Computer Project 1: Image Transforms & Applications

Name: Linsong Zhan

Due Date: Mar 21 at 11:59pm

1. Introduction

1.1. Purpose of the computer project

The goal of this computer assignment is to study PCA for image data reduction and facial pattern representation.

1.2. What will be accomplished or carried out?

To study PCA for image data reduction and facial pattern representation, the "Yalefaces" image database will be used for this assignment (yalefaces.zip with 165 images and its extended version with 5850 images). First, I will start by choosing only 15 training images (same condition or pose) and train the PCA-based system by finding the eigen-images. Next, I will test the quality of the reconstruction on some randomly picked training samples using only the most dominant PC's and display the results, and I will comment on the reconstruction ability both visually and using SNR measure. Next, I will test the reconstruction ability of the trained system on some randomly picked (at least 4) test cases (remainder of data not used for training) and comment on the results. Then, I will expand my training set to include 2/3 of the data set and keep the rest as testing set, and I will repeat the same process on these expanded data sets. Last, I will try to test the capabilities of this trained system on a new face (e.g., others) taken under similar conditions.

2. Theory

Algorithm for extracting eigen-images and reconstructing images

<u>Step 1</u>: obtain face images $I_1, I_2, \dots, I_M$ (training faces, where M is the total number of faces) with the size of $K \times L$.

<u>Step 2</u>: represent every image I_i as a vector Γ_i with the size of $KL \times 1$.

<u>Step 3</u>: compute the average face vector Ψ (psi) with the size of $KL \times 1$:

$$\Psi = \frac{1}{M}\sum_{i=1}^{M}\Gamma_i$$

<u>Step 4</u>: subtract the mean face Φ_i (phi) with the size of $KL \times 1$:

$$\Phi_i = \Gamma_i - \Psi$$

Step 5: compute the covariance matrix R:

$$R = \frac{1}{M}\sum_{i=1}^{M} \Phi_i \Phi_i^T = AA^T \ (KL \times KL\ matrix)$$

$$where\ A = [\Phi_1\ \Phi_2 \ldots\ \Phi_M]\ (KL \times M\ matrix)$$

Step 6: compute the principal eigenvectors x_i of R. Since the matrix AA^T is very large, it's not practical to calculate. Thus, we consider the matrix $A^T A$ ($M \times M\ matrix$), solve the M dimensional problem,

$$A^T A x_i = \lambda_i x_i$$

Now pre-multiplying this equation by A yields

$$(AA^T) A x_i = \lambda_i A x_i$$

Implying that the eigen-images can be obtained from

$$y_i = A x_i$$

Step 7: keep only $P \leq M$ eigen-images associated with the largest eigenvalues.

$$\eta = \frac{\sum_{i=1}^{P} \lambda_i}{\sum_{i=1}^{M} \lambda_i} \geq e.g, 95\%$$

Step 8: the reconstructed version of the ith training image from only P eigen-images is

$$\underline{\hat{I}_i} = \Psi + \sum_{k=1}^{P} \frac{{y_k^*}^T \Phi_i y_k}{\lambda_k}$$

3. Results and Discussions

Source: 15 images from yalefaces.zip

To represent 90% of the total variation of all face images, only the first 9 eigenfaces are needed. Figure1 and 2 below show the original $M = 15$ training faces and first $P = 9$ eigenfaces ordered row-wise.

