Machine Learning Performance on Face Expression Recognition using Filtered Backprojection in DCT-PCA Domain.

Ongalo Pheobe[1], Huang DongJun[2] and Richard Rimiru[3]

**[1] School of Information Science and Engineering, Central South University
Changsha, Hunan, 410083, PR China**

**[2] School of Information Science and Engineering, Central South University
Changsha, Hunan, 410083, PR China**

**[3] School of Information Science and Engineering, Central South University
Changsha, Hunan, 410083, PR China**

Abstract

An accurate and robust transformed face descriptor that exploits the capabilities of filtered backprojection applied on Discrete Cosine Transform (DCT) and kernel Principal Component Analysis (PCA) methods is proposed. The method is invariant to rotation, variations in facial expression and illumination. Filtered backprojection constructs transform parameters from a set of projections through an image enhancing feature patterns that provide an initialization for subsequent DCT computations. DCT discards high-frequency coefficients that form least significant data to retain a subset of lower frequency coefficients visually significant in the image. The resulting coefficient features are mapped to lower dimensional space using PCA which extracts principal components that form the basis for the neural network classifier. Experiments were carried on JAFEE database and computed results compared with PCA and DCT approach. The results demonstrate significant improvements in results compared to other approaches.

Keywords: *Filtered backprojection, DCT, PCA, Neural Network.*

1. Introduction

Today's world where social media has taken centre stage, face to face communication is shifting towards internet, e-mail, text messaging and telephones. On the other hand faces naturally capture things that are difficult to embrace with spoken words and one such way is by use of face expressions. Face expressions are nonverbal interactive signals through which social information flows to the recipient displaying a person's affective states and intents. Interpreting these affective states gives an insight to a variety of information that can be derived from the face. Face expressions communicate emotions faster and more effective than words. As a result there is need to develop human centred user interfaces that respond readily to naturally occurring, multimodal, human communication with the capacity to receive, initiate course of action and monitor user feedback. Facial recognition accuracy depends heavily on how well the input images are compensated for pose, illumination and facial expression given that variations within a face are much more compared to variations between faces [1].

The existing facial expression recognition literature is broadly divided into three. The holistic approach [3-4], feature-based approach [5-6],[28] and hybrid approach [7]. Holistic approach also referred to as appearance-based approach identifies faces using global representations examples include Principal Component Analysis (PCA), also called Eigenfaces, Linear Discriminant Analysis (LDA), Support Vector Machines (SVM) and Neural Networks (NN). Feature-based approach extracts distinctive facial features such as eyes, mouth, nose as well as other control points from the face. These are processed to reduce input image to a vector of geometric features from which statistical pattern recognition techniques are employed to match faces. The disadvantage of feature approach is difficulty in automatic feature detection and the fact that its left upon the implementer to make arbitrary decisions to which features are relevant. In case the chosen features lack discrimination ability, no amount of subsequent processing can compensate for that intrinsic deficiency [2]. Hybrid approach separately extracts both local and global features and combines them for recognition. As a result it is expected that systems utilizing both local and global features can display more accurate results. Other systems like [8] employ the use machine based learning techniques while [9,10] make use of temporal dynamic information encoded directly in features.

Eigenfaces define feature space that drastically reduce the dimension of the original space, and face identification is carried out in this reduced space. Sirovich and Kirby [11] were the first to utilize Karhunen-Loeve Transform (KLT) to economically represent face images. They demonstrated that any face can efficiently be represented along the eigenface coordinate space, and that a face can be approximately reconstructed using a small collection of eigenfaces. Based on Sirovich and Kirby's findings Turk and Pentland [12,13] noted that projections along eigenfaces could be used as classification features to recognize faces. In [14], authors improved PCA by adding operations to standardize faces with respect to position and size. Also in [13] the authors used PCA on particular features of a face. The features became part of the "feature space," and a distance-to-feature-space (DFFS) metric was used to locate them in an image. This localization served as a pre-processing stage for later normalization, cropping, and classification [1]. Since then, PCA has become a popular method for face recognition. Daw-Tung lin [15] proposed the use of Hierarchical Radial Basis Function Network model to classify facial expressions based on local feature extraction by PCA technique.

Discrete Cosine Transform (DCT) has excellent energy compaction property for highly correlated data this helps in reduction of feature vector dimension. Previously DCT has been used for feature extraction either in a holistic or feature based approach to provide compact subspace representation. In [16] Radon transform was exploited to enhance low frequency components. DCT was used to yield lower dimension feature space while the nearest neighbor classifier was used for classification. Ramasubramanian and Venkatesh [17] combined the use of DCT, PCA and the characteristics of the Human Visual System for face recognition. In [18] Dattatray and Raghunath exploited the use of Radon, DCT and kernel based learning for face recognition using three images per subject they obtained 99.05% on FERET database, 99.32% on ORL database and 99.6% using Yale database. In [19] PCA and LDA were used in DCT domain to derive facial features with reduced dimensionality while in [20] Radon transform based on Particle Swarm Optimization followed by PCA and LDA techniques for face recognition were used. PCA was used for dimension reduction while LDA was used to extract a set of basis vectors which maximize the ratio between class scatter and within class scatter. They achieved a recognition rate of 97.5%.

