International Journal of Computer Science and Engineering Communications- IJCSEC. Vol.2.Issue.1, Feb 2014. ISSN: 2347–8586

Novel Feature Extraction for Face Recognition using Multiscale Principal Component Analysis

K.M.Poornima
JNN College of Engineering
Shimoga, Karnataka, India
kmpoornima@sify.com

Ajit Danti
JNN College of Engineering
Shimoga, Karnataka, India
prs192003@yahoo.co.in

S.K.Narasimhamurthy
Kuvempu university
Shankaraghatta, Karnataka,
n.murthy@yahoo.com

ABSTRACT
A method of face recognition based on multiscale principal component analysis (MSPCA) is presented in this paper. Initially face area is extracted from the given face image using Adaboost face detection algorithm. From the face area, regions of interest such as eyes, nose and mouth part are extracted by dividing it along horizontal and vertical directions. Then MSPCA is employed on these regions of interest to extract the features. Multiscale Principal Component Analysis (MSPCA) combines the ability of PCA to decor relate the variables by extracting a linear relationship with that of wavelet analysis to extract deterministic features and approximately decor relate the auto correlated measurements. MSPCA computes the principal component analysis (PCA) of the wavelet coefficients at each scale, followed by combining the results at relevant scales. K-Nearest Neighbor (k-NN) classifier is used for recognition. The proposed methodology exhibits better recognition rate when compared to conventional principal component analysis.

Keywords: *Face Recognition, Feature extraction, PCA, MSPCA, k-NN Classifier*

INTRODUCTION

Recently face recognition is attracting much attention in society of network multimedia information access. Areas such as network security, content indexing and retrieval, and video compression benefits from face recognition technology. Network access control via face recognition not only makes hackers virtually impossible to steal one's "password", but also increases the user-friendliness in human-computer interaction. Indexing and/or retrieving video data based on the appearances of particular persons will be useful for users such as news reporters and moviegoers. For the applications of videophone and teleconferencing, the assistance of face recognition also provides a more efficient coding scheme. A good survey on face recognition is found in [1]. Face recognition methods can be roughly divided into two different groups: geometrical features matching and template matching. In the first method, some geometrical measures about distinctive facial features such as eyes, mouth, nose and chin are extracted [1]. In the second method the face image is represented as a two-dimensional array of intensity values and is compared to a single or several templates representing a whole face. The earliest methods for template matching are correlation-based which are computationally very expensive and require great amount of storage. Since a few years, the Principal Components Analysis (PCA) method also known as Karhunen-Loeve method, is successfully used as feature extraction technique and also used to perform dimensionality reduction [2, 3, 4, 5]. The problem with PCA is high computation complexity.

Paper Type: Research. Category: Computer Engineering
Corresponding Author: K.M.Poornima, Email: kmpoornima@sify.com

In this paper, we propose a new method for face recognition based on multiscale principal component analysis (MSPCA). From the input face image face area is extracted using Adaboost face detection algorithm. Regions of interest (ROI) such as eyes, nose and mouth part are extracted by dividing the detected face area along horizontal and vertical directions. Features are extracted by employing MSPCA on these regions. Then k-NN classifier is used for classification by considering different values for k. Experimental results are presented for ORL, Grimace and Faces94 databases and demonstrated the efficiency of the proposed method.

SYSTEM ARCHITECTURE

The proposed system consists of two stages: training stage and testing stage. In the training stage face area and regions of interest (ROI) are extracted from the given input image. For each region MSPCA is applied to extract the features and same procedure is applied for all the input face images. A meaningful feature matrix is constructed from the extracted features. Same steps are repeated in the testing stage. Then k-NN classifier is employed in order to classify the test image into known person class.

The block diagram of the proposed face recognition (FR) system architecture is shown in figure 1.

