Irwin King and Lei Xu. Localized principal component analysis learning for face feature extraction and recognition. In *Proceedings to the Workshop on 3D Computer Vision '97*, pages 124–128, Shatin, Hong Kong, 1997. The Chinese University of Hong Kong.

Localized Principal Component Analysis Learning for Face Feature Extraction and Recognition

Irwin King
king@cse.cuhk.edu.hk

Lei Xu
lxu@cse.cuhk.edu.hk

Department of Computer Science & Engineering
The Chinese University of Hong Kong
Shatin, New Territories, Hong Kong

Abstract

We present a novel face feature extraction approach using localized Principal Component Analysis (PCA) learning in face recognition tasks. The localized PCA approach produces a set of fine-tuned feature specific masks from a constrained subset of the input distribution. This method is a guided-learning based on a set of pre-defined feature points over a short training sequence. The result is the set of eigenfeatures specifically tailored for face recognition. The procedure and result of our feature extraction approach and face recognition are illustrated and discussed.

1 Introduction

Face recognition is a specialized image processing task. The recognition typically is performed in two steps: (1) feature extraction and (2) classification or matching based on the features extracted. One of the main goals of the feature extraction procedure is to seek a reduced representation of the source data that best describe each face. What is the best representation used often depends on the input distribution, in our case of face images, $\{\mathbf{x}\} = \{\mathbf{x}_l\}_{l=1}^{N}$ where each $\mathbf{x}_l$ is an $n \times n$ matrix with $x_{i,j} \in \mathcal{R}$. For natural images which include face images, there is much redundancy which leaves room for reduction in the representation [1, 3].

Several feature extraction approaches for face recognition are: model based, e.g., deformable templates [14, 15], wavelet based [2, 6, 7], unsupervised PCA based [8, 12, 11, 5], and the global PCA approach [4].

We propose a method based on a guided localized PCA approach that generates a set of eigenfeatures useful for representing faces. Section 2 will give a short background review of word done by other researchers. We will present the localized PCA approach in Section 3. We demonstrate the result from experimenting with the localized PCA methods in Section 4 and summarize in Section 5.

2 Previous Work

On face recognition, [13, 12] used the eigenface approach by finding the Principal Components (PC) of the training data. Once the linear PCA procedure is performed and the result is obtained, the system projects each training face $\mathbf{x}_i$ unto the face space spanned by the eigenvectors, $\{\vec{e_i}\}$ found from the covariance matrix. This is achieved through the vectorization of $\mathbf{x}_i$ by stacking each column vector of $\mathbf{x}_i$ on top of each other serially to form a vector $\vec{x}_i$.

This results in a representation of each face, $\vec{x}_i$ as $\mathbf{\Omega}_i = [w_1, w_2, \cdots, w_k]^t$ where $w_j = \vec{e}_j(\vec{x}_j - E(\{\vec{x}\})$ and k is the number of principal components one selects to represent the face images. The matching of a new image $\mathbf{x}_{new}$ is then performed by finding the image, $\mathbf{x}_l$, such that $||\mathbf{\Omega}_{new} - \mathbf{\Omega}_l||^2$ is the minimum.

To summarize, what is kept for the representation of each face is then the PCA learned code-book, i.e., the eigencomponents and the associated $\mathbf{\Omega}$.

O'Toole [10, 9] demonstrated an autoassociative memory based on the linear PCA to store and retrieve faces. The autoassociative memory matrix, $\mathbf{A}$, is defined as $\mathbf{A} = \sum_i \vec{f_i}\vec{f_i^t}$, such that $\vec{f_i^t}\vec{f_i} = 1$. Here the face images are stored in $\mathbf{f}_i$s. Since $\mathbf{A}$ is a symmetrical matrix, it can be decomposed by spectral analysis as $\mathbf{A} = \sum_i \lambda_i \vec{e_i}\vec{e_i^t}$ where λ_i is the ith eigenvalue, and $\vec{e_i}$ is the ith eigenvector. The recall then can be written as $\vec{f'}_i = \mathbf{A}\vec{f_i} = \lambda_1(\vec{f_1} \cdot \vec{e}_1) + \cdots + \lambda_n(\vec{f_n} \cdot \vec{e}_n)$.

However, these procedures, taking the whole face image into account, are sensitive to changes in face size, location, illumination, and background. In particular, the size and location of each face image must remain similar in order for the method to work properly. Furthermore, large illumination variation skews the PCA learning procedure which may leads to erroneous results. Lastly, using the whole image for training is indiscriminate respect to its recognition objective since the image also includes non-essential background information other than the face information. Hence, a more precise and localized method is preferred.

