

Homework #4

Alexander Shender 328626114

Vladimir Tchuiev 309206795

Technion - Israel Institute of Technology

I. Eigen-Faces and PCA for face recognition

A. Section 1

We computed the average image from the grayscale training set of face images. The result is presented in Fig. 1. The resulting average image is a vague male face; notice the black lines on the top and bottom of the average image, they are presented in all training set images.



Figure 1: Average image from training set.

B. Section 2

From all the training set images, we computed the 25 eigenfaces with the largest eigenvalues. We flattened each 243×320 pixel image into a vector with 77760 elements. We concatenated all these vectors into a single $A \in \mathbb{R}^{77760 \times 150}$ matrix. In order to compute the eigenfaces, we will have to create the matrix $AA^T \in \mathbb{R}^{77760 \times 77760}$, the covariance matrix that corresponds to each pixel, and compute its eigenvalues and eigenvectors. This matrix is extremely large and computing its eigenvectors is costly, thus we will compute

the eigenvalues and vectors of $A^T A$ and use it to compute the eigenvectors of AA^T instead:

$$A^T A x = x \lambda \quad (1)$$

where λ is the eigenvalue of $A^T A$ and x is the corresponding eigenvector. Then:

$$(AA^T)(Ax) = (Ax)\lambda \quad (2)$$

thus λ is a (non-zero) eigenvalue of AA^T as well, and its corresponding eigenvector is Ax . Therefore we only need to compute the eigenvectors of $A^T A \in \mathbb{R}^{150 \times 150}$, which is a significantly more manageable task.

In Fig. 2 we present the 5 eigenfaces with the largest eigenvalues. The images are normalized such that the 0 value is 50 percent gray.

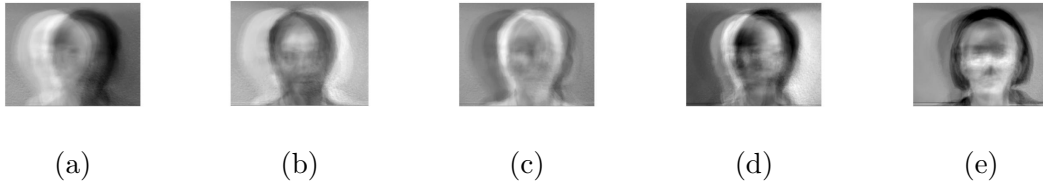


Figure 2: From left to right: eigenfaces with largest eigenvalues in descending order.

C. Section 3

In this section we reconstruct the training set images using the reduced representation, and reconstruction errors. Note that for the reconstruction we normalized the W matrix of eigenfaces such that $W^T W = I$.

In Fig. 3 we present training set images on the left column, compared to a reconstruction using 25 highest eigenvalue eigenfaces. We can see that the reconstruction is lacking detail, but overall pretty close to the training set. If we were to take all 150 eigenvalues, we would have a perfect reconstruction.

In Fig. 4 we present a graph of errors relative to the serial number of the image. We can see that dynamic range RMSE error has a significantly smaller value than RMSE error, as it tends to normalize large value swings. The average RMSE error is $79.54 \cdot 10^{-3}$, while the average Dynamic Range RMSE error is $3.337 \cdot 10^{-3}$.

D. Section 4

In this section we reconstruct the test set images using the reduced representation, and reconstruction errors.

