ISSN NO. 2320-5407

Journal homepage: http://www.journalijar.com
Journal DOI: 10.21474/IJAR01

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH

RESEARCH ARTICLE

Improved Face Recognition Method usingPCA.

Ramesh Kumar Verma[1], Arun Kumar[1], Deepak Kumar[1], Lucknesh Kumar[2].
1. Department of CSE.
2. Assistant Professor, Department of CSE.

Manuscript Info

Manuscript History:

Received: 14 February 2016
Final Accepted: 19 March 2016
Published Online: April 2016

Key words:
Face detection, faces recognition: feature based face detection, PCA, SIFT.

***Corresponding Author**

Ramesh Kumar Verma.

Abstract

This paper provides an example of the face recognition using PCA method and effect of Graph Based segmentation algorithm on recognition rate. Principle component analysis (PCA) is two or more variabletechniquethatanalyzesafacedatainwhichexperiencearedescribedbyseveralinter-correlateddependentvariables.Thegoalisto extract the important information from the face data, to represent it as a set of new statistically independent variables called principal components. The paper presents a proposed methodology for face recognition based on preprocessing face images using segmentation algorithm and SIFT (Scale Invariant Feature Transform) descriptor. The algorithm has been tested on 50 subjects (100images). The proposed method first waste stedon ESSEX face data base and next on own segmented face data base using SIFT-PCA. The experimental result shows that the segmentation in combination with SIFT-PCA has a positive effect for face recognition and accelerates the recognition PCA technique.

Introduction:-

Face recognition has been studied extensively for more than 10years. The face is the primary focus of attention in the society, playing a major role in intimate identity and emotion. Human face recognition plays significant role insecurity applications for access control, real time video surveillance systems, and robotics. The main problem of automatic human face recognition can be stated as follows: given an image of a human face (test set), compare it with stored models of a set of face images labeled with the person's identity (the trained data), and report the matching result. Face recognition can be divided into two core approaches namely, content-based and appearance based. Content-based Face recognition is based on the relationship between facial features like eyes, mouth and nose etc.[1]. In appearance based recognition the face is treated as a two dimensional pattern of intensity variation. The problem of human face recognition is a complex and highly challenging one having a variety of parameters including lightness, position orientation, expression, aging, head size, image darkening (eye glass effect), affectation, and face background [2]–[5]. This paper employed a new feature projection approach based on PCA method using segmentation algorithm-Belief Propagation. PCA is doing the optimum transformation for the differences between the classes.

The outline of the paper is as follows. In the Section II, an over view of image segmentation algorithm is given. Principal Component Analysis is described in SectionIII. In section IV and V, the process of SIFT algorithm and face recognition system are presented. Finally experimental results and implementation of the SIFT-PCA using image segmentation are introduced in SectionVI, and followed by conclusion in Section VII.

The problem of automated face and nodes recognition is relatively new and has not yet been fully volition. In recent years has been proposed a number of different approaches in treatment, localization and recognition of objects, such as the Gabor filters [1], principal component analysis, neural networks, evolutionary algorithms, Ada Boost algorithm, support vector machines, convolutional neural network [2] etc.

However, these approaches for object recognition have inadequate accuracy, credibility and speed in the real complex environment characterized by the presence of noise on images.

Image segmentation:-

The main goal of the image segmentation is split the entire image into set of segments that cover image. In this chapter, Graph Based segmentation algorithm will be presented

.

A. Graph Based Segmentation:-

The efficient graph based algorithm presented in [6] and [7] deals about problem in terms of a graph $G = (V, E)$ where nodesv_i Vrepresents pixels in the image, and the edges (v_i,v_j)Econnect certain pairs of neighboring nodes. Each edge (v_i, v_j) has responding weight $w\ (v_i, v_j)$ that is non-negativedis similarity measure between connected nodes by the edge (e.g. the difference in color, location, intensity, motion etc.). There are several techniques for image segmentation quality measurement. One of them is based on similarity of pixels in the same segment, and dis similarity if pixels in the different segments. Consequently, edges between two vertices in the same segment should have low weights and high weights for edges between two vertices in different segments here.