3 Localized PCA for Feature Extraction

The localized PCA learning approach further confines the input samples by only training on image patches around salient points on the face image. Whereas [13, 10] used a linear PCA learning to obtain a complete representation of the full image including the background, this approach focuses on specific feature points and their neighboring region. In other words, the localized PCA replaces the eigenfaces by eigenfeatures for better details.

This localized version of PCA typically renders smaller reconstruction error since it pays special attention to the local structures of the face instead of focusing on the overall face image.

One may localize on different modality of the input depending on the domain of application. For example, the sample training set may include facial salient features, e.g., mouth, eye, nose, etc. Or it can be subdivided further to a specific gender's spatial features for better classification during recognition.

The localized PCA learning is a form of directed and guided training. During the training phase, one is assumed to know *a priori* the set of facial salient points, $\{\mathbf{f}\}$, and only feeds to the network the surrounding image patch for learning. In reality, this type of learning forms a specific feature detector according to the localized subset of input distribution. In other words, if we know ahead that we want to have as face features such as the eyes, nose, and mouth then the learning statistics is constrained by the statistical distribution of these areas to produce principal components for these feature points.

Furthermore, the convergence of this localized approach is much quicker than the global approach [4] since this is a one pass learning procedure, i.e., the trial number is the same as the input training set.

3.1 Procedure for Localized PCA Learning

Input: Given $\{\mathbf{x}_i\}_{i=1}^N$ of N training face images and $\{\mathbf{f}'\} = \{\mathbf{f}'_{ij} = (x'_{ij}, y'_{ij})\}_{i=1,j=1}^{N,Q}$ of feature points. There are a total of $N \times Q$ feature points. If there are k PCA masks used then the whole face database can be represented by $N \times Q \times k$ values.

Step 1: Find out about the average matrix.

Figure 1. Some sample input training images (top row) and their corresponding test images (bottom row).

(a) Select a patch $\mathbf{x}'_i$ from $\mathbf{x}_i$ centered upon $\{\mathbf{f}'_{ij}\}$ where $\mathbf{x}'_i$ is a $p \times p$ array with $p < n$.

(b) Find $\bar{\mathbf{x}}' = \frac{1}{N}\sum_{i=1}^N \mathbf{x}'_i$ over all N sample images.

Step 2: Calculate the covariance matrix $\mathbf{C}$ using all N local image patches.

(a) Calculate $\mathbf{C} = \frac{1}{N}\sum_{i=1}^N (\mathbf{x}'_i - \bar{\mathbf{x}}')^t(\mathbf{x}'_i - \bar{\mathbf{x}}')$.

Step 3: Find the eigenvectors and the associated eigenvalues from the covariance matrix $\mathbf{C}$ to obtain p^2 eigencomponents as

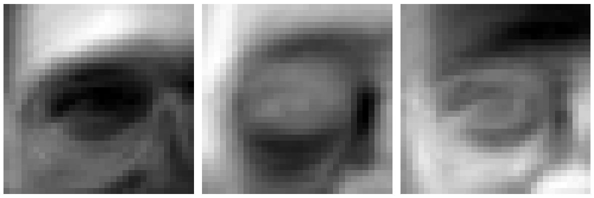
$$\mathbf{C}\vec{e}_i = \lambda_i \vec{e}_i$$

where λ_i is the eigenvalue and $\vec{e}_i$ is its associated eigenvector.

The first three principal components calculated from the localized PCA method is shown in Fig. 2 to Fig. 5. Here, $Q = 4$ which corresponds to the center-of-the-left-eye, center-of-the-right-eye, tip-of-the-nose, and center-of-the-mouth feature points. Notice that the first principal component found (subplot (a)) of each figure looks extremely similar to the ideal features we desire. Subsequent eigenvectors are found to be some variations on the same theme. These filters serve to extract other useful information from the neighborhood of the feature point.

4 Experiment and Results

The face image database contains a total of 226 face images of 113 individuals. Each person has two images, one for the training procedure and one for the testing procedure. The image database contains monochrome face images of both males and females and from different ethnic groups. All images in the database are 128×128 with 256 gray levels. A sample of few face images are shown in Fig. 1.



(a) (b) (c)

Figure 2. Lefteye masks 32 $\times$ 32 obtained from PCA learning with only lefteye images.

(a) (b) (c)

Figure 3. Righteye masks 32 $\times$ 32 obtained from PCA learning with only righteye images.

Initially, no particular care was given to control the illumination and placement of the face in the image although the width of the head is at least $1/3$ and no larger than $3/4$ of the full image. The face is roughly located near the center of the image array. A summary of parameters and their values is listed in Table 1.