In Fig. 5 we present test set images on the left column, compared to a reconstruction using 25 highest eigenvalue eigenfaces trained on the original 150 images. We can see

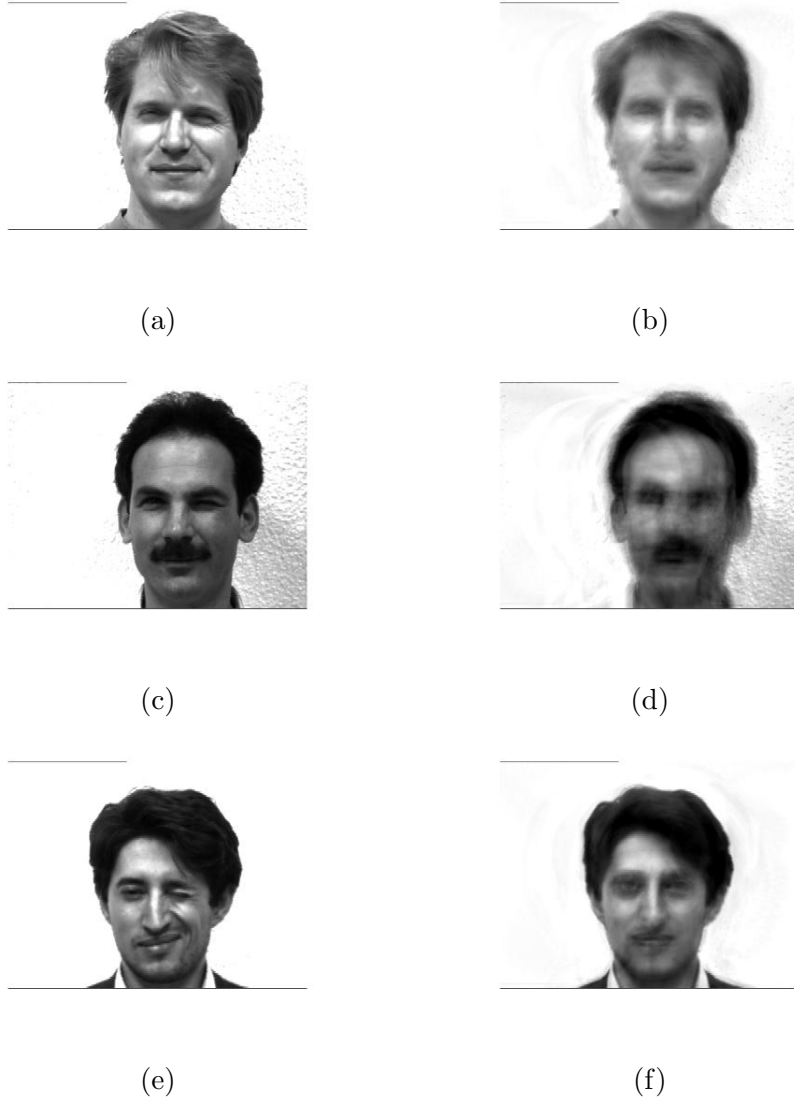


Figure 3: Training set; Left column: original images from training set. Right column: image reconstruction based on 25 largest eigenvalue eigenfaces.

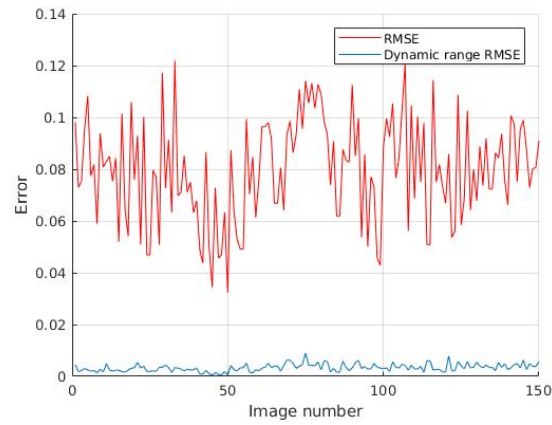


Figure 4: Training set; RMSE and Dynamic Range RMSE errors vs. image serial number.

that the reconstruction captures faces' general form well but fails to reconstruct objects that are not faces, such as the Google Chrome symbol. This reconstruction is enough for classifying faces relatively well, as the classification accuracy we reached for the eigenfaces algorithm is **81.82 percent**. We performed the classification over all face images, and ignored the non-face images as the inferred class has no meaning for these cases.

