Programming Assignment 3

CS 479

`https://github.com/alexander-novo/CS479-PA3`

Nikhil Khedekar	Mehar Mangat	Alexander Novotny
33%	33%	33%
??	??	??

Due: April 26, 2021
Submitted: April 25, 2021

Contents

1 Theory

This assignment considers the usage of Principal Component Analysis for dimensionality reduction for a dataset of images for the purpose of face recognition.

1.0.1 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a method for reducing the dimensionality of high dimensional data while preserving as much information as possible. The way this works is by the projection of the sample data along its principal distribution directions. Intuitively, this can be understood by finding the directions that capture the highest variance of the data and projecting the data along these directions as seen for a 2D to 1D case in.

For some given sample data $\mathbf{X} = \{\vec{x}_i \in \mathbb{R}^N, 1 \leq i \leq M\}$, where M is the number of samples in the data set, the sample covariance matrix Σ_x can be found using eq. (1) and eq. (3), where $\bar{x}$ is the sample mean and Φ_i are the normalized data points:

$$\bar{\mathbf{x}} = \frac{1}{M}\sum_{i=1}^{M}\mathbf{x}_i \tag{1}$$

$$\mathbf{\Phi_i} = \mathbf{x}_i - \bar{\mathbf{x}} \tag{2}$$

$$\mathbf{\Sigma_x} = \frac{1}{M}\sum_{i=1}^{M}(\mathbf{x}-\bar{\mathbf{x}})(\mathbf{x}-\bar{\mathbf{x}})^T = \frac{1}{M}\sum_{i=1}^{M}\mathbf{\Phi}_i\mathbf{\Phi}_i^T \tag{3}$$

Equation (3) can be further simplified to eq. (4) using eq. (5).

$$\mathbf{\Sigma_x} = \frac{1}{M}\mathbf{A}\mathbf{A}^T \tag{4}$$

$$\mathbf{A} = [\Phi_1\Phi_2...\Phi_M] \tag{5}$$

The eigenvectors and eigenvalues of the covariance matrix have the special property that the eigenvectors form an orthogonal basis in R^N for the sample data with the corresponding eigenvalues representing the variance of the data in these directions. The eigenvectors and corresponding eigenvalues are obtained using eq. (6).

$$\mathbf{\Sigma_x u_i} = \lambda_i u_i \tag{6}$$

In our notation, we assume that the eigenvalues are indexed in descending order, namely, $\lambda_1 > \lambda_2 > ... > \lambda_N$. Using the eigenvectors as the new basis we can represent our data samples $\mathbf{x} \in \mathbb{R}^N$ as:

$$test \tag{7}$$

1.0.2 Eigenfaces and Facial Recognition

The first widely successful method of solving the facial recognition problem is the eigenface method. An eigenface