4.1 Image Pre-processing

From our experiments, we found that image preprocessing of the training samples is a very crucial step in removing unwanted bias which exists in the input distribution. Therefore, a simple normalization step will make the system to obtain more useful PCA learned masks and also to make the system more robust. We used a simple dynamic range normalization step to ensure that the input distribution is not biased by a few input samples that have large intensity variations in their dynamic range.

A linear normalization of the dynamic range is calculated as, $A' = 255\frac{(A-\min(A))}{(\max(A)-\min(A))}$ where A is an arbitrary matrix with the final elements of $A' \in [0, 255]$, $\min(A)$ and $\max(A)$ are the minimum and maximum value of the matrix A respectively.

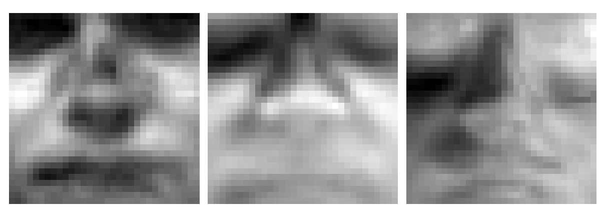
From an information-theoretic point of view, the individual face image's entropy is not changed for each of the image but the global entropy for the whole input distribution over the whole training set is modified by the normalization procedure.

The justification of this step is biologically motivated since the iris acts like an illumination normalization mechanism by automatically adjusting the contrast (dynamic) level of the input sensory information. Hence, this normalization is plausible and should be performed.

4.2 Matching Results

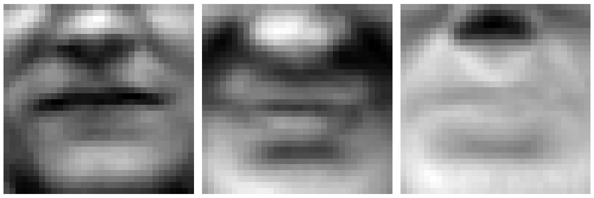
We use a simple matching strategies in our study to demonstrate the feature extraction from PCA learning. More elaborate matching schemes will yield better results, but it is not the main objective of this paper.

In the localized PCA method, after convolving the original image with the localized PCA masks, we obtain the coefficients for center-of-the-left-eye, center-of-the-right-eye, tip-of-the-nose, and center-of-the-mouth points. The result is summarized in Table 2. The table presents the ranking of the match, i.e., a perfect match of a new face to a model face of the same person would be 1, the worst would be 113 in our case. This is obtained by sorting the error values calculated from the matching and finding the correct position in the sorted list for the target image. The second column gives the frequency of the rank in a particular interval. The third column gives the percentage of the frequency in that ranking interval to the total number of test images. The last column presents the accumulated percentage from the previous column. This column tells us that, for example, the percentage of rank in the 1-9 range is 53.98% while the accumulated percentage for 1-19 is 73.45%.



(a) (b) (c)

Figure 4. Nose masks 32 $\times$ 32 obtained from PCA learning with only nose images.



(a) (b) (c)

Figure 5. Mouth masks 32 $\times$ 32 obtained from PCA learning with only nose images.

5 Discussion

The performance of the matching can be improved by increasing the feature set, $\{\mathbf{f}\}$ or by increasing the number of PCA learned masks used. Furthermore, since the computational resource required is relatively little after the learning is done this can be a very quick way to get an approximated result to see whether a finer matching is required.

Although we have chosen the batch algorithm to implement our PCA learning, an on-line version of the PCA algorithm based on iterative method can also be used to achieve the same result. The iterative algorithm can be used when the face images are not known *a priori*. The advantage to use an on-line PCA algorithm is that a precise knowledge of the input set is not required making the system extensible to new face images. On the other hand, previous learned coefficients will need periodic updating since the PCA learned masks are changing with each update. Hence, a proper procedure and schedule to re-calculate the representation of previously generated face images still can be troublesome.

Lastly, these approaches can be adopted to other domain specific object recognition task quite easily without much modifications. This is because the system "learns" what is important and what is not from the training sample. Therefore, this method is potentially very useful as a generalized vision framework.

6 Conclusion

We have demonstrated a localized PCA method for extracting features for matching face images. This method is a guided-learning based on a set of pre-defined feature points over a short training sequence. The result is the eigenfeatures specifically tailored for extracting facial features. From the result, we have shown that the localized PCA method used is quite efficient in extracting representational features from face images for matching. These extracted coefficients can be used in an initial matching stage to obtain a plausible subset of possible targets for a more detailed matching.

Table 1. A summary of parameters and their associated values.

Parameters	Value
N	113
Q	5
k	10
$\mathbf{x}$	128×128 matrix
n	128
p	32

Acknowledgement

We thank the USC Face Recognition group for providing some test images. This research is supported in part by the Earmark Grant # CUHK 485/95E, Hong Kong Research Grant Council.