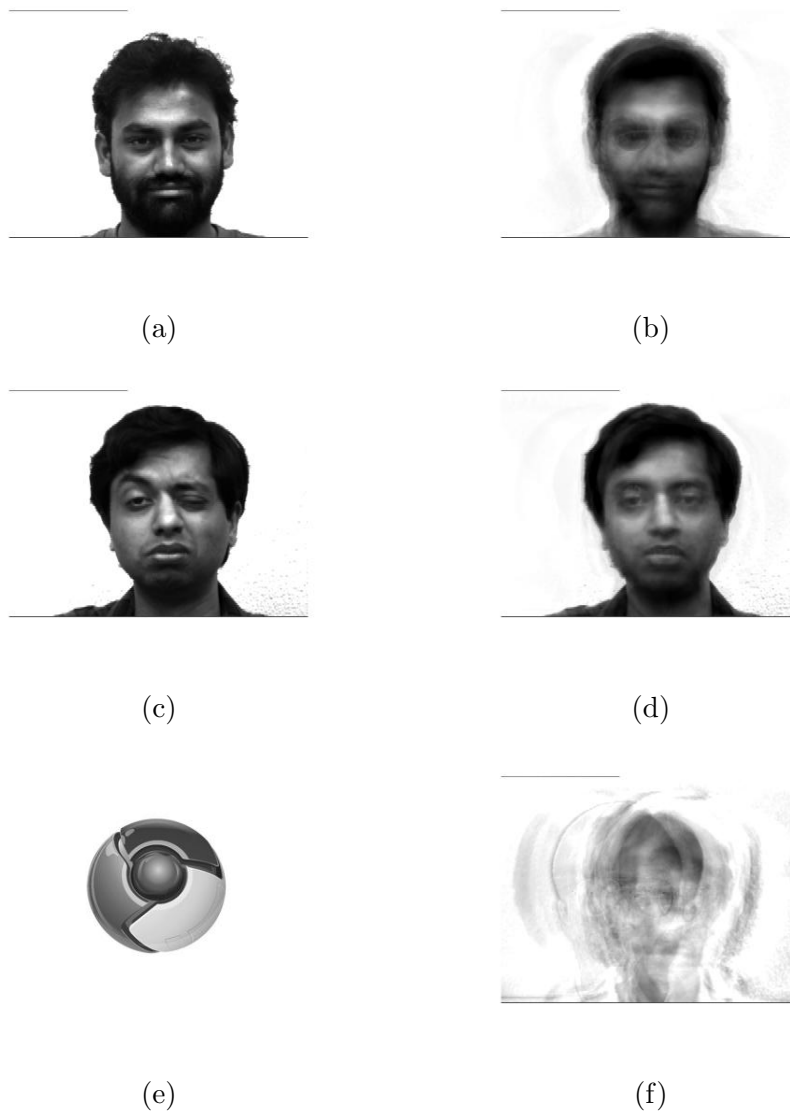


Figure 5: Test set; Left column: original images from training set. Right column: image reconstruction based on 25 largest eigenvalue eigenfaces.

In Fig. 6 we present a graph of errors relative to the serial number of the image. The average RMSE error is $143.8 \cdot 10^{-3}$, while the average Dynamic Range RMSE error is $6.890 \cdot 10^{-3}$. As expected, the reconstruction error for the test set is significantly larger than the training set, also due to the fact that the test set contains images that are very far from the training set (high model uncertainty).

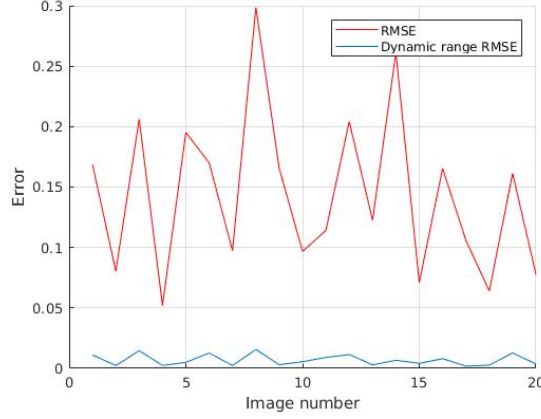


Figure 6: Test set; RMSE and Dynamic Range RMSE errors vs. image serial number.