The efficient graph based algorithm has two important tasks, definition of difference between two components or segments, and these Condimportant task is definition of threshold function. The algorithm starts with the step where each segment contains only one pixel. In the next step, segments are iterative merged by the following conditions:

$$Diff(C_1,C_2)\leq Int(C_1)+T(C_1) \qquad (1)$$

$$Diff(C_1,C_2)\leq Int(C_2)+T(C_1) \qquad (2)$$

Where $Diff(C_1,C_2)$is difference between $C_1$and C_2

components, Int (C1) and Int (C2) are internal differences of C1 and C2 components, T(C1) and T(C2) are threshold functions of C1 and C2 components [7]–[9].
The threshold function controls the degree of the difference between two segments. The difference must be bigger than internal difference of segments. Threshold function is defined:

$$T(v) = \frac{k}{|C|} \qquad (3)$$

Where $|C|$ is the size of component *C*, *k* is constant, which manages size of the components. For small segments is required stronger evidence of a boundary. Larger *k* causes a creation of larger segments, smaller segments are allowed when there is a sufficiently large difference between them.

Principal component analysis:-

Principal Component Analysis (PCA) is mapping data to new space. PCA allows us to compute a linear transformation that maps data from a high dimensional space to a lower dimensional sub-space. The goal of PCA is to reduce the dimensionality of the data while retaining as much as possible of the variation present in the data set. Traditionally each image is first converted to a vector by row (or column) concatenation. Then PCA is applied for dimensionality reduction. The key idea of the PCA method is to find the vectors that best account for the distribution of face images with in the entire *m*x*n* image space. These vectors define the sub space of face images, they are called face space. Because these vectors are the eigenvectors of the covariance matrix corresponding to the original face images, and because they are face-like in appearance, they are called eigenfaces. We make this transformation images into the low sub space due to the speed up the computational time and recognition time. PCA sub space is defined by the eigenvectors of the covariance matrix and in this subspace, PCA algorithm examines the images. Examples of PCA sub space is illustrated in Fig. 1.

The following steps summarize the process PCA. Let a face image X (*x*, *y*) be a two dimensional *m*x*n* array of intensity values. An image may also be considering the vector of dimension m,n, so that a typical image of size 112 x 92 becomes a vector of dimension 10304. Let the training set of images $\{X1, X2, X3\ldots XN\}$. The average face of the set is defined by:

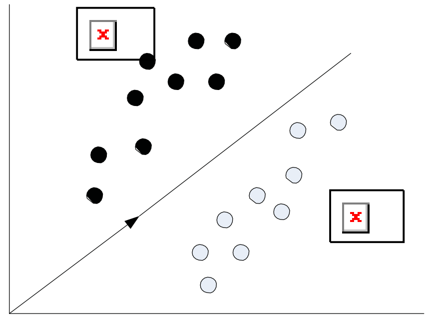


Fig.1.Effective separation data in PCA subspace.

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i \qquad (4)$$

Calculate the estimate covariance matrix to represent the scatter degree of all feature vectors related to the average vector. The covariance matrix C is defined by [11]:

$$C = \frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X}).(X_i - \bar{X})^T \qquad (5)$$

The Eigenvectors and corresponding eigenvalues are computed by using

$$C.V = \lambda.V, \ (V \in \boldsymbol{R_n}, V \neq 0) \qquad (6)$$

where V is the set of eigenvectors matrix C associated with its eigenvalue . Project all the training images of ith person to corresponding eigen-subspace:

$$Y_k^i = w^T(x_i), \qquad (1, 2, 3, 4\ldots.) \qquad (7)$$

In the testing phase each test image should be mean centered, now project the test image into the same eige space as defined during the training phase. This projected image is now compared with projected training image in eigen space. Images are compared with similarity measures. The training image that is the closest to the test image will be matched and used to identify. Calculate relative Euclidean distance between the testing image and the reconstructed image of ith person.

Sift algorithm:-

Scale Invariant Feature Transform (SIFT) is a local descriptor of image features insensitive to illuminant and other variants that is usually used as sparse feature representation [14]. SIFT features are features extracted from images to help in reliable matching between different views of the same object [13]. The extracted features are invariant to scale and orientation, and are highly distinctive of the image. They are extracted in four steps. The first step computes the locations of potential interest points in the image by detecting the maxima and minima of a set of Difference of Gaussian (DoG) filters applied at different scales all over the image. Then, these locations are refined by discarding points of low contrast. An orientation is then assigned to each key point based on local image features. Finally, a local feature descriptor is computed at each key point. This descriptor is based on the local image gradient, transformed according to the orientation of the key point to provide orientation invariance. Every feature is a vector of dimension 128 distinctively identifying the neighborhood around the key point [14].

