EIGENFACE RECOGNITION USING DIFFERENT TRAINING DATA SIZES

Zhifeng Li
Department of Information Engineering
The Chinese University of Hong Kong
Shatin, N.T., Hong Kong
(+852) 26098206
zli0@ie.cuhk.edu.hk

Xiaoou Tang
Department of Information Engineering
The Chinese University of Hong Kong
Shatin, N.T., Hong Kong
(+852) 26098379
xtang@ie.cuhk.edu.hk

ABSTRACT
This paper explores the relationship between eigenface recognition performance and different training data sets. Using the Multilevel Dominant Eigenvector Estimation (MDEE) method we are able to compute eigenfaces from a large number of training samples. This allows us to compare the recognition performance using different training data sizes. Experimental results show that increasing the number of people benefits the recognition performance more than increasing the number of images per person.

Keywords
Face recognition, eigenface, KLT, training data.

1. INTRODUCTION
Face Recognition is one of the most challenging computer vision research topics since faces appear differently even for the same person due to expression, pose, occlusion and many other confounding factors in real life. In recent years, researchers have proposed several face recognition methods among which the eigenface method is among the most popular ones [1].

The eigenface approach uses the Karhunen-Loeve Transform (KLT) for the representation and recognition of face [3][5][6]. Once a set of eigenvectors, also called eigenfaces, is computed from the face covariance matrix, a face image can be approximately reconstructed using a weighted combination of the eigenfaces. The weights that characterize the expansion of the given image in terms of eigenfaces constitute the feature vector. When a new test image is given, the weights are computed by projecting the image onto the eigenface vectors. The classification is then carried out by comparing the distances between the weight vectors of the test image and the images from the database.

Since the Eigenface vectors are computed directly from the training face images, it is reasonable to expect that the recognition results may be influenced by different training data sets. However, most previous researches simply choose a small number of training samples randomly for computation of the eigenfaces without much justification [3][5][6]. In this paper, we conduct a systematic experimental study on the relationship between the face recognition performance and training data sets with different number of total samples, number of samples per class, number of classes. To significantly reduce the computational complexity involved in eigenvector computation of large number training samples, we use the Multilevel Dominant Eigenvector Estimation (MDEE) method developed by Tang [4] to approximate the KLT.

2. EIGENFACE AND MULTILEVEL DOMINANT EIGENGACE ESTIMATION
The eigenface method is based on Karhunen-Loeve transform (KLT). Kirby and Sirovich first use eigenfaces to characterize faces [3]. Later, Turk and Pentland apply the approach on face recognition [5][6]. We now briefly review the basic idea of the eigenface method.

Let $x_1^{'}, x_2^{'} \ \dots \ x_m^{'}$ represent a set of n-dimension random vectors and μ is the mean vector. The procedure of computing the Karhunen-Loeve transform is described as the follows:

(1) Form the *n* by *m* sample matrix

$$A = \begin{bmatrix} x_1(1) & x_2(1) & \dots & x_m(1) \\ x_1(2) & x_2(2) & \dots & x_m(2) \\ \dots & \dots & \dots & \dots \\ x_1(n) & x_2(n) & \dots & x_m(n) \end{bmatrix}, \qquad (1)$$

where $x_i = x_i^{'} - \mu$, *n* is the length of each vector, and *m* is the number of vectors.

(2) Estimate the covariance matrix,

$$W = \frac{1}{m}\sum_{i=1}^{m} x_i x_i^T = \frac{1}{m} AA^T . \qquad (2)$$

(3) Compute the eigenvectors of the covariance matrix and select k eigenvectors $V_1\ V_2\ \dots\ V_k$ with the largest eigenvalues to form the transform matrix,

$$B = [V_1 \quad V_2 \quad \dots \quad V_k]. \qquad (3)$$

(4) For a new n-dimension vector x, we project it in the subspace spanned by the k eigenvectors,

$$y = B^T (x - \mu), \qquad (4)$$

where y is the weight vector that characterizes the projection of the vector x in the subspace supported by the k eigenvectors.

