Using Global PCA Generated Receptive Fields for Face Recognition*

Irwin King
king@cs.cuhk.hk

Lei Xu
lxu@cs.cuhk.hk

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

Abstract

We apply the global Principal Component Analysis (PCA) learning for face recognition tasks. The global unsupervised PCA learning generates a set of plausible visual receptive fields that are ideal for image decomposition during the feature extraction process for recognition. The procedure and results of our approach are illustrated and discussed.

1 Introduction

General machine recognition typically is performed in two steps: (1) feature extraction and (2) classification or matching based on the features extracted. The main goal of the feature extraction procedure is to seek a reduced representation of the source data. 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 in natural images which leaves room for reduction in the representation [2, 4]. Principal Component Analysis (PCA) is a way to reduce the dimensionality of the input by finding a small set of basis vectors that accommodates the largest variance in the input statistics. For this reason, we choose PCA as a way for feature detection for face recognition tasks.

2 Global PCA Learning of Receptive Masks

Researchers have shown that PCA learning forms receptive fields from random inputs [9, 1, 6]. Here, we postulate that these receptive filters are better than the Gabor wavelet approach since these receptive filters are learned from the input statistics instead of being tuned manually. A recent review of PCA learning methods is shown in [11].

Hence, we generate these receptive fields as generalized feature detectors in the feature extraction process for face recognition. The output feature map is calculated by convolving the input image $\mathbf{x}$ with a generated receptive field, $\vec{e}_i$, i.e.,

$$I_i^o(\mathbf{x}) = \mathbf{x} * \vec{e}_i \tag{1}$$

where $\vec{e}_i$ is the i-th principal component corresponding to the eigenvector found in the i-th largest eigencomponent calculated from the PCA procedure. Each principal component correspons to a receptive field which is sensitive to a particular intensity distribution (spatial feature) from the input statistics.

To form these receptive filters, one requires a random sampling from a set of images over a long trial sequence. Here is the procedure.

*This work was supported in parts by the Hong Kong Research Grant Council, NO: 220500720

Input: Given $\{\mathbf{x}_i\}_{i=1}^N$ of N training face images over K learning trials. In our experiment, we have $N = 113$ and $\mathbf{x} = 128 \times 128$.

Do: Find the average matrix.

(a) Select a receptive field patch $\mathbf{x}'_l$ randomly from the set of input images $\{\mathbf{x}_i\}_{i=1}^N$ where $\mathbf{x}'$ is a $p \times p$ array with $p < n$. Here, $p = 32$ and $K = 10,000$.

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

Do: Find the covariance matrix after all K trials.

(a) Calculate $\mathbf{C} = \frac{1}{K}\sum_{i=1}^K (\mathbf{x}'_i - \bar{\mathbf{x}}')^t(\mathbf{x}'_i - \bar{\mathbf{x}}')$. More precisely, covariance matrix is calculated by $\vec{x}_i^t \vec{x}_i$ where $\vec{x}_i$ is the vectorization of the matrix resulted from $(\mathbf{x}'_i - \bar{\mathbf{x}}')$.

Do: Find the eigenvectors and its associated eigenvalues from

$$\mathbf{C}\vec{e}_i = \lambda_i \vec{e}_i \tag{2}$$

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

It has been shown that linear neural networks can perform the above procedure to extract the principal components [8, 7, 9]. Without loss of generality, here we just use the above procedure to generate the receptive filters for recognition.

A sample of the receptive field generated from the global PCA approach is illustrated in Fig. 1(b). The top 4 rows illustrate the 32 largest eigenvectors found from $\{\mathbf{x}\}$. We have arranged them from left to right and top to bottom according to their non-decreasing eigenvalues. Since the calculation of the principal components is off by ± 1, a full set of receptive field masks can be obtained by a sign change. This is illustrated in the bottom 4 rows of the figure which are the contrast reversed sets of filters mirroring the top 4 rows of filters.

Once the receptive masks are generated, these filters can then be convolved with the input image to extract local features as shown in Eq.(1). The advantage here is that fewer filters can be used than in the full Gabor wavelet approach. This is because the associated eigenvalues of these eigenvectors drop off quickly after the first few principal components indicating that only the first few principal ones are required for a good reconstruction. See [10] for the characterization and reconstruction of human faces with few PCA components.

The feature extraction is done by selecting the coefficient from the set of salient points in each of the convolved images, $I_i^o(\mathbf{x})$. An instance of this feature extraction process is demonstrated in Fig. 2.

3 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 face images from 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). 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.

