# A Hybrid of Principal Component Analysis and Partial Least Squares for Face Recognition across Pose

Ajay Jaiswal, Nitin Kumar, and R.K. Agrawal

School of Computer and Systems Sciences
Jawaharlal Nehru University, New Delhi, India
a_ajayjaiswal@yahoo.com {nitin2689,rkajnu}@gmail.com

**Abstract.** In this paper, we propose a simple and efficient hybrid approach based on the combination of principal component analysis and partial least squares. Principal component analysis is used to reduce the dimension of image in first step and partial least squares method is used to carry out pose invariant face recognition in second step. The performance of proposed method is compared with another popular method based on global linear regression on hybrid-eigenface (HGLR) in terms of classification accuracy and computation time. Experimental results on two well known publicly available face databases demonstrate the effectiveness of the proposed approach.

**Keywords:** Face recognition across pose, Partial least squares, Principal component analysis, Hybrid-eigenfaces, Linear regression.

## 1 Introduction

In controlled environment, the performance of face recognition system has already attained satisfactory levels. But still, there are many challenges associated with face recognition in presence of illumination variation, pose variation, occlusion etc. Face recognition across pose has been identified as one of the most challenging problem by research community [1]. Many methods have been proposed to tackle the problem of recognising faces across pose [2]. In recent years, few promising methods have been proposed such as global linear regression (GLR) [3], local linear regression (LLR) [3], GLR on hybrid eigenfaces (HGLR) [4] and face recognition based on partial least squares[5] etc. However, none of them is free from shortcomings.

HGLR involves generation of multiple virtual views to carry out face recognition. Hence, more storage and computation time is required. Similarly method based on partial least squares (PLS) performed on the images of larger size requires huge computation time. One way of reducing the computation time is to reduce the size of the image. But, it will lead to information loss in the image. Another approach is to reduce the dimensionality of the image with minimal loss of information. This can be achieved using principal component analysis (PCA) which reduces the redundant information of the image. Motivated by this, we propose to carry out face recognition in two steps. In first step, PCA is applied to reduce the dimensionality and in second step, face recognition is carried out with PLS. To check the efficacy, the performance of the proposed approach was evaluated in terms of classification accuracy and computation time. We also investigated the performance under different pose variations

and different size of training data. The rest of the paper is organized as follows. Section 2 describes GLR and hybrid eigenfaces briefly. PLS is described in section 3. The proposed approach is discussed in section 4. Experimental results on two well known publicly available face databases are included in section 5. Finally, section 6 contains conclusions and future work.

## 2      Global Linear Regression (GLR) and Hybrid Eigenfaces

GLR [3] is based on the assumption that there exists a linear mapping between non-frontal face images and frontal images of the same person. Therefore, it is possible to transform a non-frontal face image to its corresponding frontal view or vice versa with some approximation. Chai et al. [3] formulated a method for virtual view generation in frontal pose from an input face image under $p^{th}$ pose using GLR. This method involves two steps. In the first step, input face image in $p^{th}$ pose ($i^p_{in}$) is decomposed into linear combination of N face images under the same pose using linear regression with face images as basis vectors and is given by

$$i^p_{in} = c^p_1 i^p_1 + c^p_2 i^p_2 + c^p_3 i^p_3 + ... + c^p_N i^p_N \tag{1}$$

This can also be rewritten as

$$i^p_{in} = I^p c^p_{in} \quad \text{and} \quad c^p_{in} = (I^p)^\perp i^p_{in} \tag{2}$$

where $c^p_{in} = [c^p_1 \quad c^p_2 \quad ... \quad c^p_N]^T$ , $I^p = [i^p_1 \quad i^p_2 \quad ... \quad i^p_N]$ and $(I^p)^\perp$ is the pseudo-inverse of $I^p$ . In the second step, the corresponding virtual frontal view is obtained by linearly combining the face images under frontal pose using the same coefficient vector $c^p_{in}$ i.e. $i^0_{in} = I^0 c^p_{in}$ and $I^0 = [i^0_1 \quad i^0_2 \quad ... \quad i^0_N]$.         GLR on coarsely aligned images does not provide realistic images and contain some artefacts especially around the corner of mouth and eyes [4]. To overcome this drawback, Sharma et al. [4] proposed GLR on H-eigenfaces. Brief description of HGLR is given here. Consider a face database consisting of face images of N subject with each subject having face images under different poses. A ''mixed vector'' is formed by stacking the face image vectors of same person under different pose [4].  For example, suppose a frontal-face vector of $i^{th}$ subject is denoted by $x^0_i$, then the mix vector ($m_i$) is formed by appending the vector $x^p_i$ below vector $x^0_i$ one after another for all poses. The same was done for all the subjects and a matrix $M$ was formed having vectors $m_i$'s as columns, i=1, 2, …, N. The matrix $M$ is given as

