Face Recognition Assignment

Computer Vision (2018-500-KEN4255)

Assignment Report

Department of Data Science and Knowledge Engineering

Maastricht University

Tomasz Stańczyk (6209867)

1. Overview

This report summarizes the work done over the second assignment from the Computer Vision course. The purpose of this assignment was to perform face recognition with the eigenfaces (as eigenvectors), with the use of approaches and techniques presented during the related lecture.

The data which was provided for the sake of this assignment included:

- main data matrix storing 10 different images for 40 people (400 in total), each with original size 32x32 pixels and vectorized into 1024-element vector,
- ground truth vector containing corresponding number of people represented by each face image,
- vectors indicating indices of the training samples (referring to the main data matrix),
- vectors indicating indices of the testing samples (as above).

The last two elements mentioned were provided with different variants, respectively: 3, 5 and 7 face images per person to be used for training and the remaining ones to be used for tests.

The whole implementation has been performed in GNU Octave, a scientific programming language, syntax of which is largely compatible with MATLAB. The following files containing the documented code have been attached to this report: *eigenfacesrecognition_training.m*, *eigenfacesrecognition_test.m*, *visualisation_script.m*, *main_script.m*.

This report contains the following sections. Section 2. describes the steps taken with respect to calculating the eigenvectors. Section 3. presents relevant visualizations and contains image descriptor derivation approach accompanied by the image reconstruction. Section 4. focuses on the applied face recognition methods and Section 5. Discusses the accuracy measurements performed. Additionally, a short summary is provided in section 6.

2. Eigenvectors Calculation

Eigen analysis as described during the lecture has been applied. During the training, first k (varying value) eigenvectors were being calculated. In order to optimize the computation process, instead of calculating the covariance matrix:

$$\Sigma = \frac{1}{N} X X^T,$$

another matrix was taken into account:

$$T = \frac{1}{N} X^T X.$$

In the equations above, X stands for the processed data in form of matrix, whereas N corresponds to number of samples, in this case – faces.

Comparing the sizes of the matrices described above, Σ is of size d x d (where d stands for the dimensions), in this case: 1024x1024. On the other hand, T is of size Nx N, in this case: 400x400. This optimization approach reduces time required to compute eigenvectors significantly.

The eigenvectors were being calculated with the use of Octave's svd() function (singular value decomposition) which returned them as the columns of a matrix. By applying the following statement presented during the lecture:

$$\Sigma(Xe_i) = \gamma_i(Xe_i),$$

and referring to the general formula:

$$\Sigma u = \lambda u,$$

the eigenvectors with respect to covariance matrix Σ were received by multiplying the eigenvectors obtained for the smaller matrix T with X. In the formulae above, γ_i and e_i correspond to eigenvalues and eigenvectors of T respectively, whereas λ and u represent eigenvalue and eigenvector with respect to Σ. It is crucial to mention that the newly received eigenvectors were normalized so as to retain their norm of 1, which originally was lost after multiplication by X.

Having received a matrix storing desired eigenvectors as its columns, first k eigenvectors were being returned as a result of the training.

3. Visualization, Image Descriptors and Image Reconstruction

In order to visualize the eigenfaces as eigenvectors and demonstrate the image reconstruction based on these eigenvectors, the following steps have been applied. Firstly, the provided face image data has been normalized using the min-max normalization approach. Next, the average face μ have been computed by taking mean of all the training samples. The visual result is presented in Figure 1.

Figure 1 - Visualization of the average face (calculated over the training set)

The training data represented as X was modified by subtracting the average face image from all the face images included in the training set:

$$X = X - \mu.$$

Subsequently, k first eigenvectors were computed, as described in the previous section. Visualization of first 30 eigenfaces is presented in Figure 2. For the sake of better demonstrability, the visualized content has been min-max normalized beforehand.

