A VECTOR BASED APPROXIMATION OF KLT AND ITS APPLICATION TO FACE RECOGNITION

Nicolas Tsapatsoulis, Vassilios Alexopoulos and Stefanos Kollias
Dept. of Electrical and Computer Engineering,
National Technical University of Athens
Heroon Polytechniou 9, 157 73 Zographou, Greece
Tel: +30 1 7722491; fax: +30 1 7722492
e-mail: ntsap@image.ntua.gr

ABSTRACT

A face recognition scheme is proposed in this paper. The method utilizes a vector based approximation of KLT (VKLT) which eliminates the large memory demands and the singularity of the covariance matrix matters that are the main drawbacks of the "eigenface method" (a face recognition scheme based on KL transform). The reconstruction error of VKLT approaches the one of KLT preserving also its data dependence, which is important for discriminating face images from non-face images. Moreover, the greater advantage of VKLT over KLT is that of keeping intra-class variations small and consequently increasing the robustness of the face recognition system.

1 INTRODUCTION

In recent years numerous algorithms have been proposed for face recognition [1] and much progress has been made toward this direction. Most of these algorithms achieve high recognizing rates only under very small variations in lighting, facial expressions and perspective angle or pose [3, 4]. The inefficiency of the presented algorithms under extreme variations of the above factors is quite logical. Recent investigations have shown that the proposed face representation schemes derive greater variability in a given face under changes in illumination, perspective angle and expression, than different faces when these three factors are held constant. In other words intra-class variance is larger than the inter-class one [2].

One of the most important face recognition schemes, reporting the highest recognizing rate, is based upon the Karhunen Loeve Transform (KLT). This method, called "Eigenface method", uses Principal Component Analysis for dimensionality reduction and yields projection directions that maximize the inter-class scattering [4, 6]. However, unwanted variations in illumination, pose and facial expressions are retained since intra-class variations are also increased. Furthermore the eigenface method demands a large amount of memory to estimate the covariance matrix which is in almost all cases singular due to the limited number of samples (faces) used for its estimation.

In our method a vector based approximation of KLT is proposed (VKLT) which eliminates the large memory demands and the singularity of the covariance matrix. The reconstruction error of the VKLT approaches the one of KLT preserving also its data dependence which is important for discriminating face images from non-face images. Moreover, the greater advantage of VKLT over KLT is that of keeping intra-class variations small and consequently increasing the robustness of the face recognition system.

2 FACE RECOGNITION SCHEMES

In order to achieve higher recognizing rates face recognition systems should first convert face images to an appropriate format ("head format") in which the face is in a specific scaling and orientation, and some protuberant facial features (eyes, mouth, nose) are aligned. A generic block diagram of a face recognition scheme is given in figure 1.

The feature extraction part is the heart of the system. Recognition without feature extraction leads to correlation in the image space which is inefficient and computational expensive since any image to be tested should be compared to any image of the database.

The eigenface method uses KLT projection for feature extraction [4, 6]. In particular, consider a set of N sample images $\{I_1, I_2, ..., I_N\}$ where $I_i \in R^{n \times m}$. By lexicographic ordering of the pixel elements of each image I_i we can form a set of vectors $\{v_1, v_2, ..., v_N\}$ where $v_i \in R^L$ and $L = n \cdot m$. The KLT basis functions are obtained by solving the eigenvalue problem:

$$\Lambda = \Phi^T \Sigma \Phi \tag{1}$$

where Σ is the covariance matrix, Φ is the eigenvector matrix of Σ, and Λ is the corresponding diagonal matrix of eigenvalues.