$$M = \begin{bmatrix} x^0_1 & ... & x^0_N \\ ... & & ... \\ x^p_1 & \ddots & x^p_1 \\ ... & & ... \\ x^q_1 & ... & x^q_N \end{bmatrix} \quad and \quad g_i = \begin{bmatrix} h^0_i \\ ... \\ h^p_i \\ ... \\ h^q_i \end{bmatrix} \tag{3}$$

In next step, PCA is applied on **M** to determine k dominant eigenvectors, $\mathbf{g_i}$ (i=1,…, k) to be used as basis vectors. Further, each of these eigenvectors, $\mathbf{g_i}$, is divided into *r+1* sub-vectors, where *r* is the number of desired non-frontal viewpoints under which virtual views are required to be synthesized. The upper sub-vector belongs to the frontal face space and the subsequent lower parts belong to corresponding view-based face spaces in the same order in which they were stacked. These sub-vectors are termed as H-eigenfaces and used for generating virtual views.

## 3    Partial Least Squares (PLS)

PLS, based on the principle of statistical learning, is widely used in chemo metrics and bioinformatics etc. [6]. In recent years, it is also used in domains such as human detection [7], face recognition [8] etc. It models the relationship between two sets of variables by means of score vectors. All PLS methods work on the assumption that the observed data is generated by a process which is driven by a small number of latent variables [6]. Suppose N observed samples from two sets of variables are denoted as $\mathbf{X} = (\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_N})$ and $\mathbf{Y} = (\mathbf{y_1}, \mathbf{y_2}, \ldots, \mathbf{y_N})$ respectively, with $\mathbf{X} \in \Re^m$ and $\mathbf{Y} \in \Re^n$. Both **X** and **Y**, are normalized to zero mean. Using PLS, **X** and **Y** can be represented as [6]:

$$\mathbf{X} = \mathbf{TP'} + \mathbf{E}$$
$$\mathbf{Y} = \mathbf{UQ'} + \mathbf{F} \tag{4}$$

where **T** and **U** are N×p matrices of the p extracted score vectors. The m×p matrix **P** and the n×p matrix **Q** are the loading matrices and the **E** and **F** are the matrices of residuals whose sizes are N×m and N×n respectively. The objective of PLS is to find weight vectors **b** and **c** such that

$$[\text{cov}\,(\mathbf{t}, \mathbf{u})]^2 = [\text{cov}\,(\mathbf{Xb}, \mathbf{Yc})]^2 = \max_{|\mathbf{b}|=|\mathbf{c}|=1} \text{cov}\,(\mathbf{Xw}, \mathbf{Yr})]^2 \tag{5}$$

where **t** and **u** are the corresponding score vectors in **T** and **U**. Face recognition using PLS requires the coupled faces across different poses. If the training set is denoted as (**X**, **Y**) where **X** and **Y** correspond to two different poses, the samples in **X** and **Y** are coupled by identity [5] i.e. i[th] row in **X** and **Y** correspond to image of ith identity in frontal and non-frontal pose respectively.

## 4    Proposed Approach

The proposed approach involves PCA in first step to reduce the dimensionality without sacrificing much information and PLS in second step for better recognition. Let **IX** and **IY** represent faces in frontal and non-frontal pose respectively. First, we normalize **IX** and **IY** to have zero mean and unit variance. Then we apply PCA separately on frontal (**IX**) and non-frontal (**IY**) images to obtain two subspaces corresponding to frontal images and non-frontal images. After this, PLS is used on the coupled training set to get optimized loading vectors **P** and **Q**. In the testing phase, the loading

vectors are used to estimate score vectors for new samples. The score vectors are good pose invariant representation of faces. Face recognition across pose can be performed by directly matching the score vectors. Let us denote a gallery face (frontal) as **Ixt** and a probe face as **Iyt**. We reduce the dimension of **Ixt** and **Iyt** using PCA models obtained during training as given below.

