CS 4442B Assignment 2

Ositadinma Arimah

250 981 235

Department of Computer Science
Western University

2022-02-21

Ositadinma Arimah
CS 442B Assignment 2

1.

a. p(Water = cool|Play = yes) = 2/3
p(Water = cool|Play = no) = 1

b. p(Play = yes|Water = warm) = 2/3
p(Play = no|Water = warm) = 1/3

c. p(Play = yes|Humid = high) = 1
p(Play = yes|Humid = normal) = 1/2

d. p(Water = cool|Play = yes) = (1+2)/(3+2) = 3/5
p(Water = cool|Play = no) = (0+1)/(1+2) = 1/3

2.

a.

2a $K(x,z) = a_1 K_1(x,z) - a_2 K_2(x,z)$
where $a_1, a_2 > 0$ are real numbers

By Mercer's theorem (symmetric positive semi-definite function)

$$K(x,z) = a_1(z^T K_1 z) - a_2(z^T K_2 z)$$
$$= z^T(a_1 K_1)z - z^T(a_2 K_2)z$$
$$\geq 0 \iff z^T(a_1 K_1)z \geq z^T(a_2 K_2)z$$

$\therefore$ b/c the statement is not always true, K is not a valid Kernel

b.

2b K is symmetric due to scalar multipiclation commutitivity

$\therefore$ from Kernel definition

K_1 is a Kernel $\rightarrow$ $\exists\, \phi(1)$ so $K_1(x, z) = ((\phi(1)(x))^T \; (\phi(1)(z)))$

K_2 is a Kevel $\rightarrow$ $\exists\, \phi(2)$ so $K_2(x, z) = ((\phi(1)(x))^T \; (\phi(1)(z)))$

$$K(x, z) = \sum_i \phi_i(1)(z) \sum_j \phi_j(2)(x)\, \phi_j(2)(z)$$

$$= \sum_i \sum_j \phi_i(1)(x)\, \phi_i^{(2)}(z)\, \phi_i^{(2)}(x)\, \phi_j^{(2)}(z)$$

$$= \sum_i \sum_j (\phi_i^{(1)}(x)\, \phi_j^{(2)}(x))\, (\phi_i^{(1)}(z)\, \phi_j^{(2)}(z))$$

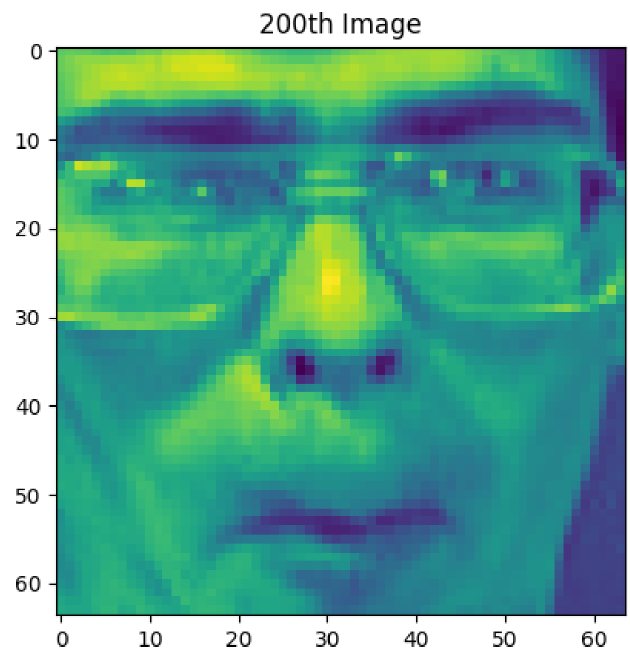
define $\alpha(\cdot) = \phi(1)(\cdot)\, \phi^{(2)}(\cdot)$