2 Implementation

3 Results and Discussion

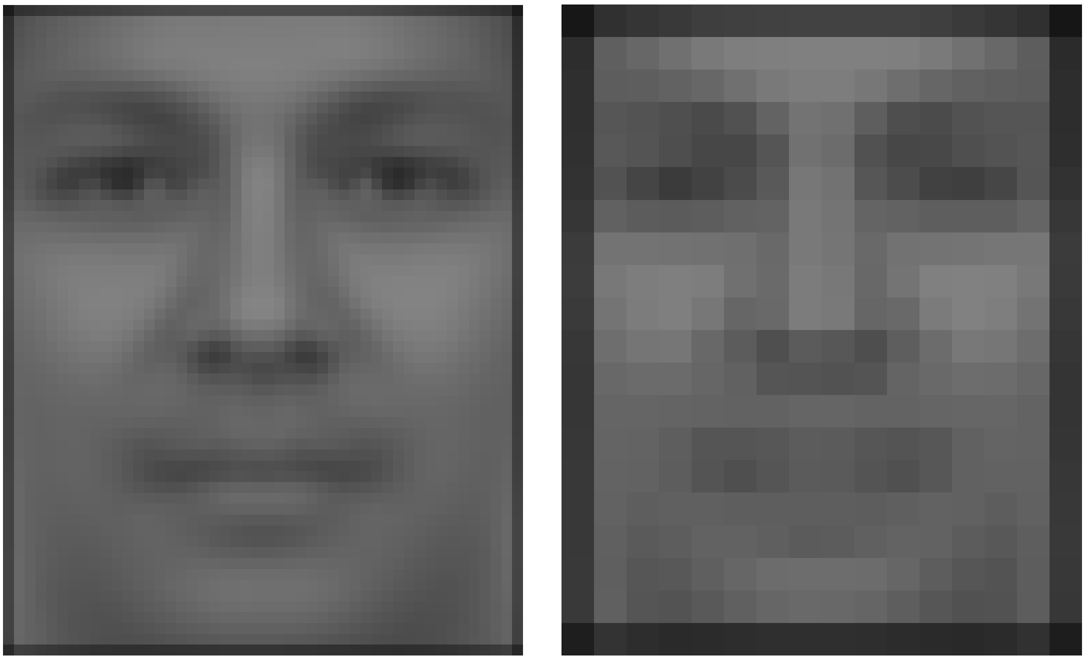

Figure 1: The sample mean face for the high resolution data set and the low resolution data set, respectiely.

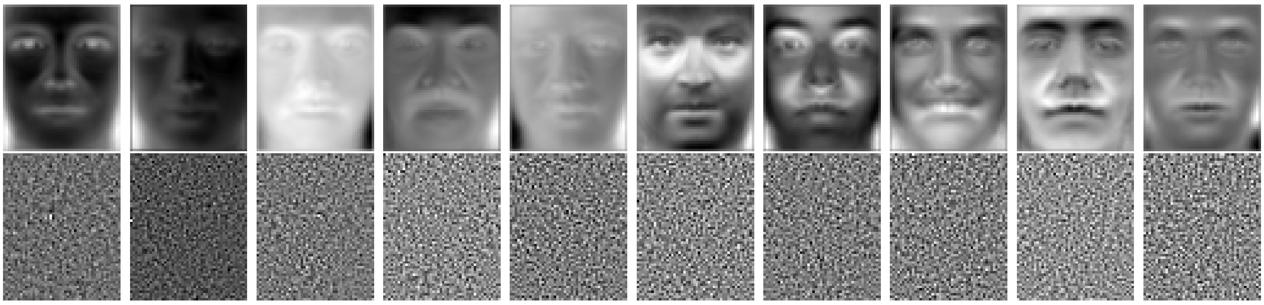

Figure 2: The 10 largest eigenfaces in decreasing order for the high resolution data set, followed by the 10 smallest eigenfaces in increasing order.

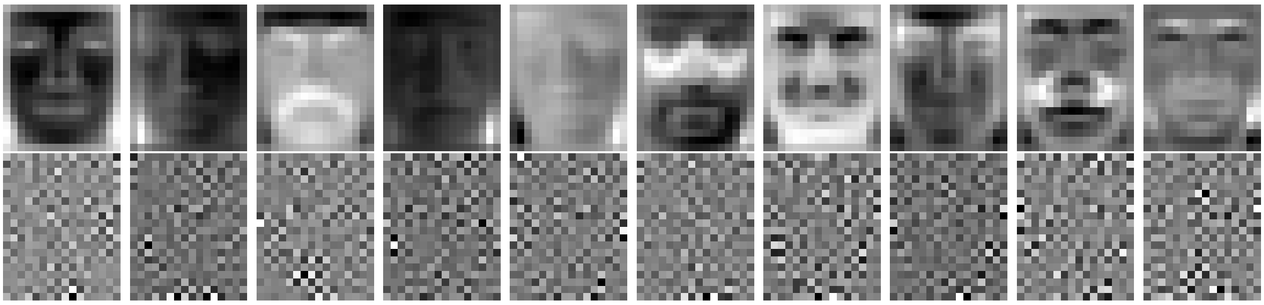

Figure 3: The same figure as fig. 2, but for the low resolution data set.