II. Fisher Linear Discriminant for face recognition

A.

The stub code 'fishertrain.m' was completed as described in the presented steps. The matrix W_{fld} was calculated using the generalized eigenvalues decomposition (GED), containing $c - 1$ largest eigenvectors. The resulting Fisher basis is indeed:

$$W = W_{pca}W_{fld}$$

. Explanation: for the Fisher basis, we did not use the original image vectors (which are of the size of $234 * 320 = 77760$), in order to avoid huge matrixes S_B and S_W . We have used instead their representation using the PCA basis, which was already calculated. By setting the size of the basis for PCA to be $r = N - C$, the scatter matrixes S_B and S_W have resulted in being the size of $[r \times r]$ (scatter matrixes are square matrixes, $[r \times r]$, where r - number of dimensions in data). Thus, the W_{fld} matrix, after taking $c - 1$ eigenvectors with largest eigenvalues, is of the size $[r (c-1)]$. The W_{pca} , on the other hand, projects the image vector to a vector of size r . Thus, its size is $[77760 \times r]$. Eventually, we obtain matrix W , which projects the image vector first on the PCA basis, and then on the $c - 1$ strongest eigenvectors of the GED. Thus, we obtain:

$$W = W_{pca}[77760 \times r]W_{fld}[r \times c - 1] = W[77760 \times c - 1]$$

And W projects the image vector to a Fisher basis.

B.

By using the parameter $c = 15$ we obtain the sizes for the matrixes (for sanity check)

$$\begin{aligned}
R &= N - c = 135 \\
S_B, S_W &\in [135 \times 135] \\
W_{pca} &\in [77760 \times 135] \\
W_{fld} &\in [135 \times 14] \\
W &\in [77760 \times 14]
\end{aligned}$$

Thus, we project each image on a vector of size 14. Using the code from the section A, we calculate the representation errors for the training set and for the test set.