References

[1] H.B. Barlow. Unsupervised learning. *Neural Computation*, 1(3):295–311, 1989.

[2] J. Buhmann, J. Lange, C. von der Malsburg, J.C. Vorbruggen, and R.P. Wurtz. Object recognition with Gabor functions in the dynamic link architecture. In B. Kosko, editor, *Neural Networks for Signal Processing*, pages 121–159. Prentice Hall, Englewood Cliffs, NJ, 1992.

[3] D. J. Field. What is the goal of sensory coding? *Neural Computation*, 6(4):559–601, Jul 1994.

[4] I. King and L. Xu. Using global PCA generated receptive fields for face recognition. In *Proceedings to the World Congress on Neural Networks*, volume II, pages 542–545, Washington D. C. USA, July 17-21 1995. International Neural Network Society, Lawrence Erlbaum Associates, Inc.

[5] M. Kirby and L. Sirovich. Application of the Karhunen-Loeve procedure for the characterization of human faces. *IEEE Trans. PAMI*, 12(1):103–108, 1990.

[6] M. Lades, J. C. Vorbruggen, J. Buhmann, J. Lange, and C. von der Malsburg. Distortion invariant object recognition in the dynamic link architecture.

Table 2. A Summary of matching results for the left eye, right eye, nose, and mouth localized PCA method.

	Rank	Frequency	Percent. Rank (%)	Accu. (%)
Left Eye	1-9	61	53.98	53.98
	10-19	22	19.46	73.45
	20-29	7	6.19	79.64
	30-39	4	3.53	83.18
	40-49	10	8.84	92.03
	50-69	3	2.65	94.69
	70-89	3	2.65	97.34
	90-113	3	2.65	100
Right Eye	1-9	67	59.29	59.29
	10-19	16	14.15	73.45
	20-29	6	5.30	78.76
	30-39	7	6.19	84.95
	40-49	4	3.53	88.49
	50-69	7	6.19	94.69
	70-89	5	4.42	99.11
	90-113	1	0.88	100
Nose	1-9	54	47.78	47.78
	10-19	15	13.27	61.06
	20-29	6	5.30	66.37
	30-39	6	5.30	71.68
	40-49	5	4.42	76.10
	50-69	14	12.38	88.49
	70-89	9	7.96	96.46
	90-113	4	3.53	100
Mouth	1-9	69	61.06	61.06
	10-19	14	12.38	73.45
	20-29	6	5.30	78.76
	30-39	4	3.53	82.30
	40-49	6	5.30	87.61
	50-69	7	6.19	93.80
	70-89	6	5.30	99.11
	90-113	1	0.88	100

IEEE Transactions on Computers, 42(3):300–311, Mar 1993.

[7] B. S. Manjunath, Chandra Shekhar, R. Chellappa, and C. von der Malsburg. A robust method for detecting image features with application to face recognition and motion correspondence. In *Proceedings. 11th IAPR International Conference on Pattern Recognition*, volume Conference B: Pattern Recognition Methodology and Systems, pages Vol. 2:208–212, Los Alamitos, CA. USA, Aug-Sep 1992. IEEE Computer Society Press.

[8] Erkki Oja and Jouko Lampinen. Unsupervised learning for feature extraction. In Rober J. Marks II Jacek M. Zurada and Charles J. Robinson, editors, *Computational Intelligence Imitating Life*, pages 13–22. IEEE Press, Piscataway, NJ, 1994.

[9] A. J. O'Toole, H. Abdi, and K. A. Deffenbacher. Low-dimensional representation of faces in higher dimensions of the face space. *Journal of the Optical Society of America. A, Optics and Image Science*, 10:405–11, Mar 1993.

[10] Alice J. O'Toole and Jamie L. Thompson. An X Windows tool for synthesizing face images from eigenvectors. *Behavior Research Methods, Instruments, and Computers*, 25(1):41–47, 1993.

[11] L. Sirovich and M. Kirby. Low-dimensional procedure for the characterization of human faces. *J. Opt. Soc. Am. A*, 4(3):519–524, Mar 1987.

[12] Matthew A. Turk and Alex Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, Win 1991.

[13] Matthew A. Turk and Alex P. Pentland. Face recognition using eigenfaces. In *Proceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 586–591, Los Alamitos, CA USA, Jun 1991. IEEE Computer Society Press.

[14] A. L. Yuille. Deformable templates for face recognition. *J. Cogn. Neurosci.*, 3:59–70, 1991.

[15] Alan L. Yuille, Peter W. Hallinan, and David S. Cohen. Feature extraction from faces using deformable templates. *International Journal of Computer Vision*, 8(2):99–111, 1992.