A. SIFT-PCA:-

PCA is a standard technique for dimensionality reduction, which is well-suited to represent the key point patches and enables us to in early-project high-dimensional samples into a low-dimensional feature space [2].In other words, PCA-SIFT use PCA instead of histogram to normalize gradient patch. The feature vector is significantly smaller than the standard SIFT feature vector, and it can be used with the same matching algorithms. PCASIFT, like SIFT, also used Euclidean distance to determine whether the two vectors correspond to the same key point in

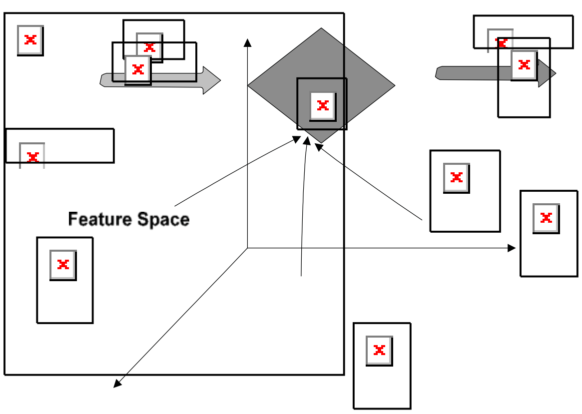


Fig.2.The example of Face Recognition System.

Different images. In PCA-SIFT, the input vector is created by concatenation of the horizontal and vertical gradient maps for the 41x41patch centered to the key point, which has 2x39x39=3042 elements. According to PCA-SIFT, fewer components requires less storage and will be resulting to a faster matching, they choose the dimensionality of the feature space, n = 20, which results to significant space benefits [2].

Face recognition system:-

An overview of the face recognition process is illustrated in Fig. 2. In the figure the gallery is the set of known individuals. The images used to test the algorithms are called probes. A probe is either a new image of individual in the gallery or an image of an individual not in the gallery. To compute performance, one needs both a gallery and

probe set. The probes are presented to an algorithm, and the algorithm returns the best match between the each probe and images in the gallery. TOestimated identity of a probe is the
Best match [16].

Experimental results:-

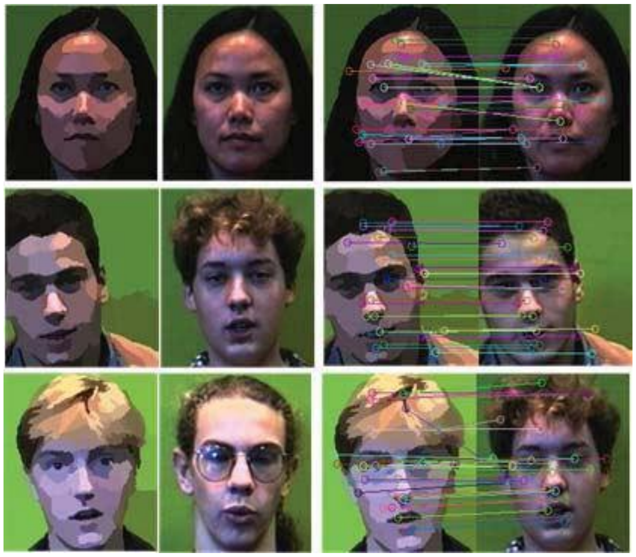
The ESSEX Face Database [12] has been used for the proposed experiment. For some subjects, images were taken at different times varying the lighting, facial expression (open, closed eyes, smiling or not smiling) and facial details(glasses, no glasses). It is of 100 images, corresponding to 50 subjects (namely, 2 images for each class). Each image has the size of 180 x 200 pixels. Some face images from the ESSEX database [9] are shown in Fig. 3.

The whole ESSEX Face Database [12] is used in training phase. As test phase we used segmented face images database. Some face images from the segmented database are shown in Fig. 4.
The proposed approach is implemented on the Pentium IV

2.8GHz. The experiment has been done on face data bases ESSEX with 50 training subjects (2 images per subjects). Firstly, we applied Graph Based segmentation algorithm for segmenting 50 images (one of each subject), as is shown in Fig. 4. Next, we have integrated this segmentation algorithm with the SIFT-PCA method. SIFT-PCA consists of two parts: SIFT part and PCA part. A s already stated, our experimental results proved that the SIFT descriptor is a very robust and reliable representation for the local neighborhood of an image

Fig.3. A pair of face images from ESSEX database.