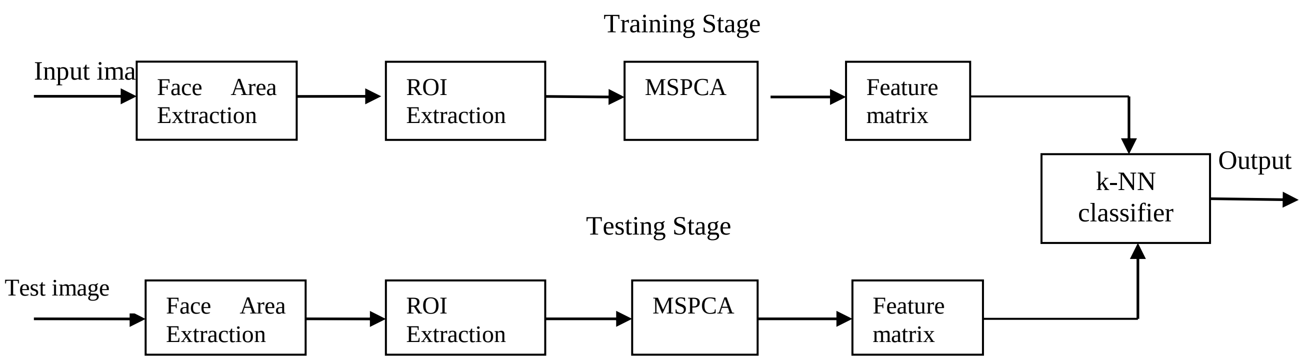

Figure 1. The proposed FR system architecture

Face area extraction

Face area is detected and extracted from the given input face image using Adaboost face detection technique [10]. AdaBoost is a machine learning algorithm which is formulated by Yoav Freund and Robert Schapire. AdaBoost is a Boosting algorithm, used to improve the performance of face detection. AdaBoost face detector includes: sub-window, Haar-like feature, classifier and classifier cascade. Initially an image is divided into sub-windows on which the detection algorithm is applied. Each sub-window has a classifier $h_t()$ that is a binary threshold function constructed from a threshold θ_t and a rectangle filter $f_t()$which is a linear function of the image, t being the identifier for each unique feature. In equation (1), x is the sub-window of an image; α_t and β_t are Boolean values signifying positive or negative votes in the cascade set during AdaBoost learning process.

$$h_t(x) = \begin{cases} \alpha_t & if \;\; f_t(x) > \theta_t, \\ \beta_t & Otherwise \end{cases} \qquad (1)$$

Haar-like features[10] is used to obtain the features from an image. Haar-like features have scalar values that represent differences in average intensities between two rectangular regions. Haar wavelet was first suggested by Alfred Haar, in which Haar wavelets uses the equation (2) to generate an expression for intensity over a function $\psi(t)$ and supports a scaling function (equation 3) for rapid summary of values.

$$\psi(t) = \begin{cases} 1 & 0 \le t \le \frac{1}{2} \\ -1 & \frac{1}{2} \le t \le 1 \\ 0 & Otherwise \end{cases} \quad (2)$$

$$\phi(t) = \begin{cases} 1 & 0 \le t \le 1 \\ 0 & Otherwise \end{cases} \quad (3)$$

Classifiers assigns each input value to a given set of classes, the classes are "Face" and "Not-Face". Threshold of the sums and difference of rectangular regions of data produced by any feature detector is used by classifiers, which includes Haar wavelets of rectangular gray-scale image values. Using the equation (4), the difference (F_{Haar}) between the fields of the rectangle can be calculated and translated to a potential detection. In the equation (4), $E(R_{black})$ is the intensity of the dark region, and $E(R_{white})$ is the intensity of the light colors in the white region of the rectangle, the difference between them is given as: $E(R_{black})$ - $E(R_{white})$. By dividing it with the standard deviation (equation (5)) on the rectangle containing all features, it is multiplied with width (w) and height (h) of the feature rectangle for scaling (i.e., the total number of pixels in the sub-window). The division is then applied to normalize the variance of the pixel value.

$$F_{Haar} = \left| \frac{E(R_{black}) - E(R_{white})}{w \times h \times \sqrt{\left| E(R_{\mu})^2 - E(R_{\mu}^2) \right|}} \right| \quad (4)$$

$$\sqrt{(\mathrm{mod}(E(R_{\mu})^2 - E(R_{\mu}^2)))} \quad (5)$$

Classifier Cascade is an algorithm for rapid detection in a sub-window, where a positive result from a classifier triggers the next classifier in line. A negative result from one classifier at any point during the cascade will immediately cause the rejection of the current sub-window. It is used to increase the computational efficiency and also to reduce false-positive rate. The detection/rejection cascade classifier algorithm works as follows, Haar-like feature is detected and then the sub-window is allowed to pass to the next step of detection. A candidate is a sub-window of the image, a rectangular section of a fixed size. The entire image is scanned for these sub-windows and a rapid summation of sub-images is carried out using equation (6), which uses a technique in which a single table is created where each pixel intensity is replaced by a value representing the sum of all the pixels contained in a rectangle of interest and the lower left corner of the image.