Approach and Motivation

This paper investigates an alternative holistic approach that can be used for face recognition and compares it to the popular PCA and DCT approach. Facial images are represented as a finite 2-D matrix having local variation in facial intensities resulting from different combinations of abrupt features found in the face. To estimate transform parameters that provide an initialization for subsequent DCT computations, we seek for an accurate mathematical model that can map facial image to the underlying low level dimension space having redundant data and noise removed. Radon transform was used along with the Fourier Slice Theorem (FST) referred to as *filtered backprobagation* in this study to enhance low frequency components in the image. The filtered elemental reconstructions corresponding to each frequency domain pattern are initialized for subsequent DCT computations. DCT transforms the image pattern discarding high-frequency coefficients that separate significant data from the least significant data. This retains a small subset of lower frequency coefficients that are visually significant in an image. PCA is then used to compute eigenvectors in the direction of the largest variance of the training vectors also called eigenfaces. Each eigenface is considered a feature representing a point in a high-dimensional "face". Classification is done using a neural network.

The rest of the paper is organized as follows: Section 2, gives the procedure for data acquisition and reconstruction. Mathematical relationship between Radon transform and Fourier Slice Theorem is also described. Section 3 explains Discrete Cosine Transform process, Section 4 explains PCA while Section 5 describes the applied back propagation neural network process. Section 6, describes the methodology, followed by section 7 which highlights the performance of proposed system based on experimental results. Finally section 8 gives the summary and conclusion.

2. Data Acquisition and Reconstruction

2.1 Image Acquisition and Processing

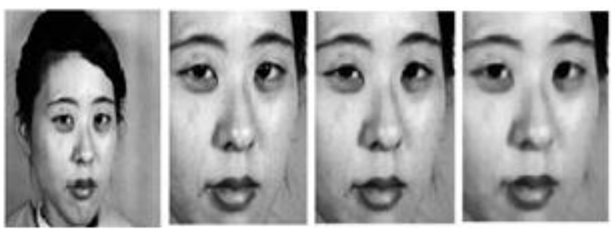
In this study static images were obtained from Japanese Female Facial Expression (JAFEE) database, which contains 213 images of 7 facial expressions posed by 10 Japanese female models [21]. To improve recognition performance against variations in face orientation, noise and illumination, various enhancement procedures were invoked to account for small perturbations in facial geometry and illumination as shown in figure1. Preprocessing measures taken include cropping images to eliminate extrinsic details like hair, neck, ears not central to face expression while eyes, mouth, and nose regions were retained. Image enhancement was done using morphological operators a process of smoothing irregularities and eliminating imperfections such as noise while distorting data of interest as little as possible [22]. Morphological operators opening and closing probe an image with a template called a structuring element. The structuring element is positioned at all possible locations in the image and is compared with other corresponding neighborhood pixels to test whether the element fits within the neighborhood. Morphological opening defined as opening of a set f by a structuring element b is the erosion of f by b followed by dilation of the result by b as given by Eq.(1).

$$f \circ b = (f \ominus b) \oplus b \quad (1)$$

Erosion filter eliminates isolated facial image details smaller than the structuring element assumed to be noise from the background, followed by dilation which thickens the object. The result of morphological opening is a smoothened object that has noise removed without affecting the shape and size of larger objects in the binary facial object. Morphological opening creates some gaps within an image this are fixed by performing a closing operation on the opening. Morphological closing defined as closing of a set f by a structuring element b is the dilation of f by b followed by erosion of the result by b as shown by Eq.(2).

$$f \bullet b = (f \oplus b) \ominus b \quad (2)$$

The closing operator smoothens the object border, merges together small features that are closer together and fills up the small gaps in the facial object. The output image Fig.1 (d) is a better quality image for the purpose of interpretation with features clearly visible.



(a) (b) (c) (d)

Fig. 1(a) Original test image (b) a cropped image from a; (c) morphological opening applied on image b; (d) morphological closing applied on image c;

2.2 Radon Transform

Radon transform collects line integrals across an image at different angles capturing directional local features present in the image. The Radon Transform of a 2 D function f(x,y) in (t,θ) plane can be defined as a series of line integrals through f(x,y) at different offsets from the origin [24]. Applying Radon transform on an image f(x,y) for a given set of angles relates to computing the projection of the image along these angles resulting into a profile of intensities. Using the geometry illustrated in Fig.2, the object is represented as a 2-D function f(x, y) and each line integral by (t,θ) parameters.

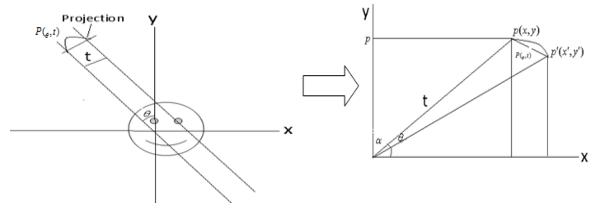


Fig. 2 Illustration of a set of projections through an object at a viewing angle forming the function P.