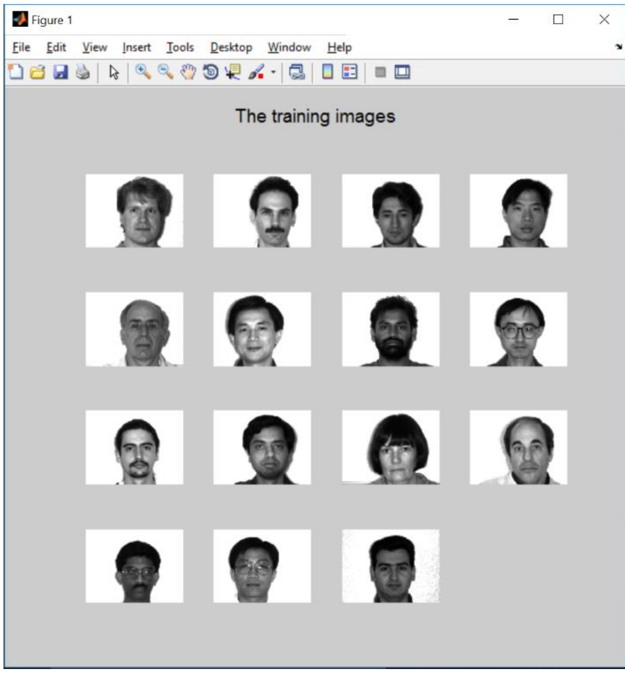


Figure 1

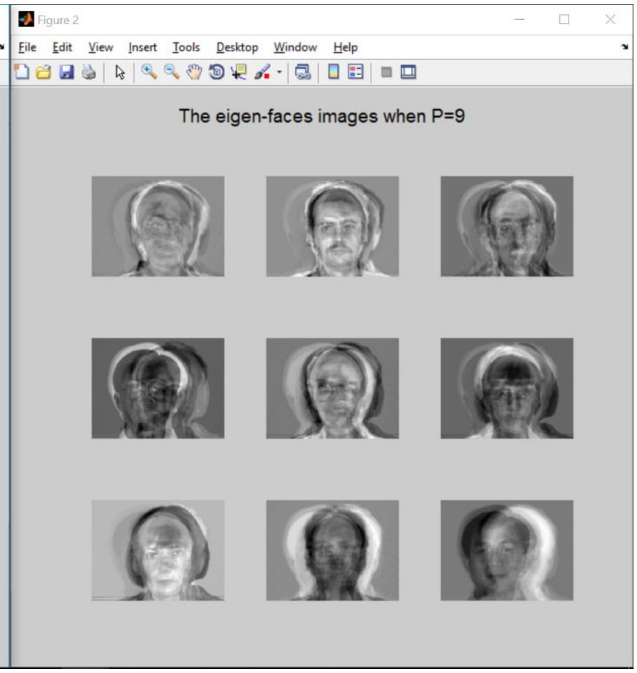


Figure 2

Figure 3 and 4 below show the process of reconstruction of two randomly picked training images using $M = 15$ training faces and only $P = 7$ eigen-faces representing 85% of the total variation of all face images.