$$\text{Sum (X, Y)} = \sum_{\substack{x \le X \\ y \le Y}} image(x, y) \quad (6)$$

$$\text{D} = \prod_{i=1}^{K} d_i \quad (7)$$

$$\text{F} = \prod_{i=1}^{K} f_i \quad (8)$$

Total positive detection rate (D) is calculated by using equation (7) and total false positive rate (F) is calculated by using equation (8), where K denotes the number of classifier stages in the cascade, d denotes each stage's positive detection rate and f denotes each stage's classifier false-positive rate.

Face area is extracted from the original image (figure 2(a)) by employing Ada Boost face detector (figure2 (b)) is shown in figure 2(c) and it is used for regions of interest extraction.

Figure 2(a). Original image. Figure 2(b). After applying AdaBoost face detector. Figure 2(c). Resultant extracted face.

Regions of interest (ROI) extraction

To extract the regions of interest, facial grid (shown in figure 3(a)) is employed on the input face image (shown in figure 3(b)) and the face image with grid is shown in figure 3(c). According to this the input image is divided in horizontal and vertical direction. Initially the input image is divided into two equal halves in horizontal direction; we call these regions as top region and bottom region. The top region is again divided in horizontal and vertical direction into two equal halves to get region 1, region 2, region 3 and region 4 respectively.

The bottom region is divided into three equal halves to get region 5, region 6 and region 7. Out of seven regions we have considered only three regions (region 3, region 4 and region 6) for our experiments. We named region 3 as Left eye (LE), region 4 as Right eye (RE) and region 6 as Nose and Mouth (NM) which is shown in figure 3(d). Extracted Left eye (LE), Right eye (RE), Nose and Mouth (NM) regions are shown in figure 3(e).

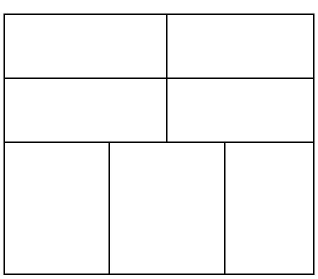

Figure 3(a). Facial grid

Figure 3(b). Input Face Image

Figure3(c) Facial grid on input Face Image

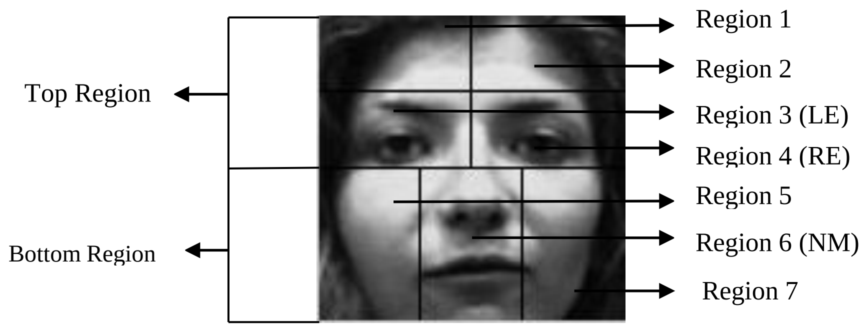

Fig 3(d). ROI Extraction using Facial grid

Figure 3(e). Extracted regions of interest (ROI).

Feature Extraction

Once the regions of interest such as eyes, nose and mouth part are extracted, the next step is feature extraction. To extract the features MSPCA is employed. The MSPCA methodology consists of decomposing each variable on a selected family of wavelets [6]. The PCA model is then determined independently for the coefficients at each scale. The models at important scales are then combined in an efficient scale-recursive manner to yield the model for all scales together. The aim of MSPCA is to reconstruct, starting from a multivariate signal and using a simple representation at each resolution level, a simplified multivariate signal [7]. The multiscale principal components generalize the normal PCA of a multivariate signal represented as a matrix by performing a PCA on the matrices of details of different levels simultaneously. A PCA is also performed on the coarser approximation coefficients matrix in the wavelet domain as well as on the final reconstructed matrix. By selecting the numbers of retained principal components, interesting simplified signals can be reconstructed. The MSPCA combines noncentered PCA on approximations and details in the wavelet domain and a final PCA. At each level, the most significant principal components are selected.