Radon transform of a 2-D function f(x,y) in (t,θ) is shown in Eq.(3).

$$p(t,\theta) = \int_{\infty}^{\infty}\int_{\infty}^{\infty} f(x,y)\delta(t - x\cos\theta + y\sin\theta)dxdy \quad (3)$$

Where t is the distance of a line from the origin, p(t,θ)is the sinogram, δ(·) is the Dirac delta, $\theta \in [0,\pi]$ is the angle of the line formed by the distance vector and $t \in [-\infty,\infty]$ is the perpendicular offset of the line from the origin. The δ function converts the two dimensional integral to a line integral dl along the line $x\cos\theta + y\sin\theta = t$. The transformed function (t,θ) is the sinogram of f(x,y). A Radon transformed 2D image projected at 125 degrees was computed in Matlab, giving directional lines present in the image see the sinogram Fig. 3. Each projection in the image forms a feature vector from which Radon transformed image intensities were extracted, backprojected and summed up to generate vectors used to approximate the shape of the original object see backprojected image in Fig. 3. To backproject a profile of intensities collected at an angle θ, we replicate the value p(t,θ) at all points along the direction normal to the profile for this angle as formally represented by Eq. 3.

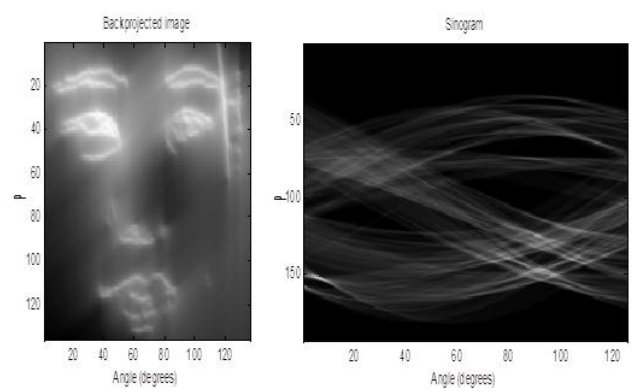


Fig. 3 Backprojected image and its Sinogram computed at 125°

From the results it is visible that backprojecting a Radon transformed 2D image as shown in Fig.3 yields unacceptably blurred image with high density at the center. This results from the overlapping of Fourier-transformed images around the low frequency region. To correct the artifact reformulation of the backprojection approach is inevitable. To do this we use Fourier Slice Theorem.

2.3 Fourier Slice Theorem (FST).

The Fourier Slice Theorem is derived by taking one-dimensional Fourier transform of a parallel projection and noting that it is equivalent to a slice of the two-dimensional Fourier transform of the original object. It follows that a Fourier transform $p_\theta(w)$ of a projection $p_\theta(t)$ through an image f(x,y) yields a two dimensional Fourier transform of the image f(u,v) evaluated along the polar lines (wcosθ, wsinθ). Given projection data, we can estimate the object by performing a 2D inverse Fourier transform as follows [24-25].

i. Calculate the inverse of 2D Fourier transforms in the Cartesian co-ordinate.

$$F(x,y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(u,v)e^{-2\pi i(ux+vy)}dudv \qquad (4)$$

ii. Then switch to polar coordinate, by letting u=wcosθ, v=wsinθ, which has Jacobin w. Changing differentials using dudv=w dwdθ. The inverse Fourier becomes;

$$f(x,y)=\int_0^{2\pi}\int_0^{\infty} wF(\mathrm{w}\cos\theta,\mathrm{w}\sin\theta)e^{-2\pi i\mathrm{w}(x\cos\theta+y\sin\theta)}dwd\theta \qquad (5)$$

iii. Manipulation of integration along full rays around a semi-circle limits yields

$$f(x,y)=\int_0^{\pi}\int_0^{\infty}\|w\|F(\mathrm{w}\cos\theta,\mathrm{w}\sin\theta)e^{-2\pi i w(x\cos\theta+y\sin\theta)}dwd\theta \qquad (6)$$

From the projection slice theorem

$$P_\theta(w)=F(w\cos\theta,w\sin\theta) \qquad (7)$$

Substitute in Eq. (6) to get,

$$f(x,y)=\int_0^{\pi}\int_0^{\infty}\|w\|P_\theta(w)e^{-2\pi i w(x\cos\theta+y\sin\theta)}dwd\theta \qquad (8)$$

iv. To ensure that Eq.(8) is indeed a filtered back-projection; filter the projections $p_\theta(t)$ using a ramp filter ||w|| to generate filtered projections $P_\theta^w(t)$:

$$P_\theta^w(t)=\int_{-\infty}^{\infty}\|w\|P_\theta(w)e^{-2\pi i w t}dw \qquad (9)$$

Build up f(x,y) by smearing the filtered projections back across the image.

$$f(x,y)=\int_0^{\pi}P_\theta^w(\mathrm{x}\cos\theta+\mathrm{y}\sin\theta)d\theta \qquad (10)$$