$$= \sum_i \sum_j \alpha_{i,j}(x)\, \alpha_{i,j}(z)$$

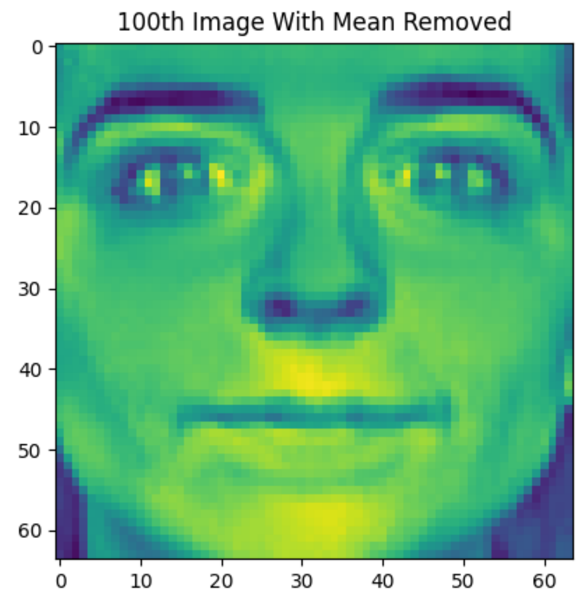
$\therefore$ K is a valid Kernel

3.

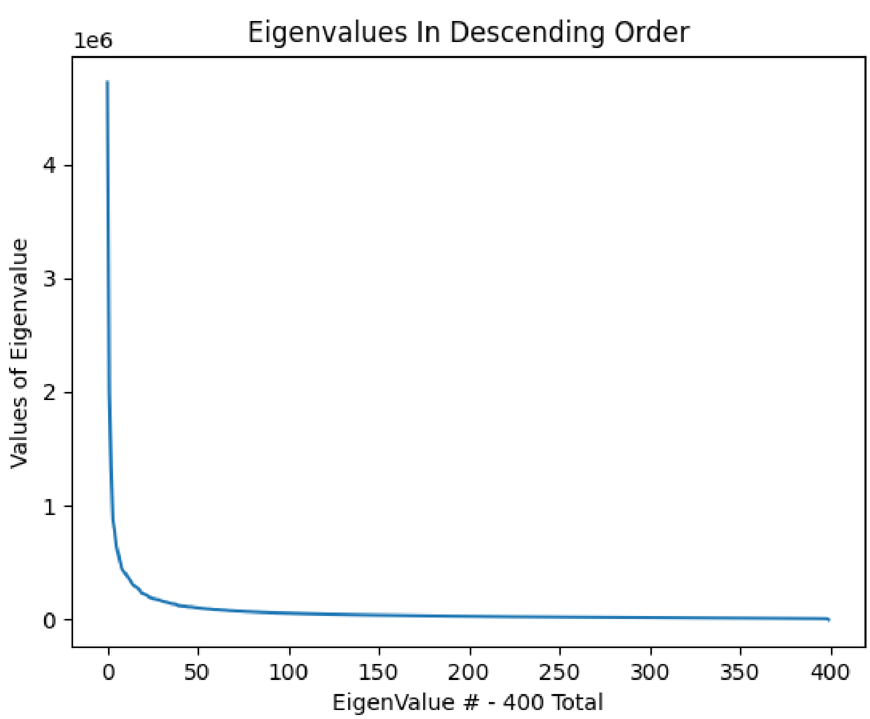
a. Display the 200th image.



b. Remove the mean of the images, and then display the 100th image.



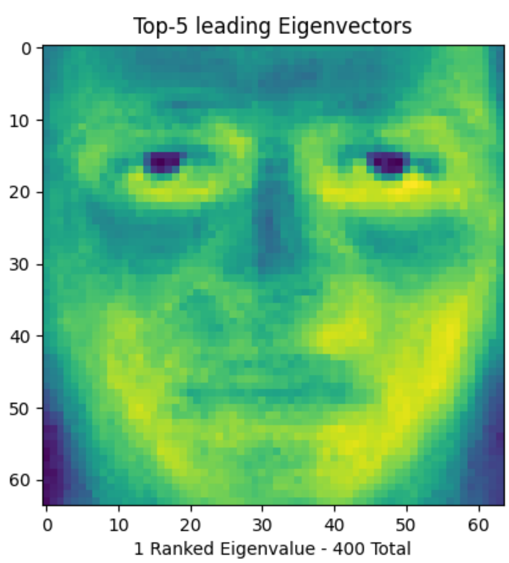
c. Perform PCA on the mean-centered data matrix.

