CS 383 - Machine Learning

Assignment 3 - Dimensionality Reduction

1 Theory Questions

1. Consider the following data:

$$\begin{bmatrix} -2 & 1 \\ -5 & -4 \\ -3 & 1 \\ 0 & 3 \\ -8 & 11 \\ -2 & 5 \\ 1 & 0 \\ 5 & -1 \\ -1 & -3 \\ 6 & 1 \end{bmatrix}$$

(a) Find the principle components of the data (you must show the math, including how you compute the eivenvectors and eigenvalues). Make sure you standardize the data first and that your principle components are normalized to be unit length. As for the amount of detail needed in your work imagine that you were working on paper with a basic calculator. Show me whatever you would be writing on that paper. (7pts).

$$\mu_1 = \frac{-2-5-3-8-2+1+5-1+6}{10} = -0.9$$

$$\mu_2 = \frac{1-4+1+3+11+5-1-3+1}{10} = 1.4$$

$$\sigma_1 = \sqrt{\frac{\Sigma(X_{i,1}-\mu_1)^2}{N-1}} = 4.23$$

$$\sigma_2 = \sqrt{\frac{\Sigma(X_{i,2}-\mu_2)^2}{N-1}} = 4.27$$

$$X_s = \frac{X-\mu}{\sigma}$$

$$w^* = argmax(Var(X_s w))$$

$$Var(X_s w) = \frac{(X_s w)^T (X_s w)}{N-1} = \frac{w^T X_s^T X_s w}{N-1}$$

$$\Sigma = Covariance = \frac{X_s^T X_s}{N-1} = \frac{1}{9}\begin{bmatrix} 9 & -3.67 \\ -3.67 & 9 \end{bmatrix} = \begin{bmatrix} 1 & -0.41 \\ -0.41 & 1 \end{bmatrix}$$

$$J(w) = w^T \Sigma w - \lambda(w^T w - 1)$$

$$w^* = argmax_w(J(w))$$

$$\frac{dJ}{dw} = 2\Sigma w - w\lambda w = 0$$

$$(\Sigma - \lambda I)w = 0 \longrightarrow \Sigma w = \lambda w$$

$$|\Sigma - \lambda I| = 0 \longrightarrow \begin{vmatrix} 1-\lambda & -0.41 \\ -0.41 & 1-\lambda \end{vmatrix} = 0 \longrightarrow (1-\lambda)(1-\lambda) - (-0.41)(-0.41) = 0$$

$$\lambda^2 - \lambda(1+1) + (1 * 1 - (-0.41)(-0.41)) = 0 \longrightarrow \lambda^2 - 2\lambda + 0.8319 = 0$$

$$\lambda_1 = 1.41 \quad \lambda_2 = 0.59$$

$$\begin{bmatrix} 1-\lambda_1 & -0.41 \\ -0.41 & 1-1-\lambda_1 \end{bmatrix} = \begin{bmatrix} -0.41 & -0.41 \\ -0.41 & -0.41 \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \longrightarrow w_1 = -w_2$$

$$\begin{bmatrix} 1-\lambda_2 & -0.41 \\ -0.41 & 1-1-\lambda_2 \end{bmatrix} = \begin{bmatrix} 0.41 & -0.41 \\ -0.41 & 0.41 \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \longrightarrow w_1 = w_2$$

$$v_1 = norm(\begin{bmatrix} -1 \\ 1 \end{bmatrix}) = \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} -0.707 \\ 0.707 \end{bmatrix} \quad v_2 = norm(\begin{bmatrix} 1 \\ 1 \end{bmatrix}) = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0.707 \\ 0.707 \end{bmatrix}$$

(b) Project the data onto the principal component corresponding to the largest eigenvalue found in the previous part (3pts).

$$Z = X_s w = \begin{bmatrix} -0.26 & -0.09 \\ -0.97 & -1.26 \\ -0.50 & -0.09 \\ 0.21 & 0.37 \\ -1.68 & 2.25 \\ -0.26 & 0.84 \\ 0.45 & -0.33 \\ 1.40 & -0.56 \\ -0.02 & -1.03 \\ 1.63 & -0.09 \end{bmatrix} \begin{bmatrix} -0.707 \\ 0.707 \end{bmatrix} = \begin{bmatrix} 0.12 \\ -0.21 \\ 0.29 \\ 0.11 \\ 2.78 \\ 0.78 \\ -0.55 \\ -1.38 \\ -0.71 \\ -1.22 \end{bmatrix}$$