Evaluation of the integrand in Eq.(9) requires one dimension interpolation at $t=x\cos\theta+y\sin\theta$. The first part of the filtered back-projection algorithm implementation uses a continuous ramp filter. The second part of reconstruction is performed as per Eq. 10, summing up elemental contributions from each filtered projection to build up the grey scale values for (x,y) pixel. The filtering process is discretized to implement the algorithm. This is accomplished via the Fourier Slice Theorem. The assumption is that the projections are band-limited with band width W, that is $P_\theta(w)$ =0 whenever $\|w\|>W$, where W depends on the sampling size τ. If $\tau\le 1/(2W)$ then sampling theorem requires:

$$P_\theta(t)=\sum_{k=-\infty}^{\infty}p_\theta(k\tau)\frac{\sin 2\pi\mathrm{W}(\mathrm{t}\text{-}k\tau)}{2\pi\mathrm{W}(\mathrm{t}\text{-}k\tau)} \qquad (11)$$

Using this representation for a projection the ramp filter in Eq.(9) is rewritten as:

$$P_\theta^w(t)=\sum_{k=-\infty}^{\infty}p_\theta(k\tau)\int_{-W}^{W}\|w\|\left[\int_{-\infty}^{\infty}\frac{\sin 2\pi\mathrm{W}(\mathrm{t}'\text{-}k\tau)}{2\pi\mathrm{W}(\mathrm{t}'\text{-}k\tau)}e^{2\pi i w t'}dt'\right]e^{-2\pi i w t}dw \qquad (12)$$

The inner integral in this expression is the Fourier transform of *sinc* function it can be reduced to $\tau e^{2\pi i w k\tau}$. The filtered back projection is expressed as

$$P_\theta^w(t)=\tau\sum_{k=-\infty}^{\infty}p_\theta(k\tau)\int_{-W}^{W}\|w\|\cos\left(2\pi w(k\tau-t)\right)dw \qquad (13)$$

Since only the even part survives integration over the interval [-W,W] continuing with discretization process, it is desirable to generate sample points of the filtered projections $P_\theta^w(t)$. Assuming that $\tau=1/(2W)$, we set t=jτ, in the previous equation to get;

$$p_\theta^w(j\tau)=\tau\sum_{k=-\infty}^{\infty}p_\theta(k\tau)\int_{-W}^{W}\|w\|\cos(2\pi w\tau(k-j))dw$$

$$=\tau p_\theta(j\tau)W^2+2\tau\sum_{\substack{k=-\infty\\k\ne j}}^{\infty}p_\theta\int_0^W w\cos(2\pi w\tau)(k-j))dw$$

$$=\frac{1}{\tau}\left[\frac{1}{4}p_\theta(j\tau)-\sum_{\substack{k=-\infty\\(k-j)odd}}^{\infty}\frac{p_\theta(k\tau)}{\pi^2(k-j)^2}\right] \qquad (14)$$

From Eq.(14) it is evident that apart from the constant $1/\tau$ the filtered projection data $\{P_\theta^w(j)\}, j=-\infty,\ldots,\infty$ can be obtained by convolving the projection data $\{P_\theta(k)\}, k=-\infty,\ldots,\infty$ with the kernel h(j) given by

$$h(j)=\begin{cases}\frac{1}{4} & \mathrm{j=0}\\ 0 & \mathrm{j=even}\\ \frac{-1}{\mathrm{j}^2\pi^2} & \mathrm{j=odd}\end{cases} \qquad (15)$$

The outcome of filtered backprojection is an image with contrasting features (high-frequencies) required for face expression recognition being emphasized as shown in Fig. 4, while blurring (low-frequencies) are minimized.

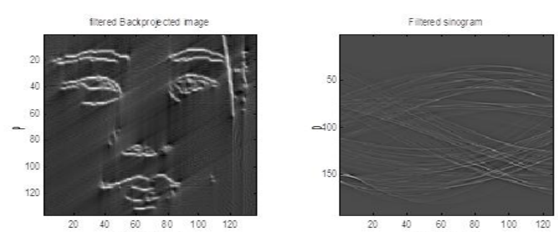


Fig. 4 Filtered backprojected image and its Sinogram computed at 125°

One advantage of filtered backprojection algorithm over frequency domain scheme is that reconstruction procedure begins as soon as the first projection is measured. This speeds up reconstruction procedure reducing the amount of data that must be stored at any given time [24].

3. Discrete Cosine Transform (DCT)

Discrete Cosine Transform is a popular linear projection technique employed in different applications for feature extraction. The superiority of DCT to PCA is that DCT can be realized in a single image or signal, while PCA depends on training samples. DCT is widely used technique in many standards of image coding and compression, like JPEG2000 and MPEG. DCT has the property that, for a typical image, most of the visually significant information about the image is concentrated in a few coefficients. Extracted DCT coefficients can be used as a type of signature that is useful for recognition tasks such as face recognition [1]. Face images have high correlation and redundant information which cause computational burden in terms of processing speed and memory utilization. DCT can be used to transform images from the spatial domain to the frequency domain. Since low frequency components are more visually significant in an image than higher frequencies DCT can been used to discard high-frequency coefficients and quantize the remaining subset of coefficients. This reduces the data volume without sacrificing too much image quality. The 2D-DCT of an M × N matrix A can be defined using Eq. (16).

$$B_{pq} = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos\left(\frac{\pi(2m+1)p}{2M}\right) \cos\left(\frac{\pi(2n+1)q}{2N}\right), \quad (16)$$

Where, $0 \le p \le M-1$, $0 \le q \le N-1$

The values B_{pq} are the DCT coefficients. DCT is an invertible transform, and its two-dimensional inverse discrete cosine transform is given by Eq. (17):

$$A_{mn} = \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} \alpha_p \alpha_q B_{pq} \cos\left(\frac{\pi(2m+1)p}{2M}\right) \cos\left(\frac{\pi(2n+1)q}{2N}\right), \quad (17)$$

Where, $0 \le m \le M-1$, $0 \le n \le N-1$.