To form a feature vector y_i for each vector v_i a partial KLT is performed:

$$y_i = {\Phi_k}^T \cdot \hat{v}_i \tag{2}$$

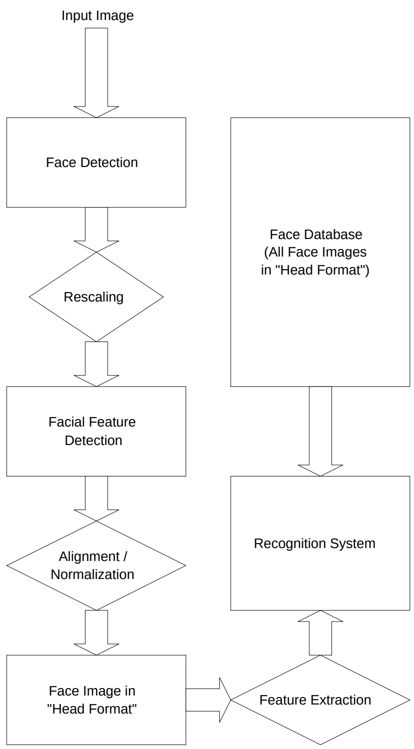

Figure 1: Architecture of a generic face recognition system.

where $\hat{v}_i = v_i - \mu$, $\mu = \sum_{j=1}^{N} v_j$ and Φ_k is the truncated eigenvector matrix in which only the eigenvectors corresponding to the k largest eigenvalues are retained.
It is observed that the dimension of the covariance matrix is extremely large i.e., $\Sigma \in R^{nm \times nm}$ and therefore its computation is impractical. For example considering images of spatial resolution 90x100, $\Sigma \in R^{9000 \times 9000}$. Furthermore, in order to avoid the singularity of Σ the number of sample images should be $N > m^2 \cdot n^2$.

3 VECTOR BASED KLT

In our approximation of KLT we consider each image in the training set $\{I_1, I_2, ..., I_N\}$ to consist of row and column vectors. We calculate two vector covariance matrices, the covariance matrix of rows:

$$R_0 = \sum_{j=1}^{N} (I_j - M) \cdot (I_j - M)^T \tag{3}$$

and columns:

$$C_0 = \sum_{j=1}^{N} (I_j - M)^T \cdot (I_j - M) \tag{4}$$

where $M = \sum_{j=1}^{N} I_j$ is the mean face image. It should be mentioned that $R_0 \in R^{n \times n}$, $C_0 \in R^{m \times m}$ and therefore their dimensions are much smaller than that of Σ.
Solving the equations:

$$\Lambda_1 = {\Phi_1}^T R_0 \Phi_1 \tag{5}$$

$$\Lambda_2 = {\Phi_2}^T C_0 \Phi_2 \tag{6}$$

we get the matrices Φ_1 and Φ_2 which define coordinate transforms which decorrelate the rows and columns of the training images respectively.
Retaining the q eigenvectors of Φ_1 and Φ_2 which correspond to the largest eigenvalues of matrices Λ_1 and Λ_2 we obtain a feature matrix Y_i for each image i.e.,

$$Y_i = {\Phi_{1q}}^T \hat{I}_i \Phi_{2q} \tag{7}$$

where $\hat{I}_i = I_i - M$.
In the VKLT approach the KLT is approximated in the sense of decorrelating the image data in two steps: first decorrelating the rows and then the columns.

4 EVALUATION PROCEDURE

Several tests were performed so as to evaluate the efficiency of VKLT. Let us consider a given set of training images $S = \{I_1, I_2, ..., I_N\}$ whose corresponding set of lexicographic ordered vectors is $V = \{v_1, v_2, ..., v_N\}$, and the matrices Φ_k, Φ_{1q} and Φ_{2l} which are the truncated versions of Φ, Φ_1 and Φ_2 (we retain only k,q and l eigenvectors respectively).

4.1 Mean Reconstruction Error in Training Set

In this test we compare the mean reconstruction error over all training images, for KLT and VKLT, and for several different values of k,q and l. To have the same length of image representation (number of element features retained) and therefore make the comparison fair, we let $k = l \cdot q$.
We define the following reconstruction errors:

$$e_{KLT}(k, j) = (\hat{v}_j - \check{v}_j)^T \cdot (\hat{v}_j - \check{v}_j) \tag{8}$$

where $\check{v}_j = \Phi_k \cdot y_j$, $y_j = {\Phi_k}^T \cdot \hat{v}_j$ and

$$e_{VKLT}(q, l, j) = vec(\hat{I}_j - \check{I}_j)^T \cdot vec(\hat{I}_j - \check{I}_j) \tag{9}$$