2 Dimensionality Reduction via PCA

$$Sklearn\ KNN\ Accuracy : 23.2558\%$$

$$My\ KNN\ Accuracy : 23.2558\%$$

$$100D\ KNN\ Accuracy : 25.3876\%$$

$$Whitened\ 100D\ KNN\ Accuracy : 33.1395\%$$

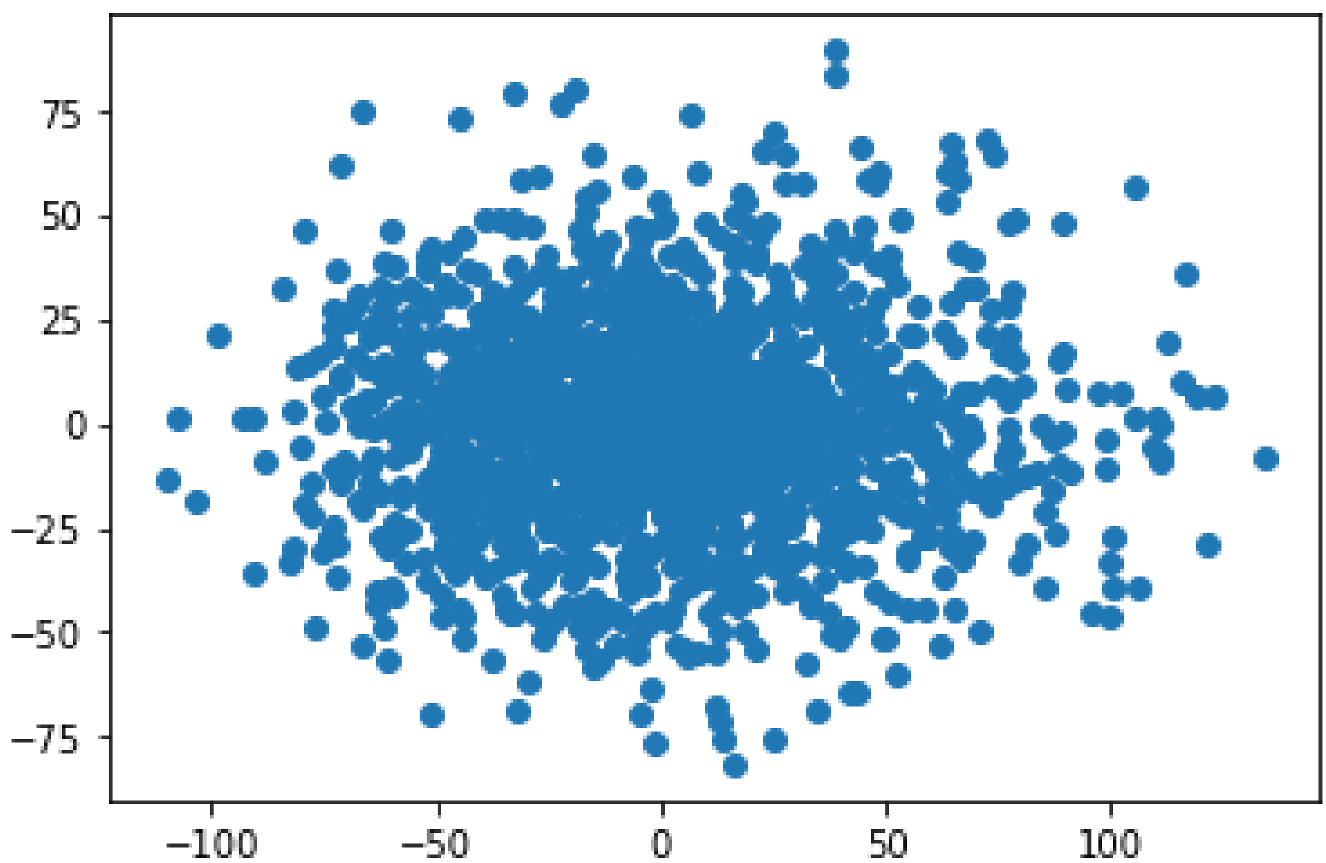


Figure 1: 2D PCA Projection of data

3 Eigenfaces

Figure 2: Max/Min PCA1: Wide Eyes + Large Mouth || $PCA2 : SmallEyes + SmallMouth$