The values $\propto_p$ and $\propto_q$ are given by Eq. (18)

$$\alpha p = \begin{cases} \sqrt{\frac{1}{M}}, p = 0 \\ \sqrt{\frac{2}{M}}, 1 \le p \le M-1 \end{cases}$$

$$\alpha q = \begin{cases} \sqrt{\frac{1}{N}}, q = 0 \\ \sqrt{\frac{2}{N}}, 1 \le q \le N-1 \end{cases} \quad (18)$$

M and *N* are the row and column size of A, respectively. DCT tends to concentrate information, making it useful for image compression applications. Applying DCT to an input sequence decomposes it into weighted sum of basis cosine sequences. To extract DCT coefficients filtered backprojected image was taken as the input image and its DCT computed. The upper left corner of a 2D-DCT matrix contains the most important values, which correspond to low-frequency components within the processed image block. These were used to give a significant reduction in number of feature vectors needed for subsequent processing. For reconstruction DCT inverse Eq. (17) is used. In order to reduce the amount of storage and compute similarity between images for face expression recognition PCA was used. PCA computes a small set of eigenvectors with top Eigenvalues used to build principal components for the feature space.

4. Principle Component Analysis (PCA)

Principal Component Analysis is known for its dimension reduction ability. It uses the least number of dimensions but keeps most of the facial information. Due to its simplicity and robustness, PCA was chosen as the baseline algorithm for face recognition grand challenge (FRGC) evaluation [26]. In this study PCA is used to achieve dimensional reduction by extracting the most representative features of facial data. By keeping low order principal components and ignoring higher ones reduces not only the image size but also the data. To compute PCA, frequency components from DCT were mapped as feature input matrix. First the mean vector of the vector population was computed followed by an approximation of the covariance matrix from a set of feature image matrix. Next eigen vectors and eigenvalues for the covariance transformation were obtained. Eigen vectors are invariant in direction during a transformation as a result they are used to form principal components that represent the dataset. These principal components are called Eigenfaces in Turk and Pentland face detection application and Eigen vehicles in Zhang et al. vehicle detection application [27] they are stored in the database during training for reference. For derivation steps used to compute PCA in this work interested readers are referred to [28]. To extract features that are invariant to illumination and rotations, PCA was used to compute a small set of eigenvectors with top Eigenvalues used to build up image characteristic. In this study, we used Q top eigenvectors where Q represents the number of important features from the eigenspace.

5. Backpropagation Neural Network

A feed forward back propagation Neural Network was used to train the network to recognize face expressions as shown in figure 5. Neural network with biases, sigmoid level, and linear output layer are capable of approximating any function with a finite number of discontinuities. Gradient descent with adaptive learning rate backpropagation was implemented using traingda function in Matlab. The input vector is weighted with appropriate weight matrix. The sum of weighted inputs and bias form input to the transfer function j. Neurons use the differentiable transfer function j to generate the output. As learning progresses across a feed forward back propagation neural network, hidden neurons discover the salient features that characterize the training data from the nonlinear transformed input data that results to a feature space. From this new feature space the classes of interest are separated. The output layer gives rise to facial expressions. Tansig transfer function was used as an activation function for hidden neurons and purelin transfer function was used for output neurons. Each column of matrix P was independent from the other columns resulting in patterns that form feature vectors. To train a NN, the input feature pattern was applied as a stimulus to the first layer of network units, which was propagated through each hidden layer adjusting the weights and bias of the network and testing the training set until an output was generated. The number of epochs were set to 20000, Mean Square Error (MSE) performance function was computed using the Eq. (19)

$$MSE = \frac{1}{N_p} \sum_{i=1}^{N_p} (k_i^a - k_i^d)^2 \qquad (19)$$

Where N_p is the number of training patterns in the training set. k_i^a is the actual output of the network for the input pattern i and k_i^d is the desired output of the network for pattern i. The actual network outputs are subtracted from the desired outputs and an error vector is produced. This error vector is the basis for the back propagation step. Errors are passed back through the network by calculating the contribution of each hidden processing layer and deriving the corresponding adjustment needed to produce the correct output. The process is repeated, layer by layer, until each node in the network receives an error signal that describes its relative contribution to the total error. Based on the error signals received, connection weights are readjusted to cause the network to converge towards a state that allows all the training patterns to be encoded.