Since MSPCA consists of wavelets, to decompose the input image we need to select the wavelet family and the number of levels. In our experiments we have considered bi-orthogonal wavelet family and varied the levels from one to five. At each level, obtained principal components are considered and feature matrix is constructed for all the training images. For a test input face image, same procedure is applied as in the training stage. Using a classification algorithm, the extracted features are compared with the trained features to find the class to which this test face image belongs to. In our implementation k-nearest neighbor (k-NN) approach has been used for classification.

Classification

k-nearest neighbor (k-NN) is a non-parametric method, proposed by Cover and Hart [8] for classifying objects based on closest training samples in the feature space. k-NN is a type of instance-based learning, or lazy learning where the function is only approximated locally and all computation is deferred until classification. In k-NN, an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its k nearest neighbors. If $k = 1$, then the object is simply assigned to the class of its nearest neighbor. The performance depends on the value of k [9] which tells the number of neighbors taken for consideration. The neighbors are taken from a set of objects for which the correct classification is known. We have a training dataset D made up of x_i, $i \in [1, |D|]$ training samples.

The examples are described by a set of facial features F and any numeric features have been normalized to the range [0, 1]. Each training example is labeled with a class label $y_j \in Y$ .Our objective is to classify an unknown face example q. For each $x_i \in D$, we can calculate the distance between q and x_i by using equation (9).

$$d(q, x_i) = \sum_{f \in F} w_f \delta(q_f x_{if}) \qquad (9)$$

Where $\delta(q_{f,} x_{if})$ is a distance metric?

In our work, we considered the Euclidean distance function, varied the k value from 1 to 5 and recognition rate is compared.

EXPERIMENTAL RESULTS

Various experiments are conducted on ORL, Faces94 and Essex Grimace databases to verify the performance of the proposed technique.

Experiments using ORL database

The ORL database [11] contains only grayscale face images. The face images are of size 92 x 112. The database contains 10 different images each for 40 distinct persons. The face images were taken at different times by varying the lighting, facial expression and facial details. All the images were taken against a dark homogeneous background. Sample images of this dataset are shown in figure 4. For training purpose 5 face images from each are considered and the remaining set is used for testing. Face area is detected and extracted using Adaboost face detection technique. Regions of interest (ROI) are extracted by employing the facial grid and MSPCA is employed on regions of interest to extract the features. Since MSPCA is the combination of wavelets and PCA, we need to decompose the face images. For decomposition bi-orthogonal wavelet family is selected and the decomposition level is varied from 1 to 5 to obtain the wavelet coefficients. PCA is applied at each level for the obtained wavelet coefficients. In the testing phase, same steps are repeated to obtain wavelet coefficients and are transformed into PCA representational basis. Then k-NN classifier is used to perform the recognition task. Series of experiments are conducted by considering k parameter value as1, 3 and 5. The recognition result is given in Table1.

Figure 4. Sample images of ORL database

Table 1. Recognition Rate vs. No. of decomposition level (ORL)

Decomposition level	Recognition Rate		
	k=1	k=3	k=5
Level1	89.5	89.5	90.5
Level2	92	91.5	91
Level3	91	90.5	90.5
Level4	92	91.5	91
Level5	96	96	95

Experiments using Essex Grimace database

Essex Grimace database [12] consists of a sequence of 20 images each for 18 individuals consisting of male and female subjects. During the sequence, the subjects move their head and make grimaces which get more extreme towards the end of the sequence. Images are taken against a plain background, with very little variation in illumination. The images are in '.jpg' format and of size 180X200. Sample images of this database are shown in figure 5. Out of 20 images of each class, 10 images are used for training and the remaining used for testing. For feature extraction and classification same procedure is used as that of ORL database. The recognition result is given in Table2.