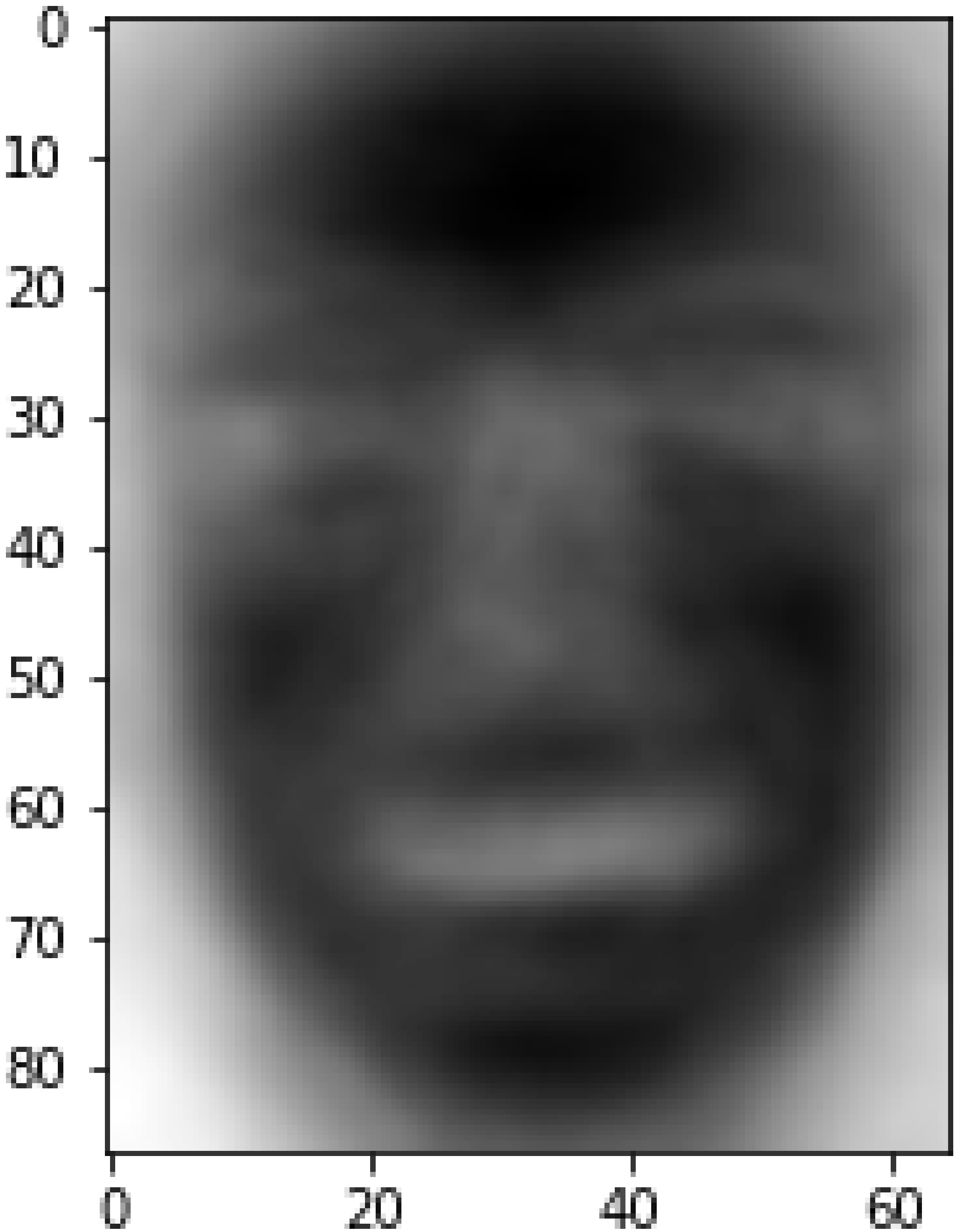
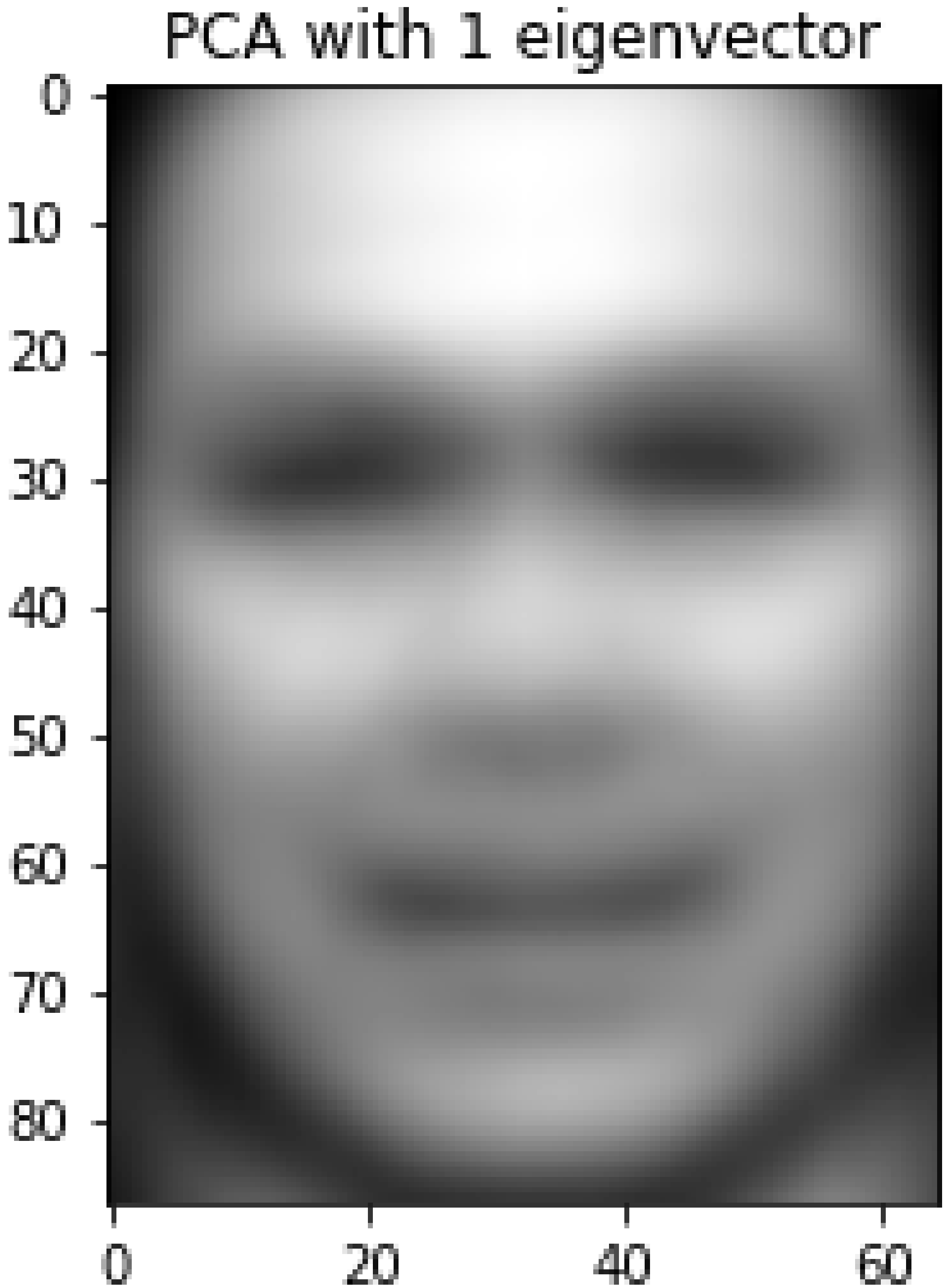