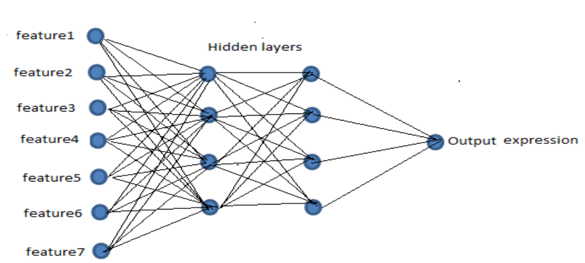


Figure 5. An illustration of neural network training architecture.

6. Methodology

To assess the validity and efficiency of the proposed approach experiments were conducted on Japanese Female Facial Expression (JAFFE) database. The database contains 213 images of 7 facial expressions posed by 10 Japanese female models. Ten individuals posed 3 to 4 expressions of each of the seven facial expressions; happiness, sadness, surprise, anger, disgust, fear and normal expression. JAFFE database presented include these images with minor rotation of camera axis and variations in head poses. Face recognition accuracy depends heavily on how well the input images have been compensated for pose, illumination and facial expression.To improve on direction low-frequency components, Radon transform was applied on preprocessed image data. Being a line intergral Radon transform acts like a low pass filter amplifying low frequency components in the image. Fourier Slice Theorem was then used before image intensities were extracted back-projected and the vectors summed up to generate the approximate shape of the facial object. This output was used to initialize DCT computations. DCT was used to discard high-frequency coefficients and quantize the remaining lower frequency coefficients reducing the data volume without sacrificing too much image quality. To recognize an input face, PCA is used on the subset of DCT coefficients to compute eigenvectors in the direction of the largest variance of the training vectors. During training 136 feature vectors were used while testing phase consisted of 70 feature vectors. Images that were used in the testing set were not included in the training set. A neural network was then trained using the outcome of the PCA process. The output generated was used as a feature map to provide an indication of the presence or absence of face expression feature combinations at the input. For purposes of evaluation and comparison, the following investigations were carried out.

i. Testing the number of Radon projections on face expression recognition
ii. Testing the effect of Morphological opening and closing on filtered backprojection approach for determination of emotional state from facial expression recognition.
iii. Testing the performance of PCA approach used alone for determination of emotional state from facial expression recognition.
iv. Testing the performance of DCT approach used alone for determination of emotional state from facial expression recognition.
v. Testing the effect of white noise robustness on facial data for determination of emotional state from facial expression recognition.

7. Results and Conclusion

7.1 Effect of number of Radon projections

In the proposed approach Radon transform is used to derive directional features with facial image data being projected at 125 degrees. To derive the best projection value, a series of experiments were carried out varying the projection degree between 30 and 180. Table 1 shows results obtained from reconstruction of facial image data for filtered backprojected enhanced data. The results reveal a steady increment in percentage recognition as the number of projections increase until 125 thereafter recognition rate starts to decline with any additional projections. Recognition accuracy was computed using Eq. (20).

$$\text{Recognition Rate} = \frac{\text{Number of Correct Images Classified}}{\text{Total Number of Classifications}} \times 100 \quad [20]$$

Table1: Effect of number of Radon projections on face expression recognition

No of Radon Projections	*Recognition Rate*
30	94.49
40	93.73
50	94.76
60	95.97
70	96.95
80	96.97
90	97.99
100	98.19
125	**98.99**
150	97.79
180	97.39

7.2 Effect of morphological processing

To evaluate the impact of pre-processing undertaken, we experimented with both morphologically enhanced and non-enhanced images. Non enhanced images gave their highest recognition rate at 97.79% while morphological processed image data gave 98.99% accuracy in determination of emotional state from facial expression recognition as shown in table 2.The confusion matrices presented by table 3 and table 4 present best performances of the experiments carried out in more details. The diagonal entries show the percentage of correct classification for each class and the scores off the diagonal entries show misclassification.

Table 2: Effect of Morphological processing on facial data for expression recognition

No of Radon Projections	**(%) Recognition Rate**	
	Non enhanced data	*Enhanced data*
30	93.69	94.49
40	94.51	93.73
50	95.96	94.76
60	96.36	95.97
70	96.59	96.95
80	96.97	96.97
90	97.39	97.99
100	**97.79**	98.19
125	97.77	**98.99**
150	97.58	97.79
180	*97.15*	*97.39*

Table 3: Confusion Matrix for enhanced filtered back propagation facial data in DCT-PCA domain.

Face expression	***Predicted Emotions***						
	Ang	Dis	Fea	Hap	Nor	Sad	Sur
Ang	100						
Dis		100					
Fea		1.4	97.2				1.4
Hap				98.6	1.4		
Nor					100		
Sad				2.9		97.1	
Surp							100

Recognition rate =98.99%

Table4: Confusion Matrix for non-enhanced filtered backpropagation facial data in DCT-PCA domain.

Face expression	**Predicted Emotions**						
	Ang	Dis	Fea	Hap	Nor	Sad	Sur
Ang	95.8	1.4		1.4		1.4	
Dis		98.6				1.4	
Fea		1.4	95.7				2.9
Hap				97.2	1.4	1.4	
Nor					100		
Sad			1.4	1.4		97.2	
Surp							100

Recognition rate =97.79%

7.3 Performance of PCA and DCT approach

For comparative analysis we implemented PCA and DCT algorithms. In both the algorithms morphologically enhanced facial data were used. The results obtained are illustrated using the confusion matrix given in Table 5 and

Table 6. In all these algorithms the neural network is used as the classifier.