For face recognition, a 2-dimensional N by M face image is usually represented by a one-dimensional face vector with the length $n=N*M$, where n is usually a very large number. In our experiments n is of size 81*101=8181. This means the size of the covariance matrix W is 8181 by 8181. It is impractical to calculate the eigenvectors from such a large matrix W directly. However, since there are only m samples in the sample matrix A, the rank of the covariance matrix is in fact m-1 [2]. Assuming that m is in general much smaller than n, the eigenface method first computes the eigenvectors of a much smaller m by m matrix $\frac{1}{m}A^T A$, then obtains the eigenvectors of the covariance matrix $\frac{1}{m}AA^T$ by a multiplication of A with the smaller eigenvectors. However, when the number of samples m is also very large, this method encounters the same problem as the direct eigenvector computation.

To overcome the computational problem, we use the Multilevel Dominant Eigenvector Estimation (MDEE) method developed by Tang [4]. It has been shown to be a very close approximation of the standard KLT with s significant reduction of computational complexity [4].

The MDEE method first breaks the long face vector into $g = n/k$ groups of small vectors with the length k.

$$A = \begin{bmatrix} B_1 \left\{ \begin{bmatrix} x_1(1) & x_2(1) & \dots & x_m(1) \\ \dots & \dots & \dots & \dots \\ x_1(k) & x_2(k) & \dots & x_m(k) \end{bmatrix} \right\} \\ B_2 \left\{ \begin{bmatrix} x_1(k+1) & x_2(k+1) & \dots & x_m(k+1) \\ \dots & \dots & \dots & \dots \\ x_1(2k) & x_2(2k) & \dots & x_m(2k) \end{bmatrix} \right\} \\ \dots \quad \dots \quad \dots\dots \quad \dots \\ \dots \quad \dots \quad \dots\dots \quad \dots \\ B_g \left\{ \begin{bmatrix} x_1((g-1)k+1) & x_2((g-1)k+1) & \dots & x_m((g-1)k+1) \\ \dots & \dots & \dots & \dots \\ x_1(n) & x_2(n) & \dots & x_m(n) \end{bmatrix} \right\} \end{bmatrix} \qquad (5).$$

After performing KLT on each group B_i, we select the first few dominant eigenfeatures from each group and put them together to form a new feature vector. Then the final feature vector is computed by applying the KLT to this new feature vector.

MDEE can achieve considerable reduction of computing time over the standard KLT. For example, if we break a face vector of length n into g = 10 groups of small vectors and only keep the top 10% of the eigenfeatures in each group for the second-level eigenvector computation, the computational complexity is only $11(n/10)^3$. Comparing to the computational complexity of the standard KLT, we reduce the computational complexity by two orders of magnitude.

Using this method, we are no longer limited by either the size of the image or the number of training samples. Through a set of experiments we can now investigate whether using a larger number of training samples will increase the recognition accuracy.

3. EXPERIMENTS

3.1 Database

For face recognition research our lab has built a face-based video sequence database. It is divided into two sessions. The first session is composed of 172 video sequences of 172 different people. The second session is composed of 72 video sequences of 72 different people who also appeared in the first session. All video sequences are captured under the same conditions. There is a time gap of more than one month between the first and the second sessions. The duration of each video sequence is 20 seconds. The person in the video is asked to read a short paragraph of text with normal expression. For each video sequence, 50 face images are intercepted evenly during the 20 seconds long.

3.2 Preprocessing Procedure

Preprocessing is an important step in face recognition. To better compare the recognition performance we first process face images through the following steps.

1. Rotate the face image to align the vertical face orientation.
2. Scale the face image so that the distances between the two eyes are the same for all images.
3. Crop the face image to remove the hair region.

After preprocessing, each face image has a size of 81 by 101. Fig 1 shows 20 samples from two persons after preprocessing.

3.3 Selection of Training Data Sets and Testing Data Sets

For the experiments, we use images of 100 people in the first session as training data, and use images of the other 72 people as testing data. There is no overlap between the two data sets.

In order to evaluate the influence of different training data sets on the recognition accuracy, we select different subsets from the training data set for the experiments. We design two training data sets with each containing 3 subsets, as shown in Table 1. For the first training data set, we fix the number of total training samples and then change the class number and samples per class in each training subset. For the second training set, we fix the number of classes and change the number of samples per class in each training subset.

For testing data, we use the same testing data set for all experiments. The testing data set is composed of a gallery set and a probe set. The gallery set contains 72*10 face images of 72 different persons from the first session. The probe set contains 72*10 images of the same 72 persons from the second session. All the face images of the testing data set have not appeared in the training data sets.