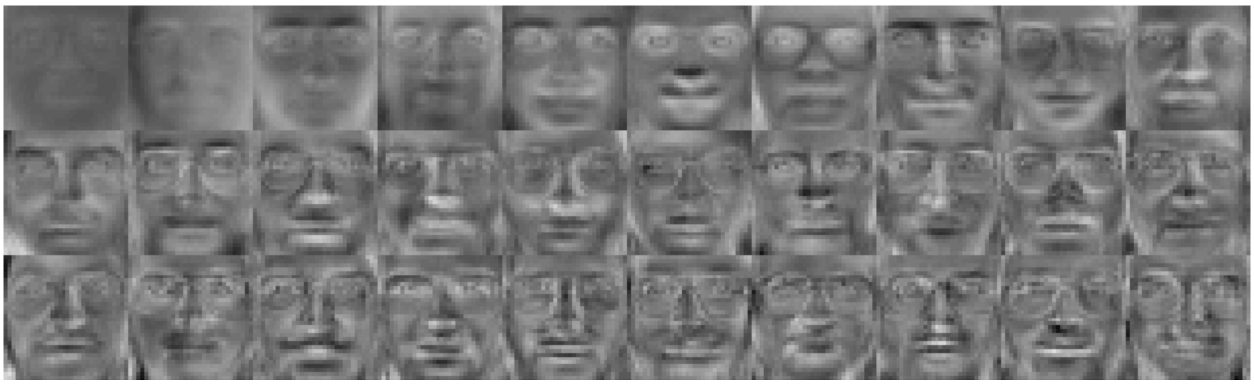

Figure 2 - Visualization of eigenfaces (k=30) with the first one in the top left corner and the last one in the bottom right corner

Further, with the use of the eigenvectors discussed, the training data has been projected on the new space and descriptors w for each training image have been derived. The following formula has been applied:

$$W = UX$$

Where W stands for a matrix representing vector descriptors for each image, U corresponds to a matrix containing k eigenvectors and X represents the training data, values of which has already been modified by subtracting the average face.

Next, based on the descriptors and eigenfaces (as eigenvectors), an image reconstruction has been performed. For this purpose, the following formula has been utilized:

$$X_{rec} = \mu + U^T W.$$

An exemplary reconstructed image along with the original image are presented in Figure 3 and 4 respectively.

Figure 3 - An exemplary reconstructed face image (k=20)

Figure 4 - The original image corresponding to the one reconstructed in Figure 3

4. Face Recognition

Following the image reconstruction and eigenfaces visualization, the face recognition has been performed. Analogously as described in the previous section, the data descriptors for both training and test set have been derived. Referring to the descriptors, for each test sample, a relevant training sample has been associated. The assignment process was based on the Nearest Neighbor approach. To be precise, Euclidean distances between the given test sample and all the training samples were being calculated and according to the shortest distance, selected training sample was related to the test sample.

Next, the ground truth values were compared for both test and selected training sample. Correct classification was indicated by the same values compared and incorrect classification was stated otherwise.

5. Accuracy Measurements

Based on the classification process described in the previous section, the accuracy was measured. For a single value of k and particular proportion between number of training and test samples per one face, the accuracy was computed as a ratio between the test samples classified correctly and all the samples in the test set. Therefore, values from the range [0, 1] have been received.

The aforementioned accuracy evaluation has been performed over 3 various sets of training and test index arrangements:

- 3 training samples (out of 10) and 7 test samples per face (the *3Train/3.mat* file),
- 5 training samples and 5 test samples (the *5Train/5.mat* file),
- 7 training samples and 3 test samples (the *7Train/7.mat* file),

and considered for each of the k value from the set: {*5, 10, 15, 20, 25, 30, 35, 40, 45, 50*} separately. Figures 5 – 7 display the plots received during these measurements, showing accuracy against the corresponding value of k.