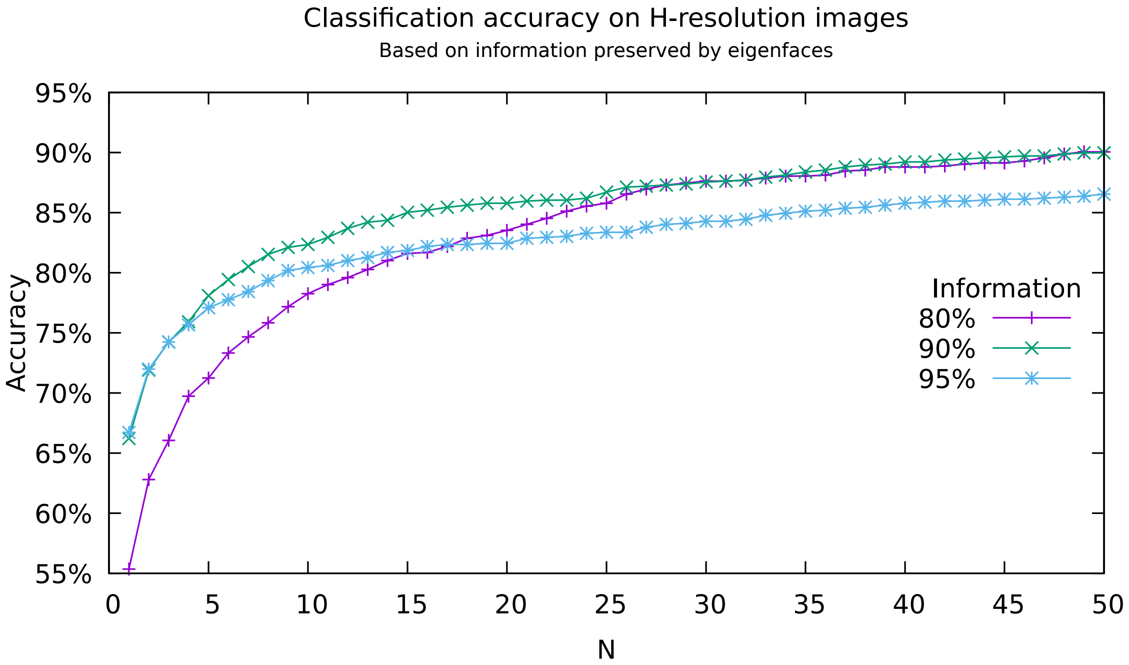

Figure 4: A comparison of classification accuracy of the high resolution images against number of images considered (N) and the percentage of information preserved by eigenfaces.