Table 5. Confusion Matrix for morphologically processed filtered backpropagation using PCA approach.

Face expression	***Predicted Emotions***						
	Ang	Dis	Fea	Hap	Nor	Sad	Sur
Ang	100						
Dis		100					
Fea		1.4	97.2				1.4
Hap				98.6	1.4		
Nor					100		
Sad				2.9		97.1	
Surp							100

Recognition rate =98.2%

Table 6: Confusion Matrix for morphologically processed filtered backpropagation using DCT approach.

Face expression	***Predicted Emotions***						
	Ang	Dis	Fea	Hap	Nor	Sad	Sur
Ang	100						
Dis		100					
Fea		1.4	97.2				1.4
Hap			1.4	95.8	1.4	1.4	
Nor					100		
Sad			1.4	1.4		97.2	
Surp							100

Recognition rate =96.6%

7.4 Noise robustness of the proposed approach

The robustness of the proposed approach towards zero mean white noise was tested. We added zero mean white noise with a variance of 0.002 in test images. No noise was added to training images and the results obtained are shown using the confusion matrix in table7.

Table 7: Confusion Matrix showing the effect of zero mean white noise on enhanced filtered back propagation facial data in DCT-PCA domain.

Face expression	***Predicted Emotions***						
	Ang	Dis	Fea	Hap	Nor	Sad	Sur
Ang	100						
Dis		100					
Fea		1.4	95.7				2.9
Hap				97.2	1.4	1.4	
Nor					100		
Sad			1.4	1.4		97.2	
Surp			1.4				98.6

Recognition rate =98.4%

Where facial expressions are depicted by; Ang=Angry, Dis=disgusting, Fea=fear, Hap=Happy, Nor=Nomal, Sur=Surprise and sad faces respectively.

8. Summary and Conclusions

In this study enhanced facial images from JAFEE database were used to accurately determine motional state facial expressions. Morphological operators were used eliminate noise and other irregularities in facial objects. Results demonstrate that preprocessed facial data combined with filtered backprojection using DCT-PCA algorithm can be used to effectively determine emotional state from facial expressions. DCT was used to extract visually significant low frequency components in an image while PCA was used to extract discriminative principle features used by a neural network to accurately determine motional state facial expressions. The feasibility of the proposed approach has been duly tested on JAFEE database. Results show that the proposed approach significantly outperforms other standard methods like DCT and PCA.

Acknowledgments

We thank JAFFE Database for providing us face images for the experiments. This work was supported by the National Natural Science Foundation of China (60873188).

References

[1] Ziad M. Hafed and Martin D. Levine, "Face Recognition the Discrete Cosine Transform", International Journal of computer vision 43(3), 167-188,2001.

[2] Rabia Jafri and Hamid R. Arabnia, "A Survey of Face Recognition Techniques", Journal of Information Processing Systems, Vol.5, No.2, June 2009 41

[3] M. Bartlett, G. Littlewort-Ford, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan,"Fully automatic facial action recognition in spontaneous behavior", In IEEE Int'l Conf. on Automatic Face and Gesture Recognition, pages 223–230, 2006.

[4] B. Jiang, M. Valstar, and M. Pantic, "Action unit detection using sparse appearance descriptors in space-time video volumes", In Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition, 2011. In print.

[5] S. Lucey, I. Matthews, C. Hu, Z. Ambadar, F. de la Torre, and J. Cohn. "AAM derived face representations for robust facial action recognition in Automatic Face and Gesture Recognition", 2006. FGR 2006. 7th International Conference on, pages 155 –160, 2006.

[6] M. Valstar and M. Pantic. Combined support vector machines and hidden markov models for modeling facial action temporal dynamics.In ICCV-HCI'07, pages 118–127, 2007.

[7] Y. Tian, T. Kanade, and J. Cohn, "Recognizing action units for facial expression analysis", IEEE Trans. Pattern Analysis and Machine Intelligence, 23(2), 2001.

[8] G. Zhao and M. Pietikainen," Dynamic texture recognition using local binary pattern with an application to facial expressions", IEEE Trans Pattern Analysis and Machine Intelligence, 2(6), 2007.

[9] Y. Tong, J. Chen, and Q. Ji,"A unified probabilistic framework for spontaneous facial action modeling and understanding", Transactions on Pattern Analysis and Machine Intelligence, pages 1–16, Dec 2010.

[10] T. Simon, M. H. Nguyen, F. D. L. Torre, and J. F. Cohn. "Action unit detection with segment-based svms. Computer Vision and Pattern Recognition", IEEE Computer Society Conference on, 0:2737–2744,2010.

[11] L. Sirovich and M. Kirby, "Low-dimensional Procedure or the Characterization of Human Faces," Journal of the Optical Society of America A: Optics, Image Science, and Vision, Vol.4, pp.519-524, 1987.

[12] M. Turk and A. Pentland, "Eigenfaces For Recognition," Journal Of Cognitive Neuroscience", Vol.3, pp.71-86, 1991.