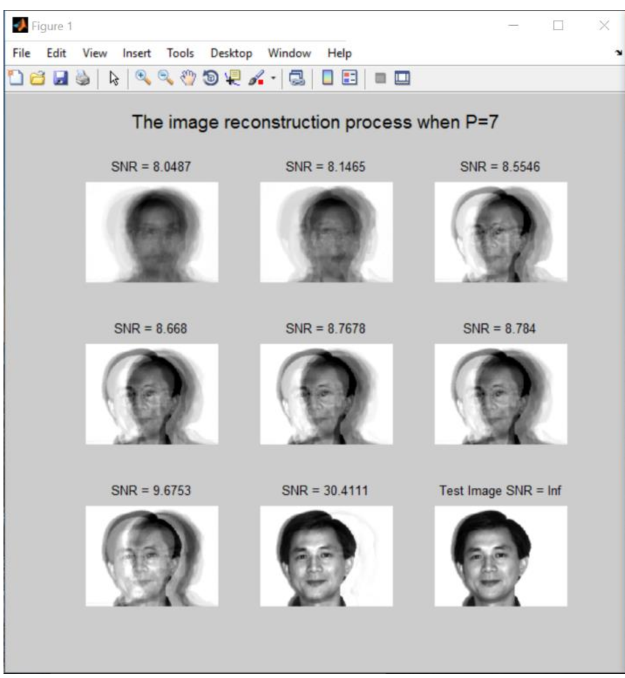


Figure 3

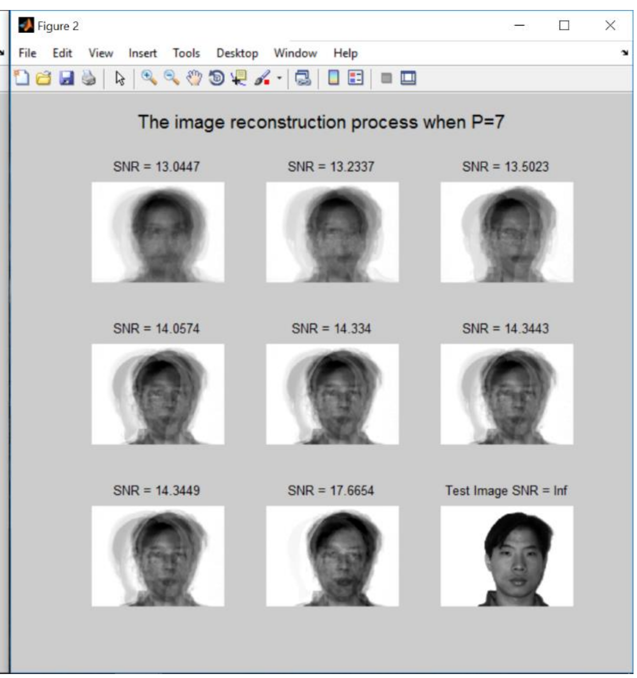


Figure 4

My comment: Both figures 3 and 4 show the images during the process of reconstruction visually. In the position (3,2) of both figures, the reconstructed images are similar with the test faces, which are in the position (3,3). In the SNR measure of an image, the greater the value of SNR is, the less percentage the noise of the image takes. As shown in the figures, the SNR is increasing during the reconstruction process because the reconstructed image becomes more and more clear. Note the SNR of test faces is infinite here because of the command in Matlab:

$$[peaksnr, snr] = psnr(double(testImage), double(testImage))$$

which leads the value of snr to infinite.

Figures 5, 6, 7 and 8 below show the process of reconstruction of four randomly picked test faces (remainder of data not used for training) using $M = 15$ training faces and only $P = 7$ eigen-faces representing 85% of the total variation of all face images.