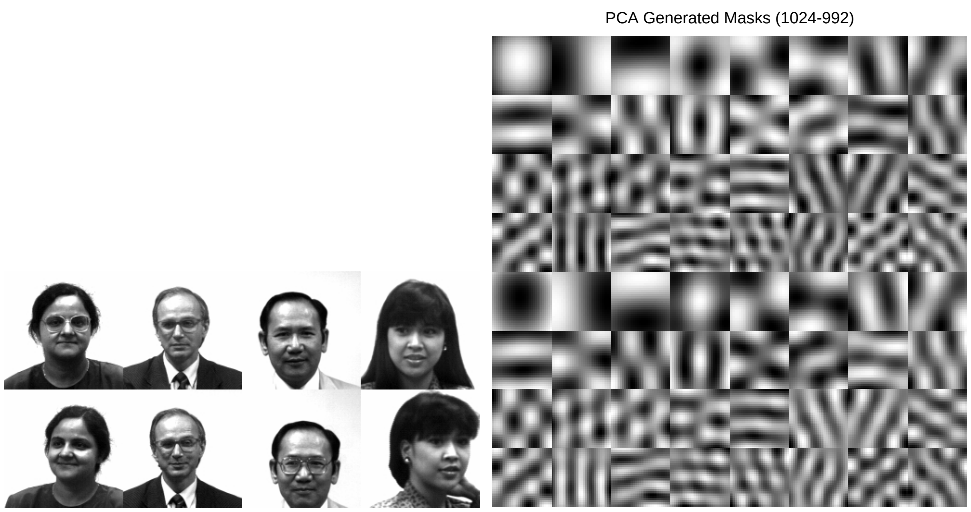
3.1 Matching Results

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.

For comparison, a general error measurement is defined as

$$\mathrm{err}(\vec{x}, \vec{x}') = \left(\sum_i (x_i - x_i')^2 \right)^{\frac{1}{2}} \tag{3}$$

After the convolution operation with the principal eigenvectors we keep only a few points for representation used in matching, e.g., left eye, right eye, nose tip, left mouth end, and right mouth. The result is summarized in Table 1. 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.



(a) (b)

Figure 1: (a) Some sample input training images and their corresponding test images. (b) The set of globally learned PCA masks generated from training of 32×32 window of 128×128 size face images.

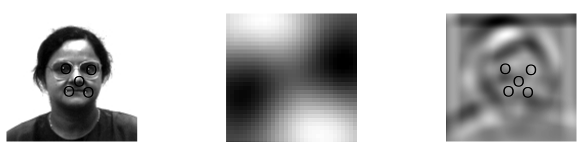


Figure 2: The feature extraction process. The left picture is an input face image. The middle subfigure is the 5th largest principal component found as shown in 1(b). We convolve the input image with the mask to obtain the right subfigure with marks around the feature points. These are the points we keep for modeling the face.

4 Discussions

The performance of the matching can be improved by increasing the feature set or by increasing the number of feature detectors generated from PCA learning. Furthermore, since the computational

Table 1: Matching results using the global masks.

Rank	Freq.	Percent. Rank (%)	Accu. (%)
1-9	50	44.2478	44.2478
10-19	13	11.5044	55.7522
20-29	8	7.0796	62.8319
30-39	7	6.1947	69.0265
40-49	10	8.8496	77.8761
50-69	11	9.7345	87.6106
70-89	9	7.9646	95.5752
90-113	5	4.4248	100

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 *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 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.

In the future, we plan to focus on implementing the elastic matching for the global approach and compare results with that obtained by [5, 3]. We also plan to extend the linear PCA learning to the non-linear case [11]. Preliminary results show non-linear PCA learning will give even better performance than the linear PCA.

5 Conclusion

We have demonstrated a global PCA-based method for feature extraction for face recognition. The global PCA method is based on the formation of receptive fields using a set of random samples from the input image set over a long training sequence. These global PCA masks are more biologically plausible and statistically more efficient to represent input samples. From the result, we have shown that this method used is quite efficient in extracting representational features from face images for matching.

References

[1] Joseph J. Atick and A. Norman Redlich. Convergent algorithm for sensory receptive field development. *Neural Computation*, 5:45–60, 1993.

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

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

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

[5] M. Lades, J. C. Vorbruggen, J. Buhmann, J. Lange, and C. von der Malsburg. Distortion invariant object recognition in the dynamic link architecture. *IEEE Transactions on Computers*, 42(3):300–311, Mar 1993.

[6] Ralph Linsker. From basic network principles to neural architecture. *Proc. Natl. Acad. Sci. USA*, 83(7508-7512, 8390-8394, 8779-8783), October-November 1986.

[7] E. Oja. Neural networks, principal components, and subspace. *International Journal of Neural Systems*, 1:61–68, 1989.

[8] Erkki Oja. A simplified neuron model as a principal component analyzer. *J. Math. Biology*, 15:267–273, 1982.

[9] T. D. Sanger. Analysis of the two-dimensional receptive fields learned by the generalized hebbian algorithm in response to random input. *Biological Cybernetics*, 63:221–228, 1990.

[10] 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.

[11] Lei Xu. Beyond PCA learnings: From linear to nonlinear and from global representation to local representation. In Myung-Won Kim and Soo-Young Lee, editors, *Proceedings to the International Conference on Neural Information Processing*, pages Vol. II:943–949, Seoul, Korea, Oct 1994.