Figure 1. Face samples after preprocessing.

Table 1. Different training data sets used in the experiments.

Training data sets		Number of all samples per subset	Number of classes	Number of samples per class
Training data set #1	Subset #1	1000	100	10
	Subset #2	1000	50	20
	Subset #3	1000	20	50
Training data set #2	Subset #4	100	100	1
	Subset #5	1000	100	10
	Subset #6	5000	100	50

3.4 Face Recognition Performance Using Different Training Data Sets

The face recognition results based on training data set #1 is shown in Table 2 and Fig. 2. For the three different training subsets, we compare their recognition performance using a number of different eigenfeature numbers ranging from 20 to1000. A probe image is considered correctly recognized if it matches any one of the ten images of the same person in the gallery set. The absolute accuracy is not important in the experiments. We intentionally use difficult data containing large facial expression changes to lower the overall recognition accuracy in order to compare the relative performance of different experiments.

From the results, we can see that the training subset #1 is slightly better than #2, which in turn is slightly better than #3, especially when the feature length is small. This shows that using images from more people can better characterize the eigenspace because of more inter-person variations in the training data set.

The face recognition results based on training data set #2 is shown in Table 3. The results seem again confirm what we observe in Table 2. If we look at the results below feature length 100, the three tests are fairly compatible. This shows that simply increasing the number of images per person will not affect the recognition results much. The number of people seems more important.

We focus more on the results of short feature lengths since they illustrate how efficient the transformation compress the large face vector. As the length of the feature vector increases, it becomes more like the original face vector. The effect of the transformation is largely lost. In fact, if we use the original face image directly for face recognition, we get an accuracy of 74.9%, which is actually the upper limit of the eigenface results. The advantage of the eigenface approach is not at improving the recognition accuracy, but rather is at improving the computational efficiency. We can use a feature vector of a few hundreds values to achieve comparable performance of the original image with thousands of pixels.

Table 2. Face recognition performance based on the three training subsets of training data set #1.

	Recognition Rate (%)		
Feature numbers	Training Data Subset #1	Training Data Subset #2	Training Data Subset #3
20	50.4	46.0	41.9
40	59.3	56.5	52.9
60	64.3	62.1	56.8
80	66.4	65.4	60.3
100	68.3	66.8	62.5
200	72.1	70.4	66.7
300	72.6	71.9	69.2
400	73.2	72.5	71.0
500	73.3	72.6	71.4
600	73.3	72.8	71.8
700	73.5	73.2	72.1
800	73.6	73.3	72.1
900	73.9	73.5	72.4
1000	73.9	73.5	72.4

Table 3. Face recognition performance based on the three training subsets in training data set #2.

<table>
<tr><td></td><td colspan="3">Recognition Rate (%)</td></tr>
<tr><td>Feature Numbers</td><td>Training Data Subset #4</td><td>Training Data Subset #5</td><td>Training Data Subset #6</td></tr>
<tr><td>20</td><td>51.7</td><td>50.4</td><td>49.7</td></tr>
<tr><td>40</td><td>57.7</td><td>59.3</td><td>59.6</td></tr>
<tr><td>60</td><td>61.1</td><td>64.3</td><td>64.7</td></tr>
<tr><td>80</td><td>64.3</td><td>66.4</td><td>66.8</td></tr>
<tr><td>100</td><td>68.1</td><td>68.3</td><td>68.2</td></tr>
<tr><td>200</td><td rowspan="13">Null</td><td>72.1</td><td>72.0</td></tr>
<tr><td>300</td><td>72.6</td><td>73.1</td></tr>
<tr><td>400</td><td>73.2</td><td>73.6</td></tr>
<tr><td>500</td><td>73.3</td><td>73.9</td></tr>
<tr><td>600</td><td>73.3</td><td>74.4</td></tr>
<tr><td>700</td><td>73.5</td><td>74.4</td></tr>
<tr><td>800</td><td>73.6</td><td>74.6</td></tr>
<tr><td>900</td><td>73.9</td><td>74.6</td></tr>
<tr><td>1000</td><td>73.9</td><td>74.6</td></tr>
<tr><td>2000</td><td rowspan="4">Null</td><td>74.7</td></tr>
<tr><td>3000</td><td>75.0</td></tr>
<tr><td>4000</td><td>74.9</td></tr>
<tr><td>5000</td><td>74.9</td></tr>
</table>