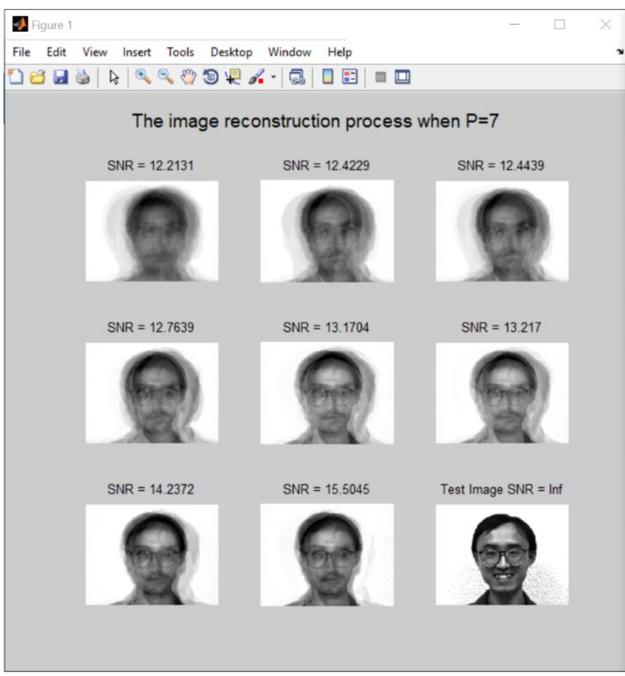


Figure 5

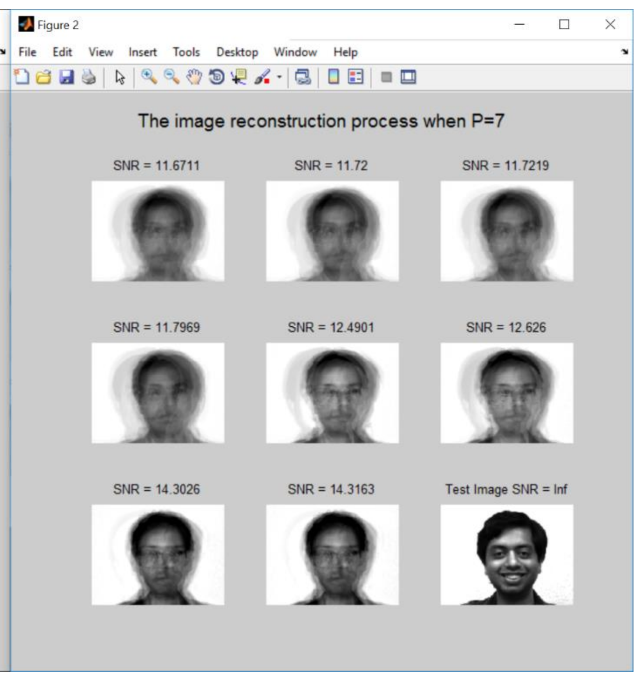
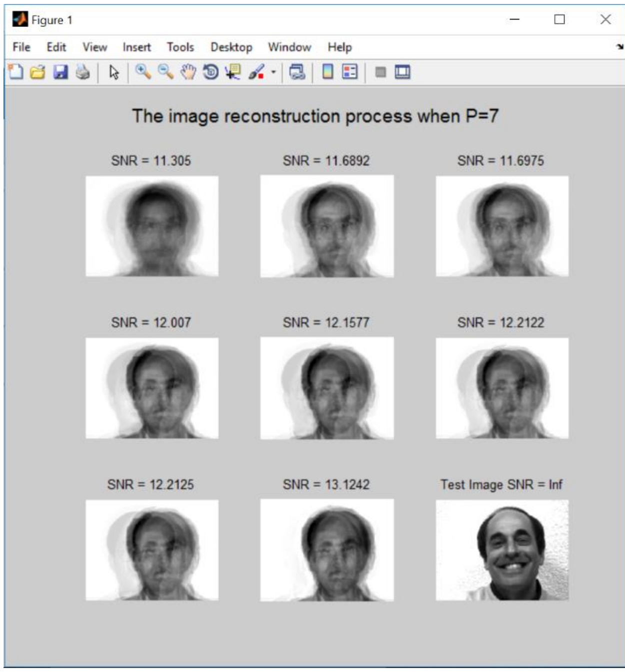


Figure 6



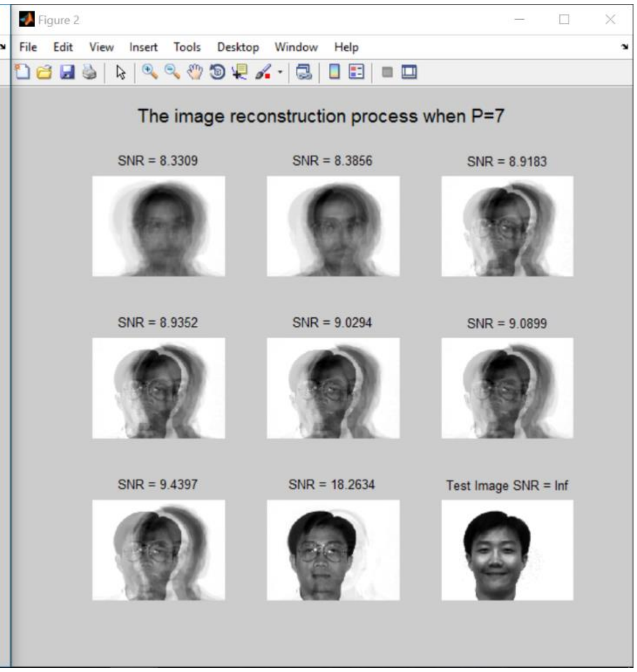
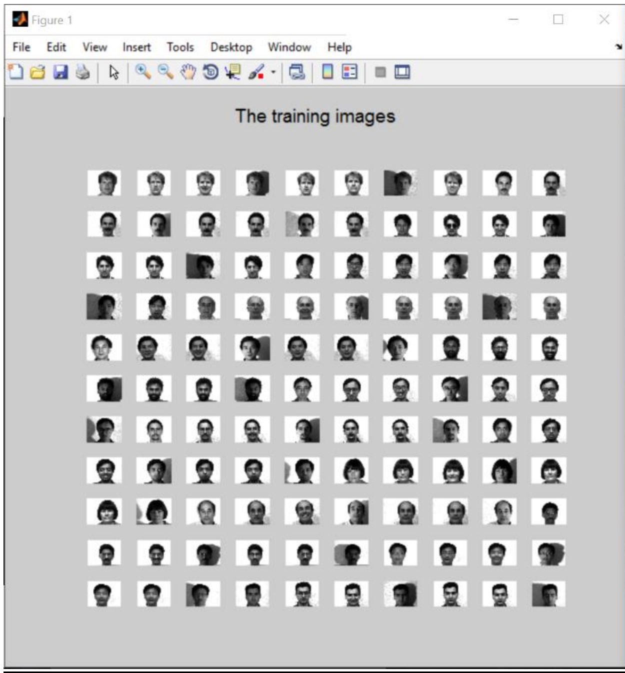


Figure 7 **Figure 8**

My comment: Figures above show the images become more and more clear visually. While the reconstructed images in the position of (3,2) are not very accurate. This is because the test faces here cannot be found in the training datasets. The SNR shows the same increasing as ones in the figure 3 and 4. In conclusion, the reconstruction of a test face that is not part of the training dataset is impossible.