$$xt = W_f Ixt$$
$$yt = W_p Iyt \tag{6}$$

where $W_f$ and $W_p$ are the transformation matrix obtained after applying PCA to **IX** and **IY** respectively. Based on Eq. (4) their corresponding score vectors **t** and **u** can be estimated using the loading vectors **P** and **Q**.

$$t = xt(P')^{-1}$$
$$u = yt(Q')^{-1} \tag{7}$$

The similarity s between **Ixt** and **Iyt** can be simply measured by the correlation between **t** and **u** which is given by $s = \dfrac{\langle t, u \rangle}{\|t\| \cdot \|u\|}$, where $\langle t, u \rangle$ denotes the inner product of

**t** and **u**. In the recognition, test face is classified to the gallery face with highest similarity value (s).

## 5      Experimental Setup and Results

To check the efficacy of the proposed approach, we carried out experiments on two publicly available face databases PIE [9] and FERET [10]. For all the experiments, first the images were coarsely aligned according to two eyes position and two extreme mouth corners. Then images were cropped and resized to 50×50. The performance is evaluated in terms of classification accuracy and computation time. The classification accuracy is obtained in terms of 2- fold and 3-fold cross-validations. Each cross-validation is repeated 10 times to obtain average classification accuracy. In our experiments, images of seven poses from the PIE database are used, which covers the pose yawing over ±45 degree and the pitching variations in depth [9]. The poses used in our experiments a0re: P07, P09, P05, P29, P37, P11, and P27. Each pose class includes 68 subjects. The FERET database subset used in our experiments contains images of 196 subjects under nine poses (ba, bb, bc, bd, be, bf, bg, bh and bi). Example faces of both the databases are shown in figure 1.



**Fig. 1.** Example faces of PIE (left) and FERET (right) database

### 5.1      Experiments Using HGLR

To obtain H-eigenfaces, more explicitly, in 2-fold cross-validation, images in one pose are divided into two parts. One part is used to generate H-eigenfaces. The virtual

views for other part are synthesized using GLR on H-eigenfaces. The virtual images are constructed for all the subjects with different number of H-eigenfaces. All experiments are carried out using single frontal image as a gallery image and two other poses as test images. Using gallery image, virtual frontal and non-frontal images in the poses present in test set are synthesized. Then, for each pose (including frontal), fifteen images under different illumination were generated using gray level transform resulting in a total of 15×3=45 virtual images. Then we used PCA+LDA to carry out face recognition. PCA is applied on the gallery images to obtain PCA subspace. All the virtual images are used as training samples for LDA. Before applying LDA, all the training images were first projected with PCA consisting of top few eigenfaces. The experiment is carried out with different number of eigenfaces. Euclidian distance is used to find the closest match of a probe image in all the experiments. Among the classification accuracies obtained with different combination of number of H-eigenfaces and number of eigenfaces, the best accuracy is reported. Similarly, we carry out 3-fold cross-validation experiments. Experiments on both the face databases are carried out in 3 modes. For PIE face database in mode 1, gallery set contains single frontal pose P27 and testing set comprises of images under poses P07, P09. Test set consists of poses P05 and P29 in mode 2. In mode 3, test set is replaced with images in poses P37 and P11. For FERET face database in mode 1, gallery set comprises of single frontal pose ba and testing set comprises of images under poses be and bf. In mode 2, probe set is replaced by poses bd and bg. In mode 3, test set consists of images in poses bc and bh. All the experiments were executed on Dell T3500 workstation running MATLAB 2010 on Windows 7 (64-bit) platform.