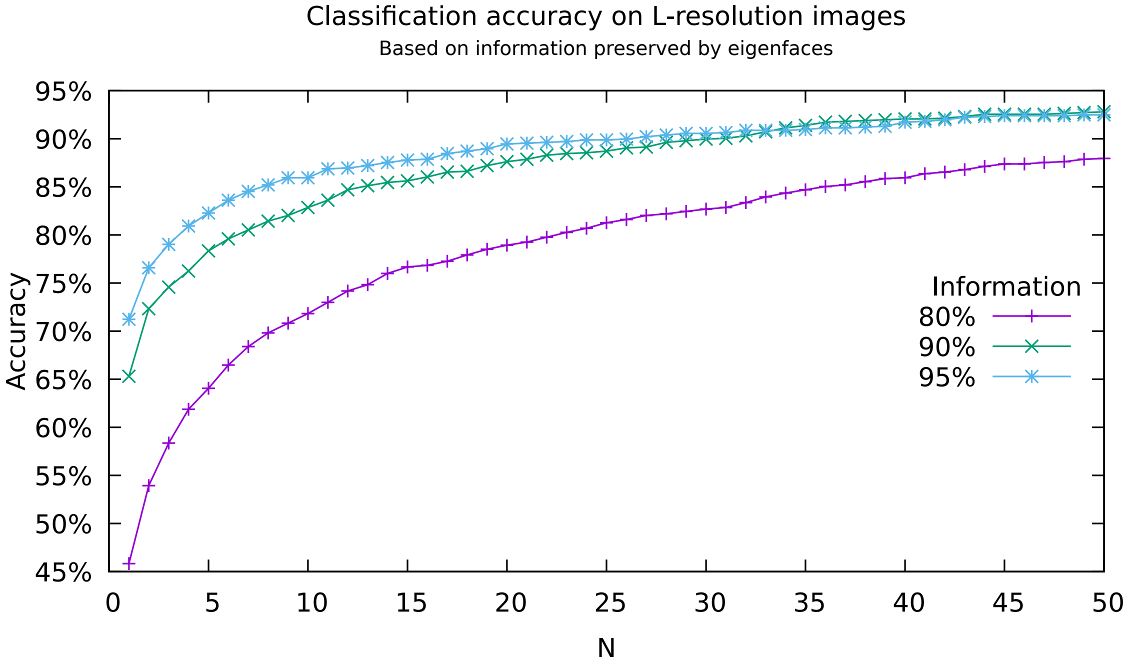

Figure 5: A comparison of classification accuracy of the high resolution images against number of images considered (N) and the percentage of information preserved by eigenfaces.

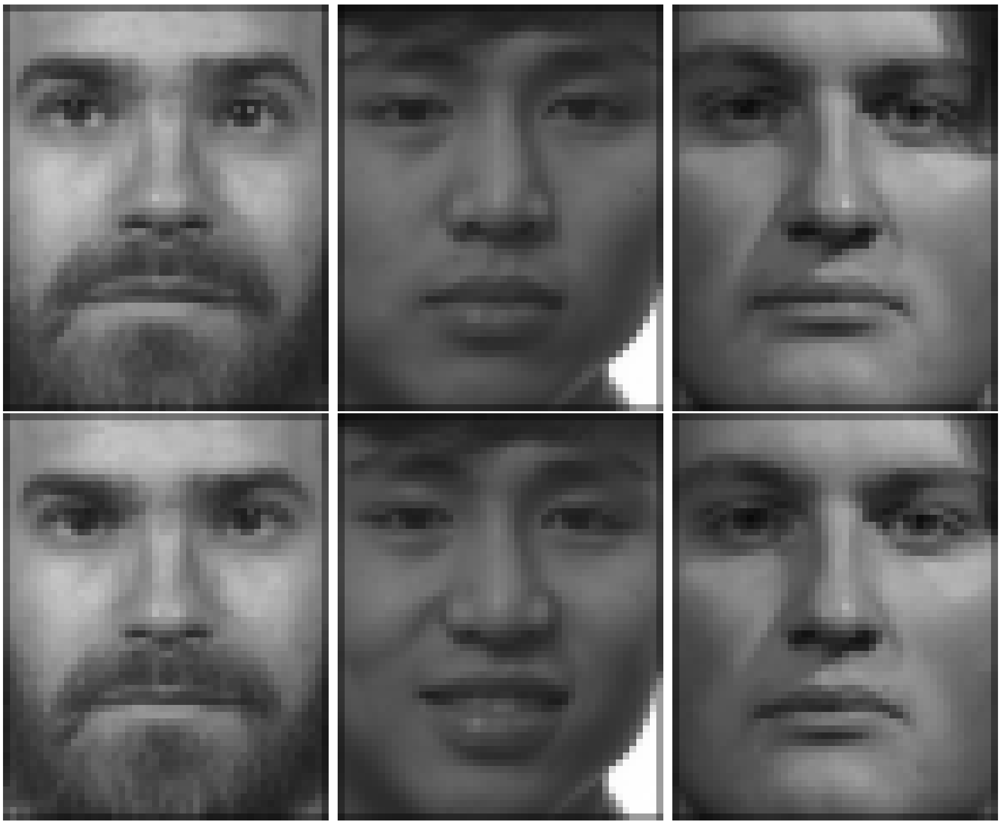

Figure 6: Three high-resolution test images (top) and the training images they were correctly matched against (bottom).