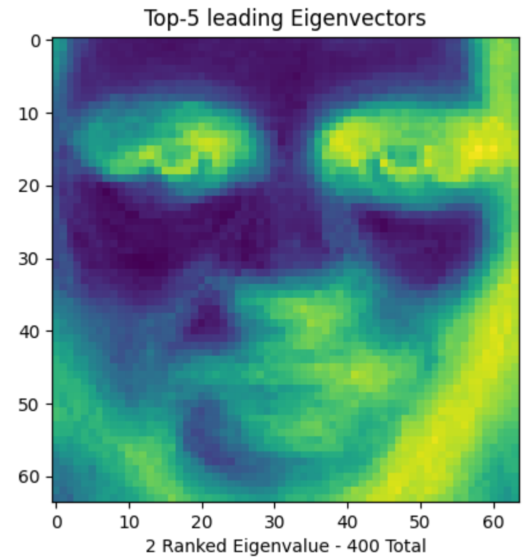


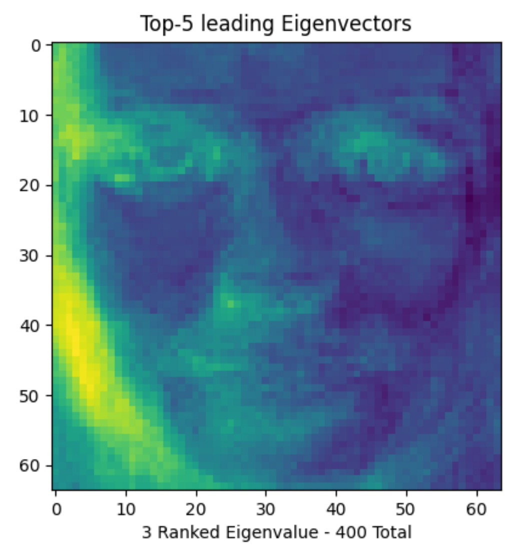
d. 400th eigenvalue is 0 means that the null space is nontrivial; in other terms, the corresponding dimension, otherwise known as the component, has a variance of 0 and, consequently does not exist. Furthermore, each eigenvector is assigned an eigenvalue whose magnitude determines the data's variability that is explained by its eigenvector. In turn, the corresponding component for that eigenvalue does not because it is assigned 0.

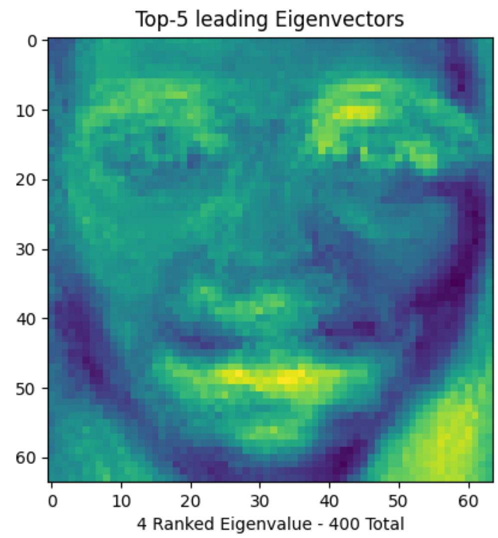
e. Following the graph results, it can be derived that the first twenty eigenvalues contribute to most of the variance in the data. Therefore, I used 83% for the cut-off point as this number is the value is a few deviations above the mean for a stand distribution, with only 35/400 components being displayed.

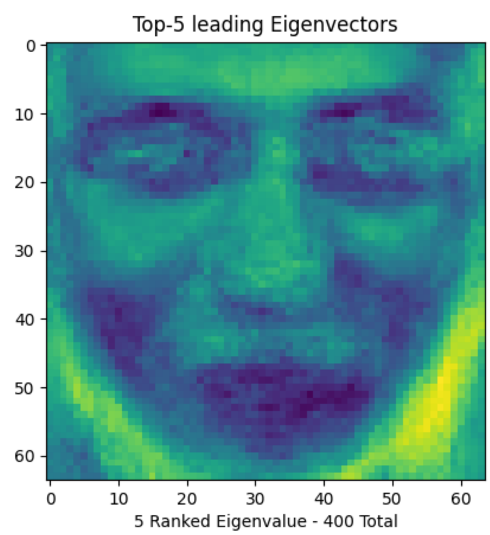
f. Display the top-5 leading eigenvectors (corresponding to the top-5 largest eigenvalues) in 5 figures.