Fig.4. Segmented faces images from ESSEX database.



PCA **SIFT-PCA**

Fig.5. Correct and incorrect face recognition results for segmented data base. Point (see Fig. 6).

We made two experiments to test the impact of segmentation for face recognition. In first experiment, we used classic images from ESSEX data base in the testing phase. We used one image per subject (therefore testing database contains 50 images). In second experiment, we used segmented images from the database in the testing phase. Eigen faces are calculated by using PCA algorithm and experiment is performed by varying the number of eigen faces used in face space to calculate the face descriptors of the images [11].

Fig.6.Test images with SIFT features shown as circles.

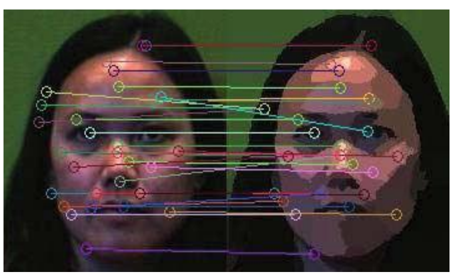


Table 1 processing time comparison

Method	TotalMatches	Total time [s]	100 matches time[s]
SIFT	352	7,2	4,9
SIFT-PCA	39	5.3	3.1

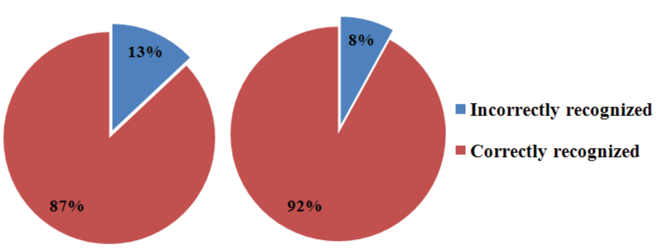


Fig.7.The results for segmented images using a) PCA and b) SIFT-PCA.

Matlab is used to implement eigenfaces. Eigen faces are computed for each face in the database and the eigenface of the query face is compared with all faces in the database. Comparison is done by computing Euclid and distance between two eigenfaces. Nearest neighbor of the query is retrieved which has got minimum distance.
The code for extracting SIFT features is available in Lowe's[14] website. The SIFT features are extracted from all faces in the database [17]. Then given a new face image, the features extracted from that face are compared against the features from each face in the database. A feature is considered as matched with another feature when the distance to that feature is less than a specific fraction of the distance to the next nearest feature. Further spatial topology is verified by Angle - Line Ratio (ALR) statistics [15] among the matched feature distributions. This ensures that we reduce the number of false matches. The face in the database with the largest number of matching points that agrees with the spatial distributions of the key points is considered as nearest face and is used for the classification of the new face.

In the first part of our experimental study, we load 100 face images into the training phase. Testing phase contains 50 face images. Next we selected one testing image and compared with other images in the training database using PCA algorithm. Moreover, proposed PCA algorithm with using images have reached recognition rate 84% and average recognition time was 10, 2 seconds. In the second part of our experimental study, we load100 face images into the training phase and using SIFT approach. Testing phase contains 50 segmented face images. We repeated the procedure and recognition rate was 92% and average recognition time was
10 seconds. The results of the experiment on ESSEXdatabase has been shown in Fig. 7.

Time is counted for the complete processing which includes feature detecting, matching and recognition. Table I show that SIFT – PCA is the fastest one, SIFT is the slowest but it finds most matches.

Table2 Faceidentificationrate

Method	SIFT[%]	PCA[%]	SIFT-PCA[%]
FindMatch	89	74	92

Some resultant face images from both PCA and SIFT are shown in Fig.5. Figure 5 shows two false positives and one correct positive retrieve during PCA and two correct positives retrieved and one false positive using SIFT - PCA approach. The face **identification rate is shown in Table II. It is evident from Table II that SIFT performs better in face identification even under deliberate**modifications.
In order to evaluate performance of the system we input each query at a time. The identification rate is computed as follows

$$M = \frac{P}{Q} \qquad (8)$$

Where P is number of times correct positives retrieved and Qisnumber of queries.