[13] A. Pentland, B. Moghaddam, and T. Starner, "Viewbased and modular eigenspaces for face recognition," in IEEE Conference on Computer Vision and Pattern Recognition, 1994, pp.84-90.

[14] Akamatsu, S., Sasaki, T., Fukamachi, H. and Suenaga, Y. 1991, "A robust face identification scheme, KL expansion of an invariant feature space", In SPIE Proc. Intell. Robots and Computer Vision X. Algorithms and Techn., 1607:71–84.

[15] Daw-Tung lin, "Facial Expression Classification using PCA and Hierarchical Radial Basis Function Network", Journal of information science and engineering 22, 1033-1046 (2006).

[16] Dattatray V. Jadhav, Raghunath S. Holambe: Radon and discrete cosine transforms based feature extraction and dimensionality reduction approach for face recognition", signal processing 88(10), 2604-2609 (2008)

[17] D. Ramasubramanian, Y.V. Venkatesh, "Encoding and recognition of faces based on the human visual model and DCT", Pattern Recogn. 34 (2000) 2447–2458.

[18] Dattatray V. Jadhav, Raghunath S.Holombe, "Rotation, illumination invariant polynomial kernel fisher discriminant analysis using Radon and discrete cosine transforms based features for face recognition", Pattern Recogn. Lett. 31 (2010) 1002–1009.

[19] W. Chen, M.J. Er, S. Wu, "PCA and LDA in DCT domain", Pattern Recogn. Lett. 26 (2005) 2474–2482.

[20] Hamid, Waleed and Dr. Majed, "Face Recognition Using Improved FFT Based Radon by PSO and PCA Techniques", International Journal of Image Processing (IJIP), Volume (6) Issue (1) 2012.

[21] http://www.kasrl.org/jaffe_download.html

[22] Ongalo P. N. Fedha , Huang Dong Jun , Richard Rimiru, "A Neural Network Based Classifier for a Segmented Facial Expression Recognition System Based on Haar Wavelet Transform", Global Journal of Computer Science and Technology Volume XII Issue VII Version I 7 april 2012.

[23] A. C. Kak and Malcolm Slaney, "Principles of Computerized Tomographic Imaging, Society of Industrial and Applied Mathematics 2001".Also available on Http://www.slaney.org/pct/.

[24] Hugh Murrell, "Computer Aided Tomography",The Mathematica Journal, Vol 6, No. 2, pp.60-65 2001.

[25] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, et al. "Overview of the face recognition grand challenge", IEEE Conf. Computer Vision and pattern recognition, San Diego,CA, 2005.

[26] DS.Z. Li, R.F. Chu, S.C. Liao, L. Zhang, "Illumination invariant face recognition using near-infrared images", IEEE Trans. Pattern Anal. Machine Intell, 29 (4) (2007) 627–639.

[27] Kyungnam Kim,"Face Recognition using Principle Component Analysis", International Conference on Computer Vision and Pattern Recognition, pp.586-591.

[28] Sushil Kumar Paul, Mohammad Shorif Uddin , and Saida Bouakaz ," Extraction of Facial Feature Points Using Cumulative Histogram", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 1, No 3, January 2012

First Author: Ongalo Pheobe. Received B.Com. Degree from Bharadhidasan government college for women in 1997 India, and Master of Computer Application(MCA) in 2001 from Pondicherry university India. Since 2003 she has been a lecturer in the department of Computer Science, Egerton University, Kenya. She is currently working towards her Ph.D. degree at the School of Information Science and Engineering, Central South University, Changsha, China. Her research interests include multimedia, neural networks, image processing and pattern recognition.

Second Author :Huang Dongjun, is a professor of Central South University and a doctoral tutor. He currently heads the Department of Computer Engineering, College of Information Science and Engineering. He obtained a master's degree in computer science and technology in May 1996.He completed a PHD degree from Central South University in 2004. In 2007 to 2008 he was at British University of Glasgow graphics and computer vision research group as a visiting scholar. He is committed to teaching and research development, his area of interest includes networking, distributed computing, multimedia systems and applications. Over the last 10 years he has presided over the completion of the National Natural Science Foundation of the State and the school-enterprise cooperation projects 8. He got a Provincial Science and Technology Progress Award (ranked first), made two software copyright, and has published more than 40 academic articles in the "Journal of Software" Electronic Journal and IEEE CVPR well-known publications and conferences, EI, 20. Taught Multimedia Technology and Application "courses for undergraduates and graduate students. Teaching Achievement of Central South University (ranked first), Directive graduated design Outstanding Thesis for school, second prize two, has won a first prize Teaching Quality Excellence Award.

Third Author: Richard Rimiru received his B.Sc. degree from the Department of Mathematics and Computer Science, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya, in 1999; received his M.Sc. degree from the Department of Computer Science, National University of Science and Technology, Bulawayo, Zimbabwe, in 2002. He is currently working towards his Ph.D. degree at the School of Information Science and Engineering, Central South University, Changsha, China. His main research areas are artificial intelligence especially in bio-inspired computing, pattern recognition and image processing.