Figure 5. Sample images of Essex Grimace database

Table2. Recognition Rate vs. No. of decomposition level (Essex Grimace)

Decomposition level	Recognition Rate		
	k=1	k=3	k=5
Level1	91	91	90.5
Level2	89.5	91.5	91
Level3	94	94.5	94
Level4	92	98.5	98
Level5	95	95	94.5

Experiments using Essex Grimace database

Faces94 database [13] contains images of 153 individuals, 20 images per person. The subjects sit at fixed distance from the camera and are asked to speak, whilst a sequence of images is taken. Faces of this database show considerable expression changes but very minor variations in lighting or head position. Sample images are shown in figure 6. Out of 20 images of each person class, 10 images are used for training and remaining 10 images are used for testing. For feature extraction and classification same procedure is used as that of ORL database. The recognition result is given in Table3.

Table3. Recognition Rate vs. No. of decomposition level (Faces94)

Decomposition level	Recognition Rate		
	k=1	k=3	k=5
Level1	91	91	90.5
Level2	89.5	91.5	91
Level3	95	94.5	94
Level4	99	98.5	98
Level5	99	99	98.5

Figure 6: Sample images of Faces94 database

For all three databases highest recognition is obtained at level5 decomposition with k value is one or three. As the k value is increased beyond three, recognition rate starts decreasing.

CONCLUSIONS

A new method of face recognition based on multiscale principal component analysis (MSPCA) is proposed. The aim of multiscale PCA is to reconstruct a simplified multivariate signal, starting from a multivariate signal and using a simple representation at each resolution level. Here MSPCA is used as feature extraction as well as dimensionality reduction technique. The proposed method is experimented on ORL, Essex Grimace and Faces94 databases and exhibiting better recognition rate than the existing techniques. The proposed method can be experimented on other face databases and also number of decomposition levels can be increased.

REFERENCES

[1] W. Zhao, R. Chellappa, A. Rosenfeld, P.J. Phillips, “Face Recognition: A Literature Survey”, *ACM Computing Surveys*, pp 399- 458, 2003.

[2] M. Kirby, L. Sirovich, “Application of the Karhunen-Loeve Procedure and the Characterization of Human Faces”, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 12(1):103–108, 1990.

[3] A. Pentland, R.W. Picard, S. Sclaroff., “Photobook:Content-Based Manipulation of Image Databases”,In *Proceedings of the SPIE Storage and Retrievaland Video Databases II*, No. 2185, San Jose, 1994.

[4] M. Turk, A. Pentland.,”Eigenfaces for Recognition”, *Journal of Cognitive Science,* 3(1):71–86, 1991.

[5] C. L.Wilson, C. S. Barnes, R. Chellappa, S. A. Sirohey, “Face Recognition Technology for Law Enforcement Applications”, *NISTIR 5465, U.S. Department of Commerce*, 1994.

[6] Bakshi, B, "Multiscale PCA with application to Multivariate Statistical Process Monitoring", AIChE Journal, 44, pp. 1596-1610,1998.

[7] Aminghafari, M.; Cheze, N.; Poggi, J-M. (2006), "Multivariate de-noising using wavelets and principal component analysis", Computational Statistics & Data Analysis, 50, pp. 2381-2398, 2006.

[8] T.M. Cover and P.E. Hart, “Nearest neighbor pattern classification”, IEEE transaction in information theory, Vol 13, pp. 21-27, Jan 1967.

[9] Yu-Long Qiao; Jeng-Shyang Pan; Sheng- He Sun, “Improved K nearest neighbor classification algorithm”, Circuits and Systems, 2004, Proceedings. IEEE Asia-Pacific Conference, vol.2, pp. 1101- 1104, Dec. 2004

[10] K.T.Talele, Sunil Kadam and Atul Tikare. “ Efficient Face Detection using Adaboost”. *IJCA Proceedings on International Conference in Computational Intelligence (ICCIA2012)* iccia (10), March 2012.

[11]. http://www.cl.cam.ac.uk/Research/DTG/attarchive:pub/data/att_- faces.zip.

[12]. http://cswww.essex.ac.uk/mv/allfaces/grimace.zip.

[13]. http://cswww.essex.ac.uk/mv/allfaces/faces94.html