Source: 110 images from yalefaces.zip

To represent 84% of the total variation of all face images, only the first 14 eigenfaces are needed. Figure 9 and 10 below show the original $M = 110$ training faces and first $P = 14$ eigenfaces ordered row-wise.



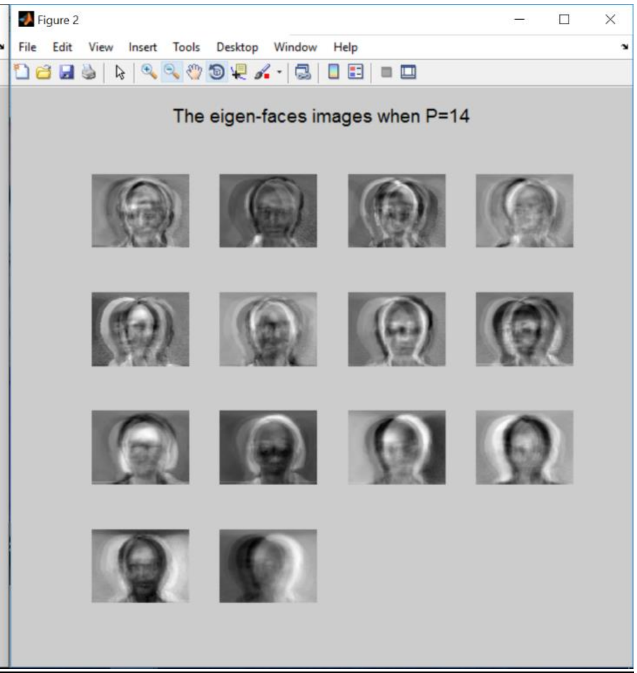
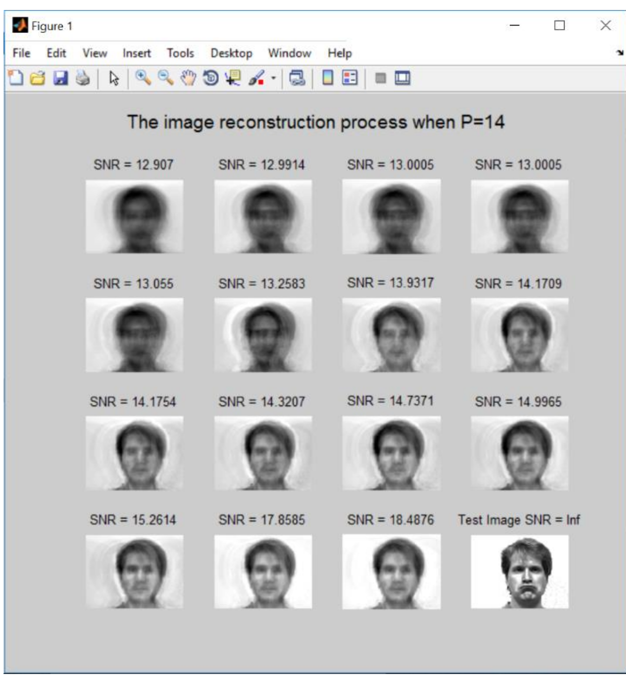


Figure 9 **Figure 10**

Figures 11, 12, 13 and 14 below show the process of reconstruction of four randomly picked test faces (remainder of data not used for training) using $M = 110$ training faces and only $P = 14$ eigen-faces representing 84% of the total variation of all face images.