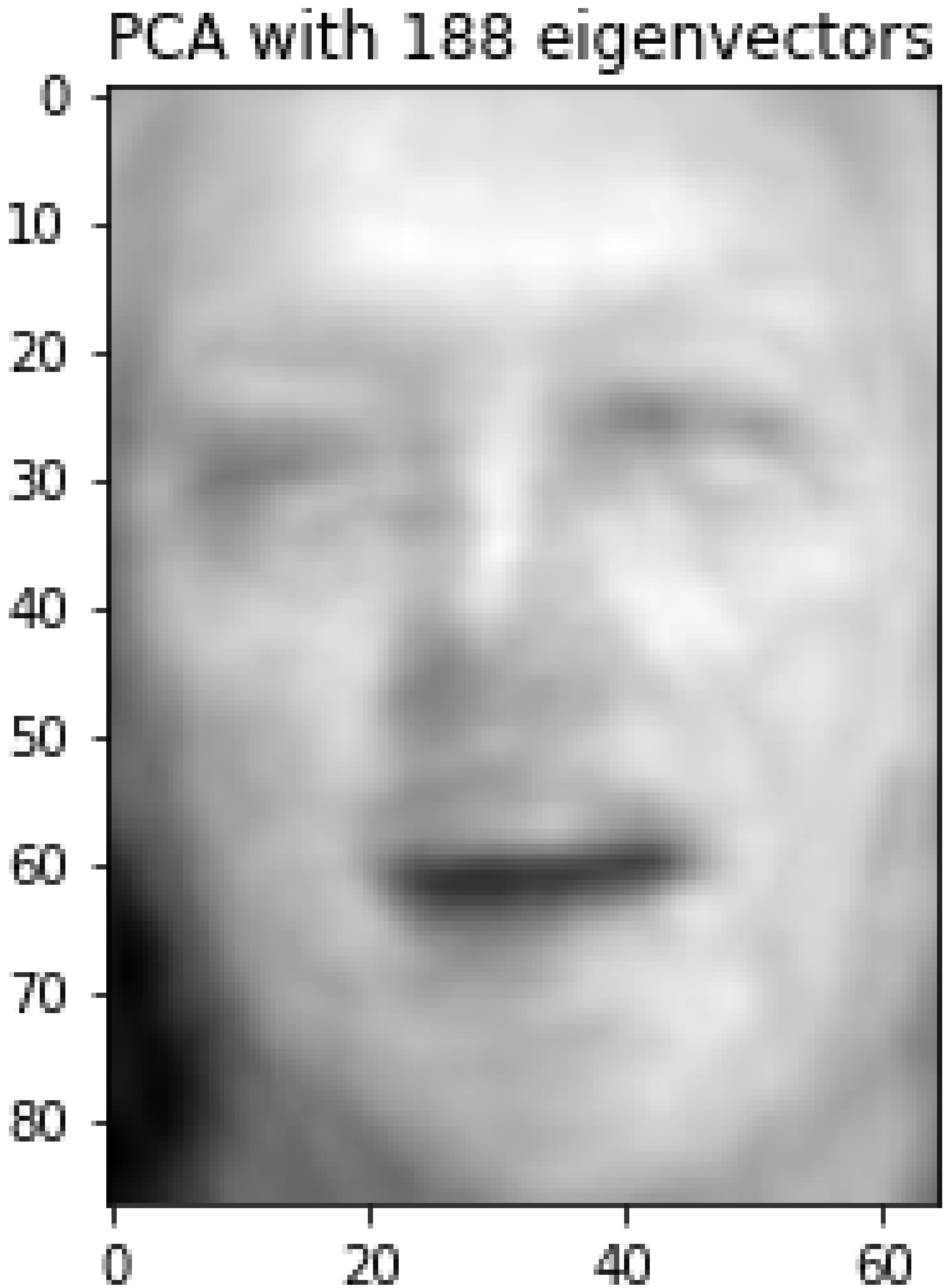


Figure 3: Most Important Component


PCA with 1 eigenvector
0
10
20
30
40
50
60
70
80
0
20
40
60


PCA with 188 eigenvectors
0
10
20
30
40
50
60
70
80
0
20
40
60