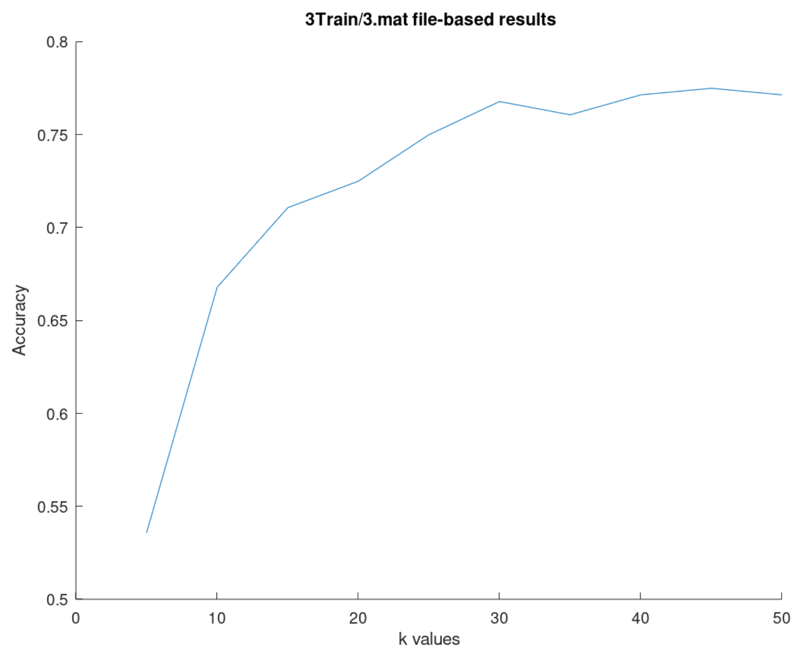

Figure 5 - Accuracy versus k - 3 training images per face

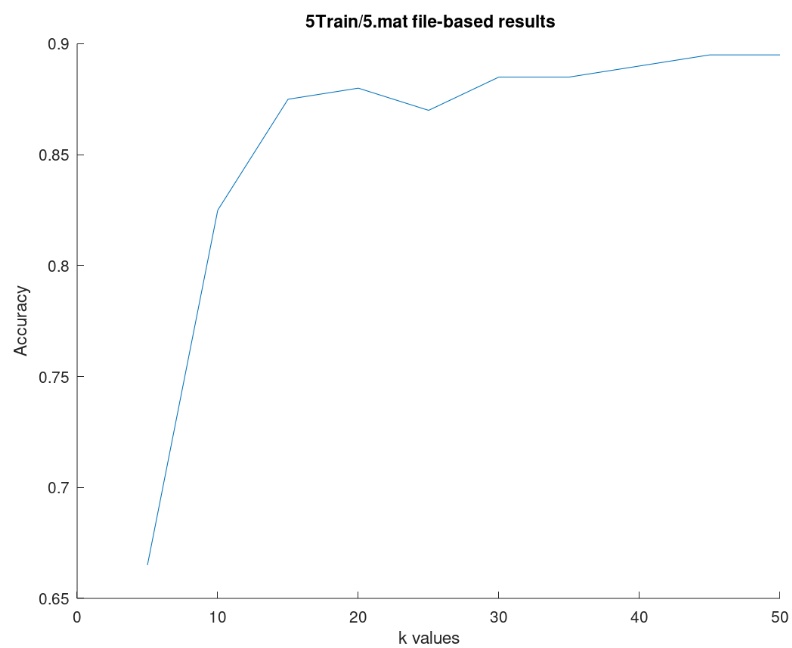

Figure 6 - Accuracy versus k - 5 training images per face

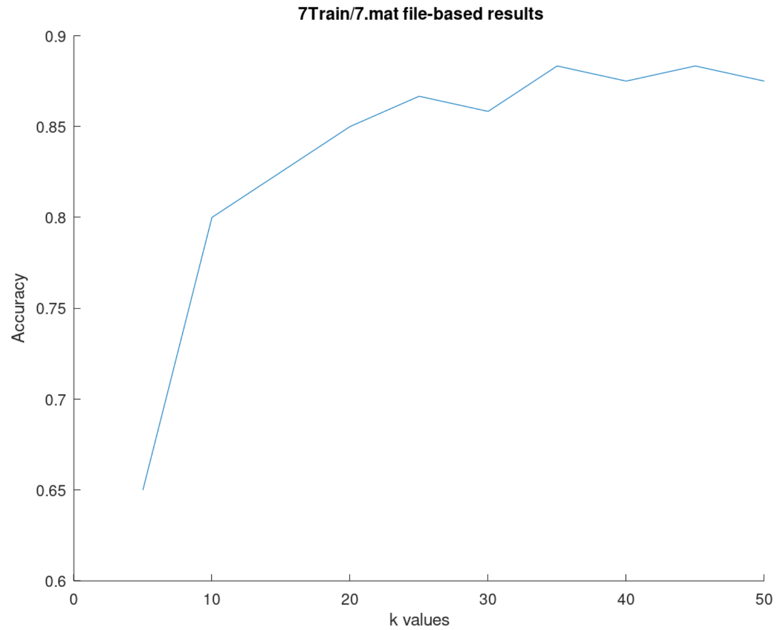

Figure 7 - Accuracy versus k - 7 training images per face

For better comparison, all the plots from Figures 5 – 7 were juxtaposed together in Figure 8.

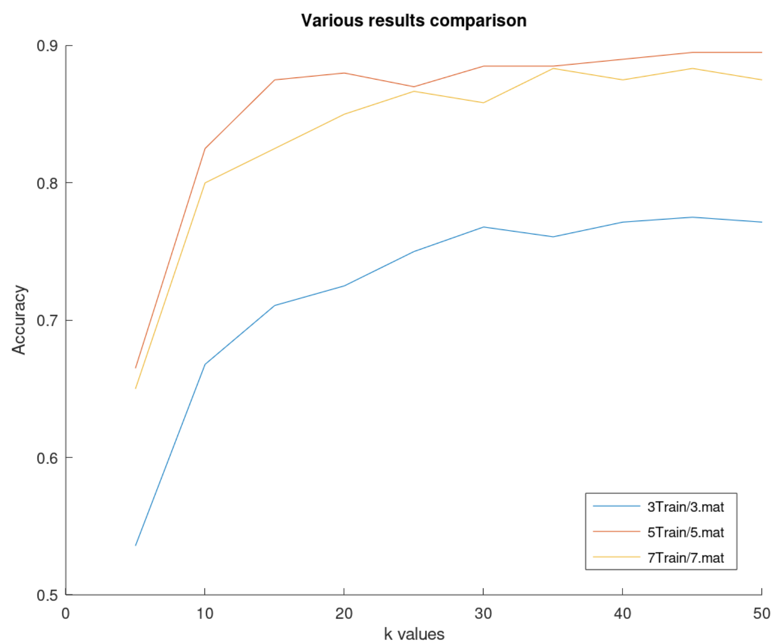

Figure 8 - Accuracy versus k - 3, 5 and 7 training images per face respectively

As it can be observed, in all three cases, the highest accuracy gains occur between k set to 5 and 10. Furthermore, starting from certain values of k, the differences between accuracies along with the increase of k are quite insignificant. These values are $k=30$ for the 3- and 5 training images per face (TIF) datasets and $k=35$ for the 7 TIF dataset. They could be considered as the most optimal values (for the corresponding dataset training and test samples ratio) regarding the accuracy and size trade-off (the smaller the value of k, the less data needs to be stored).

Interestingly, for all the k values tested, the accuracy measured over the 5 TIF dataset is always higher than the accuracy measured over the 7 TIF dataset. However, these differences are less significant than the ones observed between 7 TIF and 3 TIF dataset or even between 5 TIF and 3 TIF dataset accuracies.

This suggests that the number of training samples needs to be sufficiently large with respect to the number of test samples. Nevertheless, the former should not exceed the latter considerably.

6. Summary

This report described the steps performed over the realization of the face recognition with eigenfaces assignment. Calculation of first k eigenfaces as eigenvectors has been described and first 30 eigenfaces have been visualized. Moreover, deriving the image descriptors has been discussed and image reconstruction has been demonstrated along with exemplary image visualizations. Further, face recognition has been described and accuracy measurement have been presented. Finally, the respective observations over the varying parameters have been discussed.