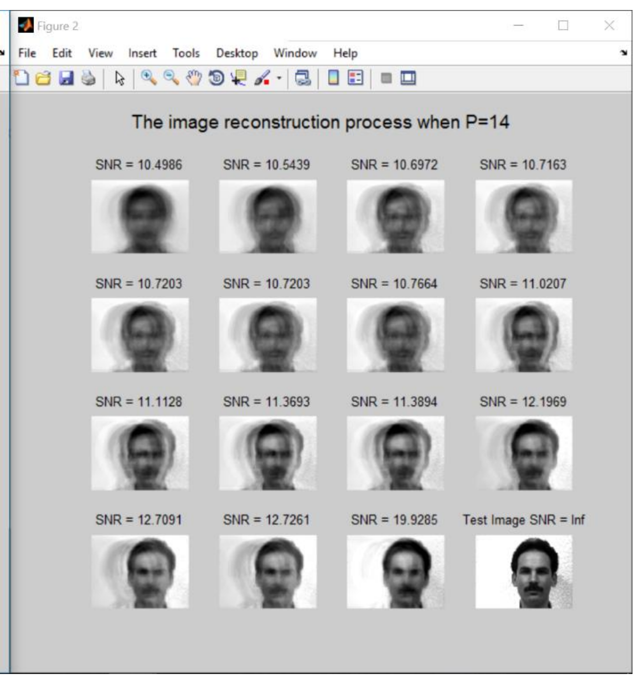


Figure 11 **Figure 12**

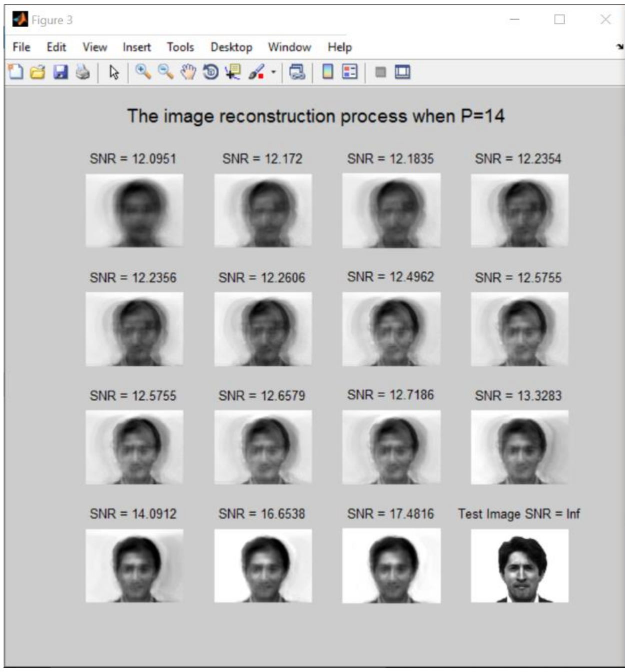


Figure 13

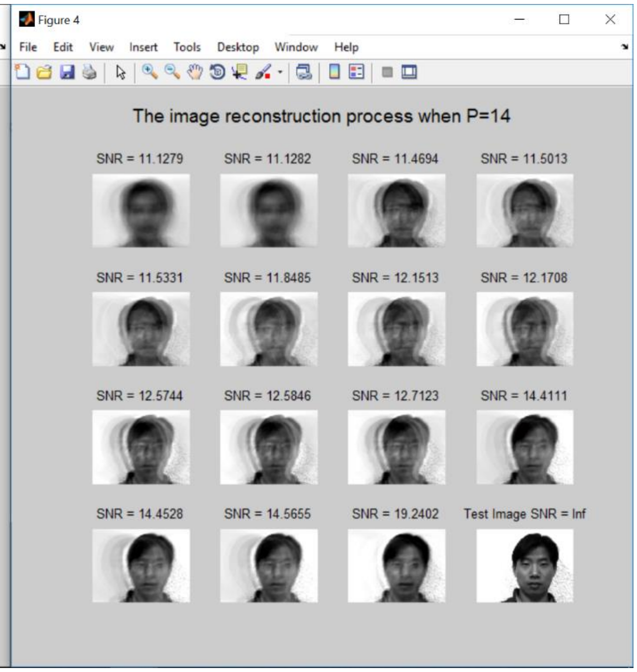


Figure 14

<u>My comment:</u> Figures above show the images become more and more clear visually. While the reconstructed images in the position of (3,2) are not very accurate. This is because the test faces here cannot be found in the training datasets. The SNR shows the same increasing as ones in the figure 3 and 4 during the process.