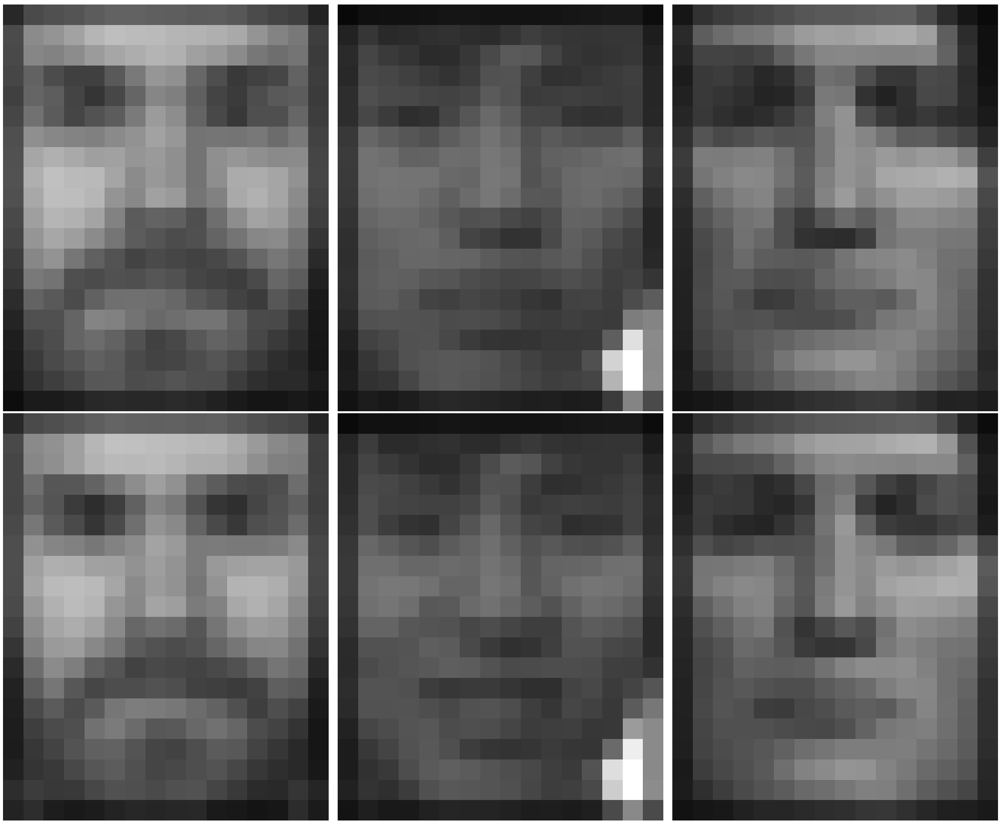

Figure 7: Three low-resolution test images (top) and the training images they were correctly matched against (bottom).