where $vec(I)$ is a vector corresponding to lexicographic ordering of image I, $\check{I}_j = \Phi_{1q} Y_j {\Phi_{2l}}^T$, and $Y_j = {\Phi_{1q}}^T \hat{I}_j \Phi_{2l}$.
Using the equations 8 and 9 we define the mean reconstruction error of KLT and VKLT respectively as:

$$me_{KLT}(k) = \frac{1}{N} \sum_{j=1}^{N} e_{KLT}(k, j) \tag{10}$$

$$me_{VKLT}(q, l) = \frac{1}{N} \sum_{j=1}^{N} e_{VKLT}(q, l, j) \tag{11}$$

4.2 Intra-class Scattering

In face recognition schemes variations in illumination, translation and facial expressions should not affect the recognition task. In other words intra-class variations should be minimized in the representation space. In the following tests KLT and VKLT are compared based on the distance between the representations of an image and its translated, noisy and differently illuminated versions. Given an image I_i and the respective vector v_i, we define its random translation T_i and its differently illuminated version L_i. We consider also the corresponding vectors t_i and l_i. The distance in KLT representations

$$te_{KLT}(k,i) = \| {\Phi_k}^T \cdot \hat{v}_i - {\Phi_k}^T \cdot t_i \| \tag{12}$$

$$le_{KLT}(k,i) = \| {\Phi_k}^T \cdot \hat{v}_i - {\Phi_k}^T \cdot l_i \| \tag{13}$$

between and or should be small. The same holds for the VKLT representations of I_i and T_i or L_i:

$$te_{VKLT}(q,l,i) = \| vec({\Phi_{1q}}^T \hat{I}_i \Phi_{2l}) - vec({\Phi_{1q}}^T T_i \Phi_{2l}) \| \tag{14}$$

$$le_{VKLT}(q,l,i) = \| vec({\Phi_{1q}}^T \hat{I}_i \Phi_{2l}) - vec({\Phi_{1q}}^T L_i \Phi_{2l}) \| \tag{15}$$

The mean distances between the above representations over the whole database can be calculated using equations 12, 13, 14, 15 and definitions equivalent to that of 10 and 11.

4.3 Recognition Rate

The most important test in face recognition schemes is the recognition task in which a face contained in the face database but in different scaling, pose, illumination or facial expression is presented to the system. The recognizing criterion is usually the nearest neighbor rule and comparisons are made in the feature space. It should be mentioned that the system should discard non-face images and this is why data dependent transforms are important in face recognition tasks.
In our simulations the recognizing rates were calculated for several representations in KLT and VKLT domains using the nearest neighbor rule and keeping always $k = l \cdot q$.

5 EXPERIMENTAL RESULTS

The experimental results for all tests, were obtained using the ORL face database. Comparisons were accomplished for both methods (KLT and VKLT) using face images in "head format". Due to memory problems arising from the covariance matrix of KLT we restricted the spatial resolution of face images to 64x72.
The performance of the recognition task is shown in Table 1 for both KLT and VKLT. The outperformance of VKLT is depicted clearly in all cases (different retained coefficients). Moreover, as it is illustrated in Figures 3,4 and 5, VKLT outperforms KLT, in all tests except that of mean reconstruction error in which KLT is by definition optimum. Particularly, VKLT has smaller mean translation and mean illumination error and therefore is more robust than KLT in intra-class variations.

Retained Coefficients	16	25	36	49
Performance %				
KLT	*54*	*63*	*75*	*82*
VKLT	*62*	*69*	*81*	*85*

Table 1: Recognition rate for KLT and VKLT.