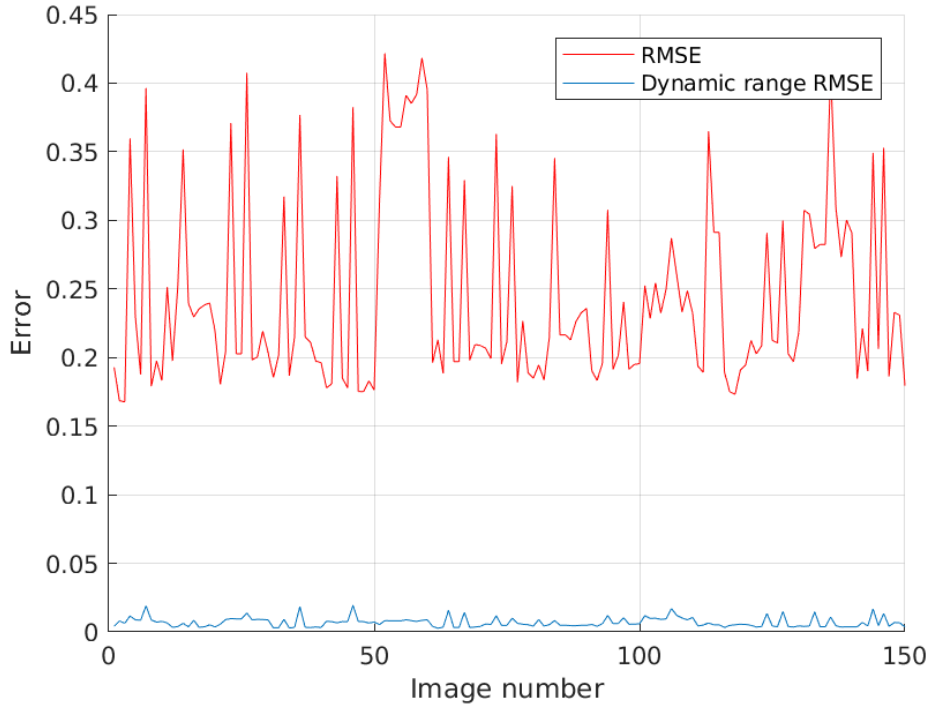


Figure 7: Training set reconstruction errors

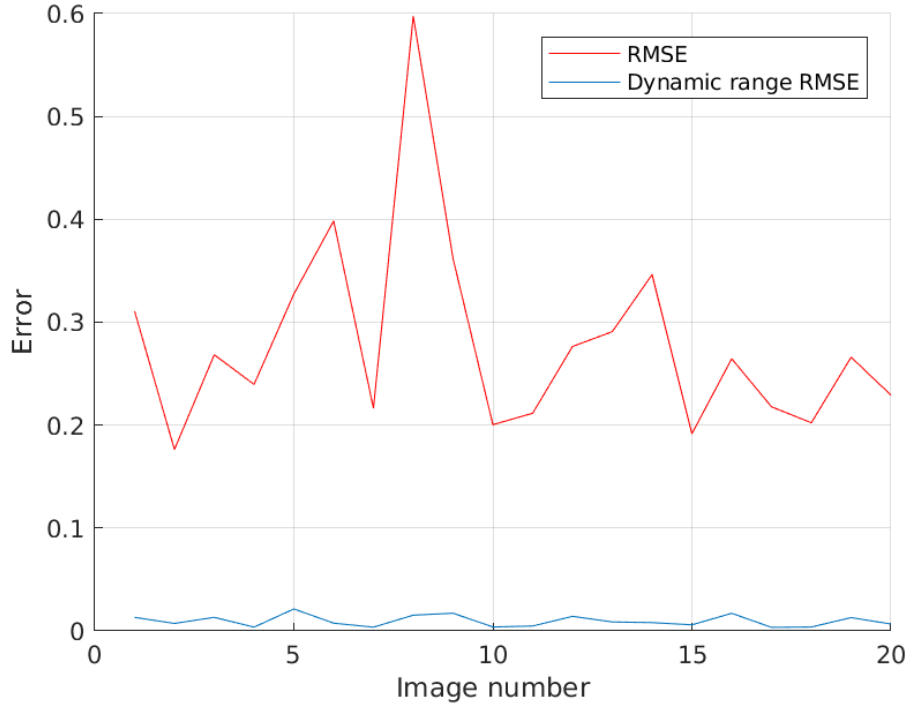


Figure 8: Test set reconstruction errors

We can see that the reconstruction errors are larger than the reconstruction from the PCA basis alone from the previous section. We can explain this by the fact the representation vector from the Section A was of size 25, while in this section we use the representation of size 14. Moreover, while PCA aims to give the best representation of the data, the FLD aims to separate the data in the most optimal way, thus the FLD representation is worse, but the classification results are better, as is demonstrated in the next subsection.

C.

Using the FLD, we achieve a fantastic classification rate of 100%! This demonstrates that using the Fisher Basis, we achieve better classification of the dataset. The explanation can be found in numerous resources.

The PCA finds the most accurate representation of the data, in a lower dimensional space. It finds the directions of maximum variances, but those may be not so good for classification, since they may be similar for different classes. What Fisher Linear Discriminator does is finds a projection of the representation of the data, where the classes are separated in an optimal way - where the separation is good. This is done by maximizing the between-classes scatter matrix, and minimizing the within-classes scatter matrix. Good separation means good classification - there is less overlapping of the classes representations on the projection subspace.