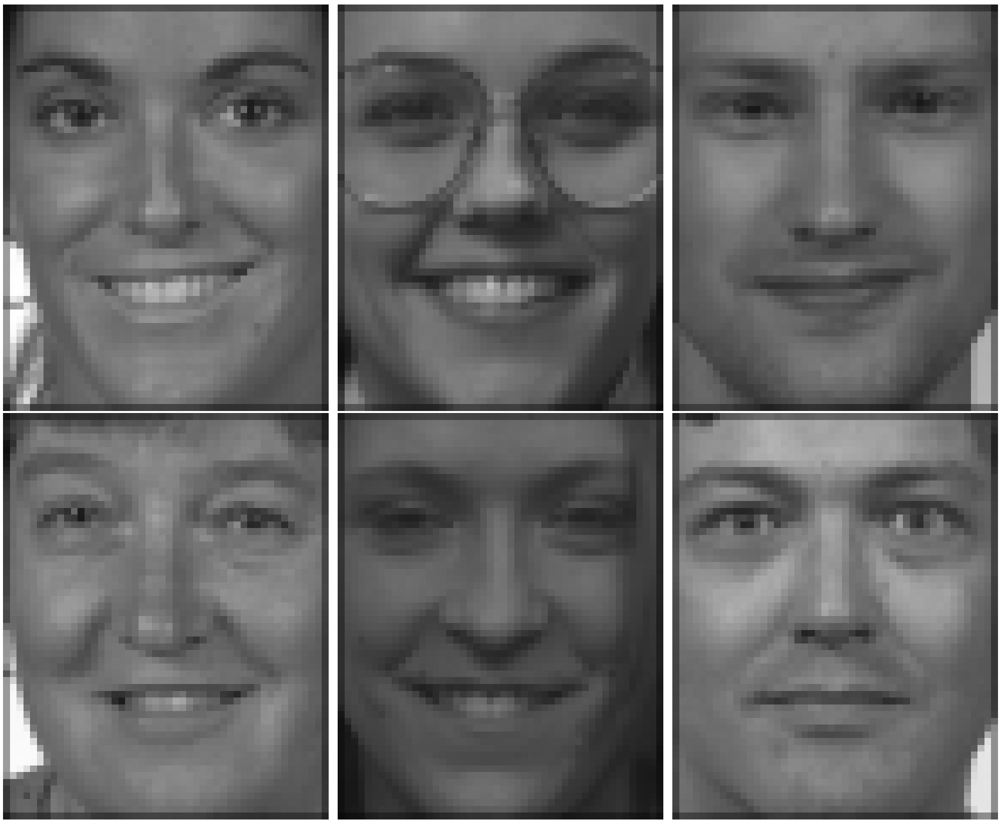

Figure 8: Three high-resolution test images (top) and the training images they were incorrectly matched against (bottom).

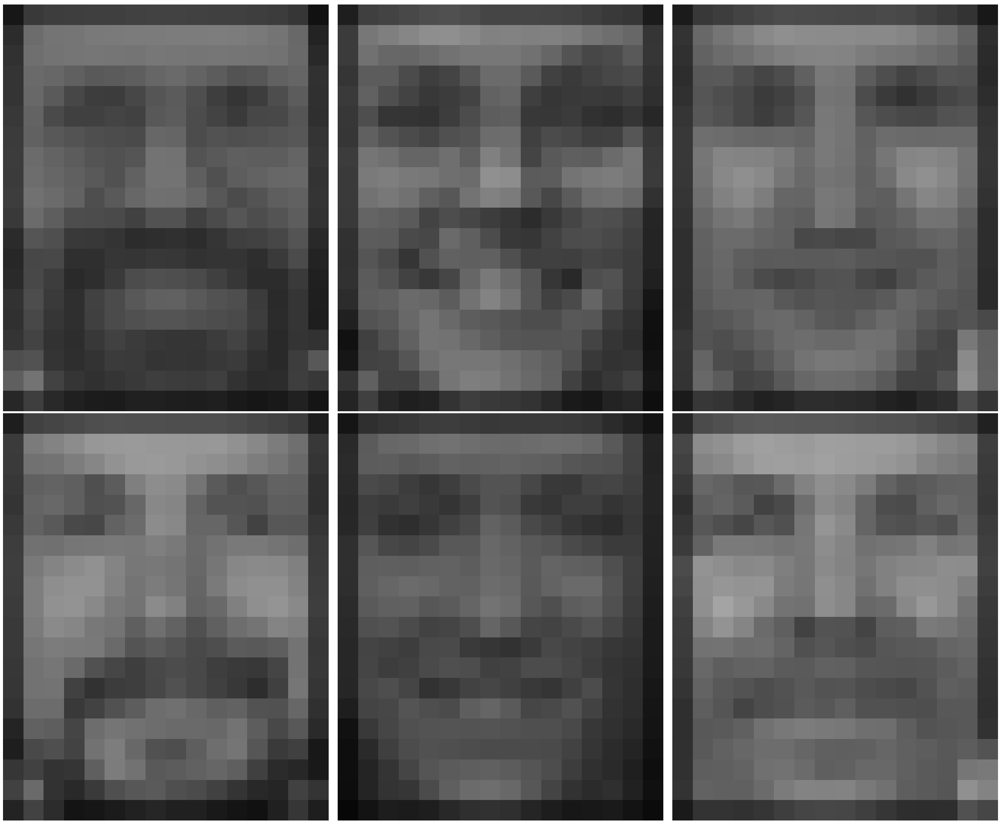

Figure 9: Three low-resolution test images (top) and the training images they were incorrectly matched against (bottom).