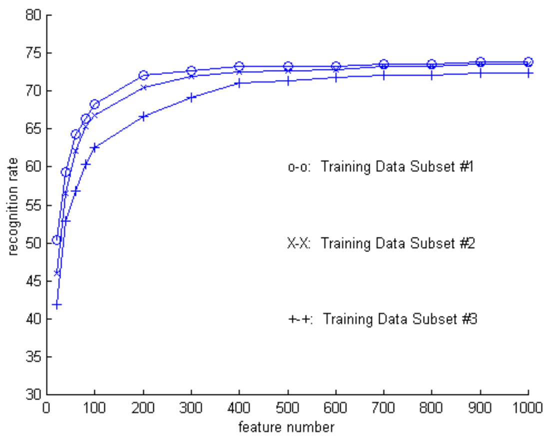

Figure 2. Face recognition performance using 3 different training subsets in training data set 1

3.5 Comparison of MDEE and KLT

In this section we use a simple experiment to illustrate that the MDEE method is a very close approximation of the KLT method. We apply MEDD and KLT separately on the same training data set: 1000 face images from 100 different people with 10 face images per person. Figure 3 shows that the values of the top 50 eigenvalues computed by the MDEE and KLT. The results of the two methods are nearly identical. The recognition results are shown in Table 4. Again, the results are nearly the same. From Fig 3 and Table 4, we can see that the performance of MDEE and KLT are very similar and MDEE is indeed a very close approximation of KLT.

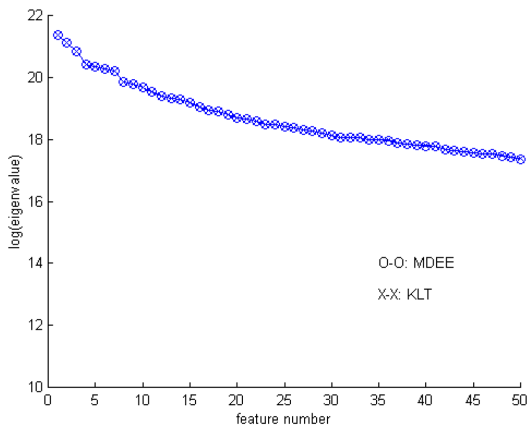

Figure 3. Top 50 eigenvalues of MDEE and KLT.

Table 4. Recognition rate comparison of MDEE and KLT

	Recognition Rate (%)	
Feature Numbers	MDEE	KLT
20	50.1	50.1
40	59.3	59.3
60	64.3	64.3
80	66.4	66.4
100	68.5	68.3
200	72.2	72.1
300	73.0	72.6
400	73.2	73.2
500	73.3	73.3
600	73.6	73.3
700	73.9	73.5
800	73.9	73.6
900	74.2	73.9
1000	74.2	73.9

4. CONCLUSIONS

In this paper, we explore the relationship between eigenface recognition performance and different training data sets. Using the MDEE algorithm we are able to compute eigenfaces from a large number of training samples. This allows us to compare the recognition performance using different training data sizes. Experimental results show that increasing the number of people benefits the recognition performance more than increasing the number of images per person. However, our results are still limited by the size of our database. Unfortunately, we only have

172 people in the database. We need a database with more people to further verify our conclusion.

5. ACKNOWLEDGMENT

We thank Dr. Qiumei Yang for many constructive comments. The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region. (Project no. CUHK 4190/01E).

6. REFERENCES

[1] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and machine recognition of faces: a survey," *Proceedings of the IEEE,* Vol. 83, pp. 705-741, May 1995.

[2] K. Fukunaga, Introduction to Statistical Pattern Recognition (Second Edition), Academic Press, New York, 1990.

[3] M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve procedure for the characterization of human faces," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 103-108, Jan.1990.

[4] X. Tang, "Texture information in run-length matrices," *IEEE Transactions on Image Processing*, vol. 7, No. 11, pp. 1602-1609, Nov. 1998.

[5] M. Turk and A. Pentland, "eigenfaces for recognition," *J. of Cognitive Neuroscience*, vol. 3, no. 1, 1991.

[6] M. Turk and A. Pentland "Face Recognition Using Eigenfaces," in *Proc. of IEEE International Conference Computer Vision and Pattern Recognition*, pp. 586-591, 1991.