayashreeKar-lekar, Manivasakan, MilindPatil, U.B.Desai, P.G. Poonacha, S. Choudhuri, "FindingFaces in Photographs," ICCV98, pp. 640-645, Jan. 1998.
[18] Joachim Hornegger, Heinrich Niemann, Pulus D. and Schlottke G., "Object Recognition using Schlottke G., "Object Recognition using hidden Markov Models," Proc. of Pattern Recognition in Practice, pp. 37-44, 1994.
[19] Samaria F.S. and Young S., "HMM-based Architecture for Face Identification," Image and Vision Computing, Vol. 12, No. 8, pp. 537-543, Oct. 1994.
[20] L. Rabiner, "A tutorial on Hidden Markov Models and selected applications in speech recognition", Proceedings of the IEEE, Vol.77, No.2, 1989, pp.257-286.
[21] Hidden Markov Model Toolkit (HTK), S. Young et al.: http://htk.eng.cam.ac.uk
[22] Jan Mazanec, Martin Melisek, Milos Oravec, JarmilaPavlovicova "Support Vector Machines, PCA and LDA in face recognition" Journal of Electrical Engineering, VIX(4), pp.203-209, 2008.
[23] M. Kirby, L. Sirovich, Application of the Karhunen-Loeve procedure for the characterization of human faces, IEEE Transactions on Pattern Analysis and Machine Intelligence 12 (1990) 103–107.
[24] L. Sirovich, M. Kirby, A low-dimensional procedure for the characterization of human faces, Journal of the Optical Society of America 4 (1987) 519–524.
[25] Paul Nicholl, Abbes Amira, DjamelBouchaffra "Multiresolution Hybrid Approaches for Automated Face Recognition", In the Proceedings of the IEEE 2nd NASA/ESA Conference, pp.89-96, 2007.
[26] W. Chen, M. Joo and S. Wu, "Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain", IEEE Trans. on Systems, Man, And Cybernetics vol. 36, no. 2, pp.458-466, April 2006.
[27] V. P. Vishwakarma, S. Pandey and M. N. Gupta," An illumination invariant accurate face recognition with down scaling of DCT coefficients", Journal of Computing and Information Technology, vol.18, no. 1,pp. 53–67, 2010.
[28] Prof. G. U. Kharat, Dr.S.V.Dudul Neural Network Classifier for Human Emotion Recognition from Facial Expressions Using Discrete Cosine Transform", IEEE, pp 653-658, 2008
[29] Prof. K V Krishna Kishore, GP.S.Varma "Efficient Facial Emotion Classification with Wavelet fusion of Multi Features", International Journal of Computer Science & Network Security, vol. 11 no, 8 pp. 160-165, Aug.2011ISSN : 1738-7906.
[30] Haralick, R. M. and Shapiro, L. G., Computer and Robot Vision, Volume I. Addison-Wesley, 1992.
[31] Rabiner L. R., "A Tutorial on Hidden Markov Modeles and Selected Application in Speech Recognition," Proceeding Of the IEEE, Vol. 77, No. 2, pp. 257-258, 1989.
[32] Zhang J., Yan Y., and Lades M. Face recognition:eigenface, Proceedings of the IEEE, Vol. 85, No. 9, elastic matching and neural nets. September 1997.
[33] Lam K. M. and Yan H., "An analytic-to-holistic approach for face recognition on a single frontal view," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 7, pp.673–686, 1998.
[34] Yuen P. C. and Lai J. H., "Face representation using independent component analysis," Pattern Recognition, Vol. 35, No. 6, pp. 1247–1257, 2002.
[35] Samaria F. and Harter A.. "Parameterization of a Stochastic Model For Human Face Identification," In Proceedings of IEEE Workshop on Applications of Computer Vision, Sarasota, Florida, December 1994.
[36] Huang R., Pavlovic V. and Metaxas D. N., "A Hybrid Face Recognition Method using Markov Random Fields," IEEE, 0-7695-2128-2, 2004.
[37] Ayinde O. and Yang Y., "Face recognition approach based on rank correlation of gaborfiltered images," Pattern Recognition, Vol. 35, No. 6, pp. 1275–1289, 2002.
[38] Lin S., Kung S., and Lin L.. "Face Recognition/Detection by Probabilistic Decision- Based Neural Network" IEEE Trans. Neural Networks, Vol 8, No 1 pp. 114–131, January 1997.
[39] Guo G., Li S. Z., , and Kapluk C.. "Face recognition by support vector machines," Image and Vision Computing, Vol. 19, No. 9-10, pp. 631–638, 2001.
[40] Bicego M., Castellani U., and Murino V., "Using Hidden Markov Models and Wavelets for face recognition," In Proceedings IEEE International Conference on Image Analysis and Processing (ICIAP), 0-7698-1948-2, 2003.
[41] Le H. S. and Li H., "Simple 1D Discrete Hidden Markov Models for Face Recognition," Springer Link Computer Science, Vol. 2849, pp. 41-49, September 2003.
[42] Eickeler S., Mller S., and Rigoll G., "Recognition of jpeg compressed face images based on statistical methods," Image and Vision Computing, Vol. 18, No. 3, pp. 279–287, 2000.
[43] Lucas S. M., "Face recognition with the continuous n-tuple classifier," In Proceedings of British Machine Vision Conference, September 1997.
[44] Kohir V. V. and Desai U. B., "Face recognition using DCTHMM approach," In Workshop on Advances in Facial Image Analysis and Recognition Technology (AFIART), Freiburg, Germany, June 1998.