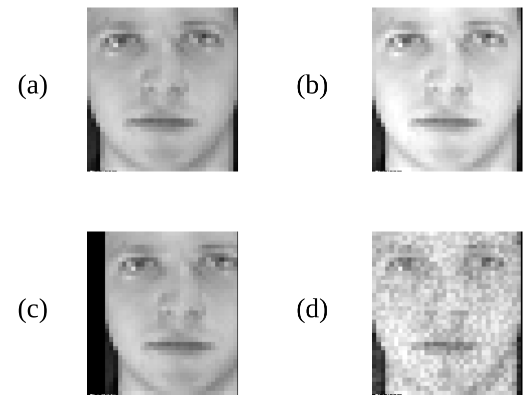

Figure 2: (a) Original face image in "head format" (b) Differently iluminated (c) Translated (d) Noisy.

6 CONCLUSION

An innovative face recognition scheme is proposed in this paper. The method utilizes a vector based approximation of KLT considering that the face images consist of row and column vectors. The VKLT eliminates the large memory demands and the singularity of the covariance matrix which are main disadvantages of the KL transform. On the other hand the presented method performs better than the "eigenface" one, under variations in illumination, translation and facial expressions. Moreover the proposed scheme keeps the intra-class variations small and therefore increases the robustness of the whole system.

References

[1] P. Chellapa, C. Wilson and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," *Proc. IEEE*, vol. 83, no. 5, pp. 705-740, 1995.

[2] J. Daugman, "Face and Gesture Recognition: Overview," *IEEE Trans. on PAMI*, vol. 19, no 7, pp. 675-676, 1997.

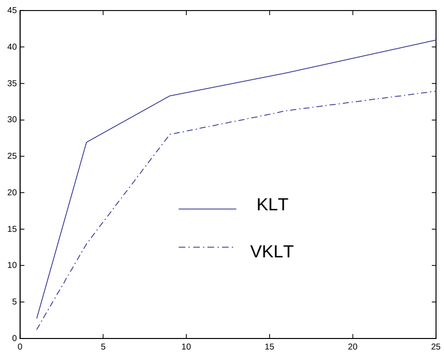

Figure 3: Mean distances in KLT and VKLT domains between the original images and their random translated versions, vs. retained coefficients.

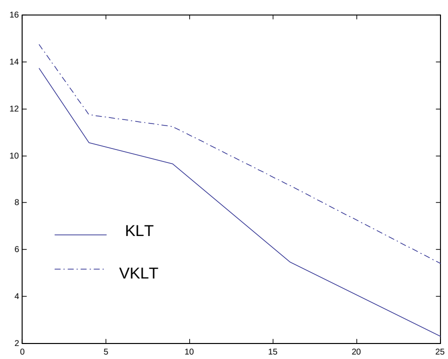

Figure 4: Mean reconstruction error in KLT and VKLT domains vs. retained coefficients.

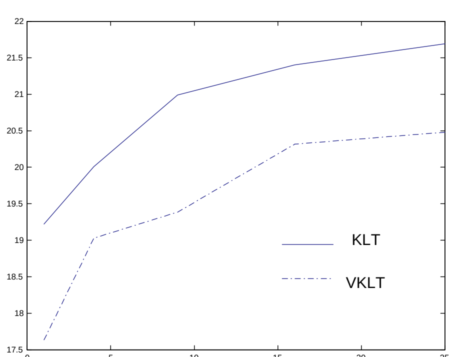

Figure 5: Mean distances in KLT and VKLT domains between the original images and their differently illuminated versions, vs. retained coefficients.

[3] A. Doulamis, N. Tsapatsoulis, N. Doulamis and S. Kollias, "Innovative Techniques for the Recognition of Faces Based on Multiresolution Analysis and Morphological Filtering," *Proc. of IWISP*, Manchester, November 1996.

[4] B. Moghaddam and A. Pentland, "Probabilistic Visual Learning for Object Representation," *IEEE Trans. on PAMI*, vol. 19, no 7, pp. 696-710, 1997.

[5] N. Tsapatsoulis, K. Karpouzis, G. Votsis and S. Kollias, "Analysis by Synthesis of Facial Images Based on Frontal and Profile Views," *Proc. of IWSNHC3DI Conference*, Rhodes, September 1997.

[6] M. Turk and A. Pentland, "Eigenfaces for Recognition," *Journ. Cognitive Neuroscience*, vol. 3, no. 1, 1991.