Conclusion:-

The paper presents Graph Based segmentation algorithm and face recognition approach for face identification based on PCA and SIFT features. The proposed approach is compared with eigenfaces and proved its excellence through experiments. As an extension, we are testing the use of SIFT features for retrieval of correct face with other forms of face representatives. Test results gave are cognition rate of about 87% for ESSEX database using PCA and 92% for our segmented database in combination with SIFT-PCA.
We have introduced a feature extraction technique from still images, which have been evaluated on database ESSEX and our segmented database. This technique has been found to be robust against extreme expression variation as it works efficiently on database. We have shown that the segmentation using SIFT-PCA has a positive effect for face recognition and accelerates the recognition PCA technique. Future work is suggested towards exploring the combination with LDA, 2DPCA or CCA.

References:-

1. P. N. Belhunmeur, J. P. Hespanha, D. J. Kriegman, "Eigen face svs Fisher faces Recognition Using Class specific linear projection", *IEEE Trans Pattern Analysis and Machine Intelligence*, 1997.
2. W. Zhao, R. Chellappa, A. Rosenfeld, "*Face recognition: a literature survey*", ACM Computing Surveys, Vol. 35, No. 4, December 2003, pp. 399–458.
3. K. C. Kwak, W. Pedryez, "Face Recognition : A Study in Information Fusion using Fuzzy Integral", *Pattern recognition Letters*, vol. 26, pp.719-733,2005.
4. H. Bernd, T. Koshizen, "Components for Face Recognition", *Proceedings of the 6th International Conference on Automatic Face and Gesture Recognition*, Seoul, Korea, pp. 153–158, 2004.
5. S.Jianbo,J.Malik,"Normalized cutsand image segmentation", *Pattern Analysis and Machine Intelligence, IEEE Transactions*, vol. 22, no. 8, pp. 888-905, 2002.
6. M. Benco, R.Hudec, "The advanced image segmentation techniques for broadly useful retrieval in large image database", *Proceedings NSSSIX*, Tatranske Zruby, Slovak Republic, pp. 40-44, 2006.
7. P. F. Felenzwalb, D. P. Klette, "*Efficient graph - based image segmentation*", *International Journal of Computer Vision*, [Online], vol. 59, No. 2, 2004.
8. R. He, Y.Zhu "A hybrid image segmentation approach based on mean shift and fuzzy c – means", *Asia – Pacific Conference on Information Processing* [Online], 2009.
9. L. Dong, P. Ogunbona, W. Li, G. Yu, L. Fan, G. Zheng "A fast algorithm for color image segmentation", *First International Conference on Innovative Computing, Information and Control* [Online], ICICIC 06, 2006.
10. M. Kirby, L. Sirovich, "Application of the Karhunen – Lo eve procedure for the characterization of humanfaces", *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 12, No.1, pp. 103-1071990.
11. J. Mazanec, M. Melisek, M. Oravec, J. Pavlovicova, "Support Vector Machine, PCA and L D Ain Face Recognition", *Journal of Electrical Engineering*, vol. 59, No. 4, pp. 203-209, 2008.
12. Essex Face Database: http://cswww.essex.ac.uk/mv/allfaces/index.html, 2008.
13. K. Mikolajczyk, D. Schmid, "Scale and Affine In variant Interest Point Detectors", *International Journal of Computer Vision, (IJCV)*, vol. 60, No. 1, pp. 63-86, 2008.
14. D. G. Lowe, "Distinctive image features from scale – in variant keypoints", *International Journal of Computer Vision*, vol. 2, No. 60, pp. 91-110, 2004.
15. H. R. Chennamma, L. Rangarajan, "Robust Near – Duplicate Image Matching for Digital Image Forensics", *International Journal of Digital Crime and Forensics*, 1(3), July, 2009.
16. W. Cooley, W. Lohnes, "PCA correlation", *P. R. Multivariate Data Analysis John Wiley & Son*, Inc., NewYork.
17. L. Bolecek, V.Ricny, "The Algorithm for shadows detection in profilometry applications", *Elektrorevue, Internet Journal for Electrical Engineering*, vol. 58, No. 4, pp. 1 - 7, November, 2011.