## 5.2    Experiments Using Proposed Approach

Images in frontal and one non-frontal pose are used at a time. Specifically, in 2-fold, images in one pose are divided into two parts. PCA is applied to first part of data to obtain transformation matrices with respect to frontal and non-frontal pose. Then PLS is applied to obtain loading vectors. Face recognition is performed on the second part of the data. Then training is done with second part and recognition is performed with first part. In this way face recognition is performed for all the subjects. In 3-fold cross-validation, two parts are used for training and face recognition is performed on third part. For PIE face database in mode 1, first we used frontal pose P27 and images under pose P07. Then images under pose P09 are used with frontal images. The average classification accuracy of two different poses (P07 and P09) is reported. Similarly, the average classification accuracy of poses in mode 2 and poses in mode 3 are determined. Similarly for FERET face database, the average classification accuracy of two different poses (be and bf) in mode 1, poses (bd and bg) in mode 2 and poses (bc and bh) in mode 3 are determined. Comparative classification accuracy of HGLR and the proposed approach for both the face databases are shown in Table 1. The comparative computation time is shown in Table 2. The following can be observed from Tables 1-2:

- The classification accuracy of the proposed method is significantly better in comparison to HGLR for both PIE and FERET database in all cases except in mode 1 of 2-fold cross-validation for FERET database.
- As pose variation increases (indicated by modes), the difference in classification accuracy of the two approaches increases significantly.
- Standalone PLS (without PCA) takes large computation time. With the use of PCA in first step of the proposed method, the computation time drops significantly while classification accuracy gets marginally affected. The computation time is significantly less in comparison to HGLR also. Similar observations are made for FERET face database.

**Table 1.** Comparative classification accuracy

| Face database | Cross-validations | mode1 | | mode2 | | mode3 | |
|---|---|---|---|---|---|---|---|
| | | HGLR | HPCAPLS | HGLR | HPCAPLS | HGLR | HPCAPLS |
| PIE | 2-fold | 88.68 | **99.63** | 80.37 | **99.49** | 57.87 | **96.25** |
| | 3-fold | 97.35 | **99.93** | 90.37 | **99.57** | 69.48 | **98.98** |
| FERET | 2-fold | **87.8** | 87.45 | 76.48 | **82.22** | 56.22 | **68.9** |
| | 3-fold | 90.66 | **91.64** | 82.7 | **88.22** | 65.56 | **76.96** |

**Table 2.** Variation in computation time for PIE face database(2-fold cross-validation)

| Approach | Approx. Time taken (in sec.) |
|---|---|
| PLS on images after applying PCA | **0.88** |
| HGLR | 98 .3 |
| PLS on image of size  50×50 | 13140 |

The classification accuracy (Standalone PLS) for PIE database in mode 1 was found to be 99.70% in 2-fold cross validation which is comparable to that of proposed approach.

## 6     Conclusions

Experimental results demonstrate that the proposed approach performs significantly better than HGLR approach in terms both classification accuracy and computation time. In the proposed approach, there is no need to generate virtual views which reduces both computational time and memory requirement. In future, some other feature extraction methods can be explored for dimensionality reduction to further improve the performance.

## References

1. Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face recognition: a literature survey. ACM Comput. Surv. 35(4), 399–459 (2003)
2. Zhang, X., Gao, Y.: Face recognition across pose: A review. Pattern Recognition 42, 2876–2896 (2009)

3. Chai, X., Shan, S., Chen, X., Gao, W.: Locally linear regression for pose-invariant face recognition. IEEE Trans. Image Process. 16(7), 1716–1725 (2007)
4. Sharma, A., Dubey, A., Tripathi, P., Kumar, V.: Pose invariant virtual classifiers from single training image using novel hybrid-eigenfaces. Neurocomputing 73, 1868–1880 (2010)
5. Li, A., Shan, S., Chen, X., Gao, W.: Cross-pose face recognition based on partial least squares. Pattern Recognition Letters 32(15), 1948–1955 (2011)
6. Rosipal, R., Kramer, N.: Overview and recent advances in partial least squares. Subspace Latent Struct. Feat. Select., 34–51 (2006)
7. Schwartz, W., Kembhavi, A., Harwood, D., Davis, L.: Human detection using partial least squares analysis. In: IEEE Internat. Conf. on Computer Vision, pp. 24–31. IEEE (2009)
8. Baek, J., Kim, M.: Face recognition using partial least squares components. Pattern Recognition 37(6), 1303–1306 (2004)
9. Sim, T., Baker, S., Bsat, M.: The CMU pose, illumination, and expression database. IEEE Trans. PAMI 25(12), 1615–1618 (2003)
10. Phillips, P.J., Wechsler, H., Huang, J., Rauss, P.: The FERET database and evaluation procedure for face recognition algorithms. Image Vision Comput. 16(5), 295–306 (1998)