Figure 15 below shows the reconstruction process of a picture with left light using the $M = 110$ training faces and only $P = 14$ eigen-faces.

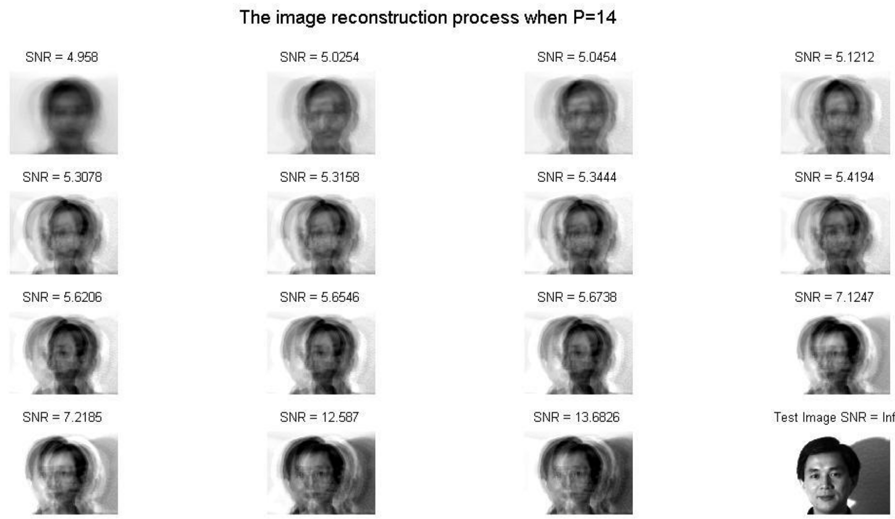


Figure 15

Figure 16 below shows the reconstruction process of a picture with left light using the $M = 110$ training faces and only $P = 14$ eigen-faces.

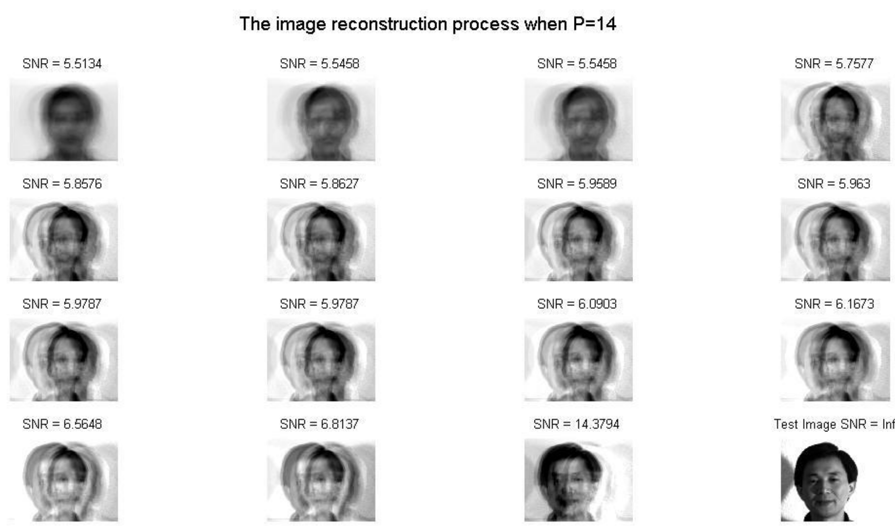


Figure 16

Visually the ability of face reconstruction using PCA decreases when the light condition changes because the reconstructed faces are ambiguous and low-contrast.

4. Conclusion

The technique of face reconstruction using PCA is proved to be visually effective when the test face can be found in the training faces. While one limitation is that the reconstructed face is not visually clear when the test face is not in the training faces, the other limitation is that the reconstruction process is also very time-consuming when it comes to 110 training images. In order to solve the first limitation, people use methods like nonlinear PCA to find test images that are different from the training images. In order to solve the second limitation, people switch to LDA (linear discrimination analysis) since LDA based algorithms achieves simply object reconstruction instead of optimizing the low dimensional representation of the objects with focus on most dominant components.

5. Reference

- http://www.vision.jhu.edu/teaching/vision08/Handouts/case_study_pca1.pdf
- ECE513 lecture slide 13-14