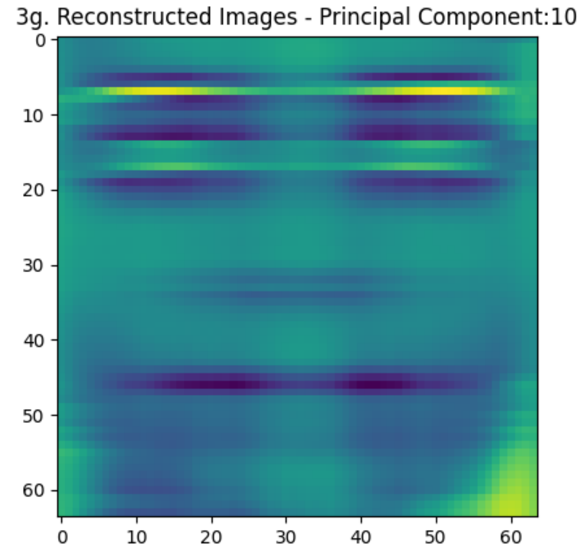

Top-5 leading Eigenvectors
0
10
20
30
40
50
60
0
10
20
30
40
50
60
2 Ranked Eigenvalue - 400 Total

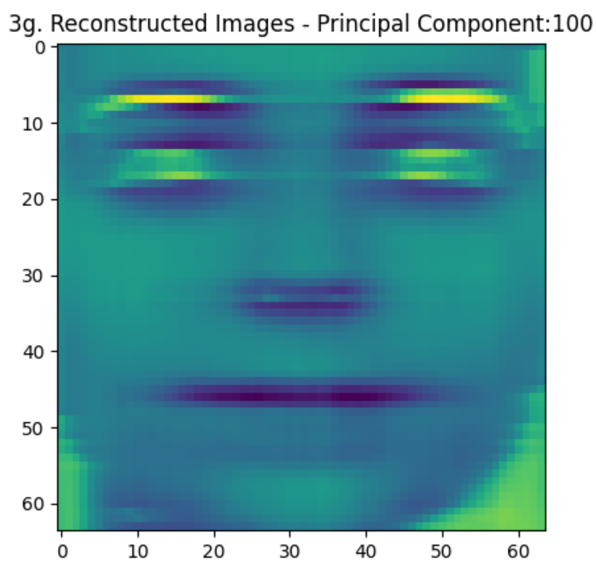

Top-5 leading Eigenvectors
0
10
20
30
40
50
60
0
10
20
30
40
50
60
3 Ranked Eigenvalue - 400 Total


Top-5 leading Eigenvectors
4 Ranked Eigenvalue - 400 Total

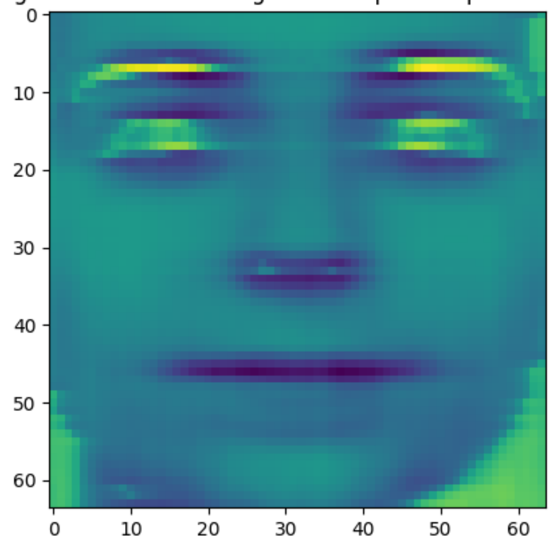

Top-5 leading Eigenvectors
5 Ranked Eigenvalue - 400 Total

g. Display, respectively, the reconstructed 100th images using 10, 100, 200, and 399 principal components.





3g. Reconstructed Images - Principal Component:200



3g. Reconstructed Images - Principal Component:399