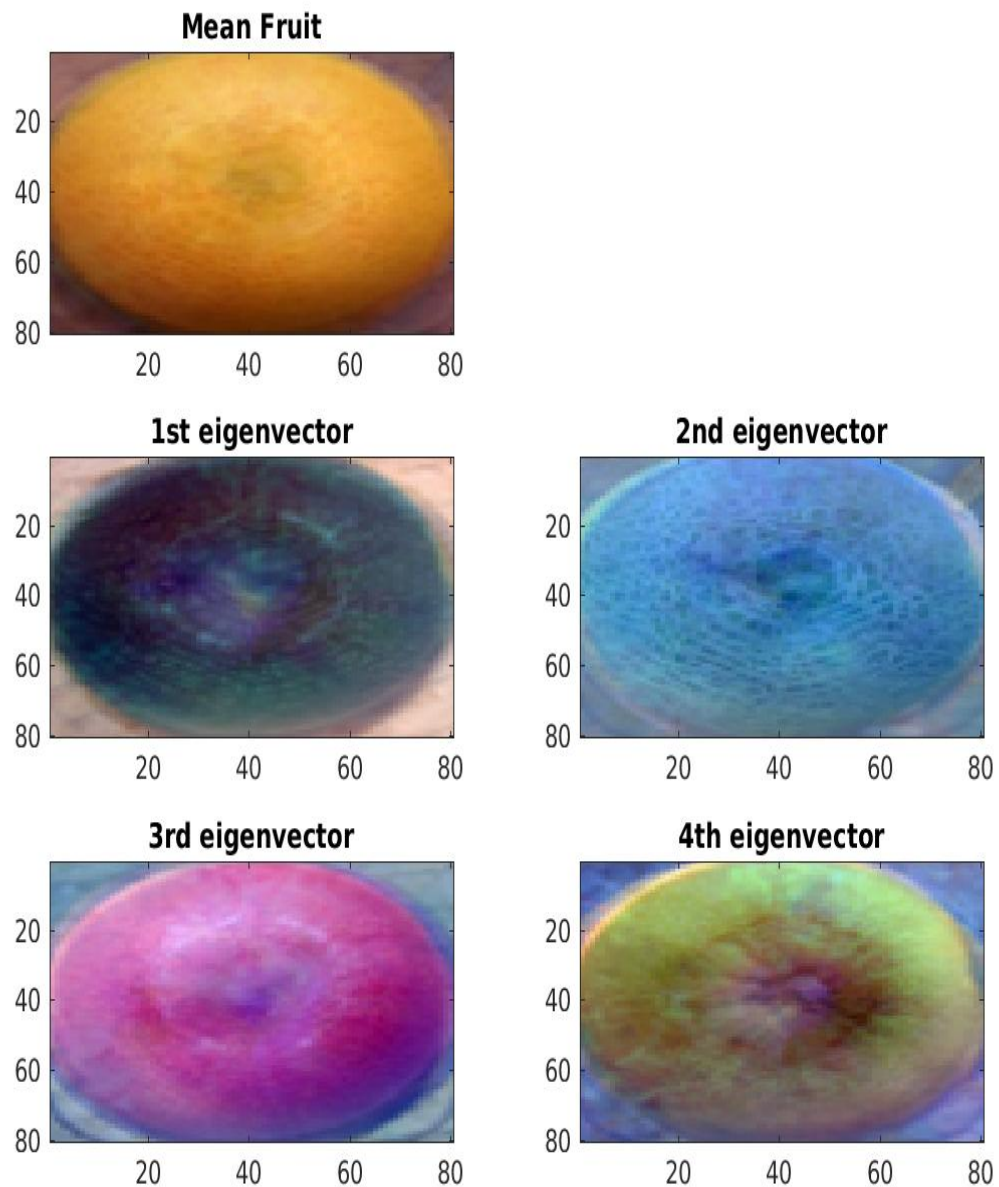


Question 5

5.1 : Part A

For the collection of images, the Mean Vectors, Covariance Matrices, eigenvalues, and corresponding eigenvectors have been calculated and stored in 'results.mat' (present in the results directory). More on it's loading in Sub-Section 5.4 (Usage of code). The mean and the first four eigenvectors have also been plotted, and the plots follow:



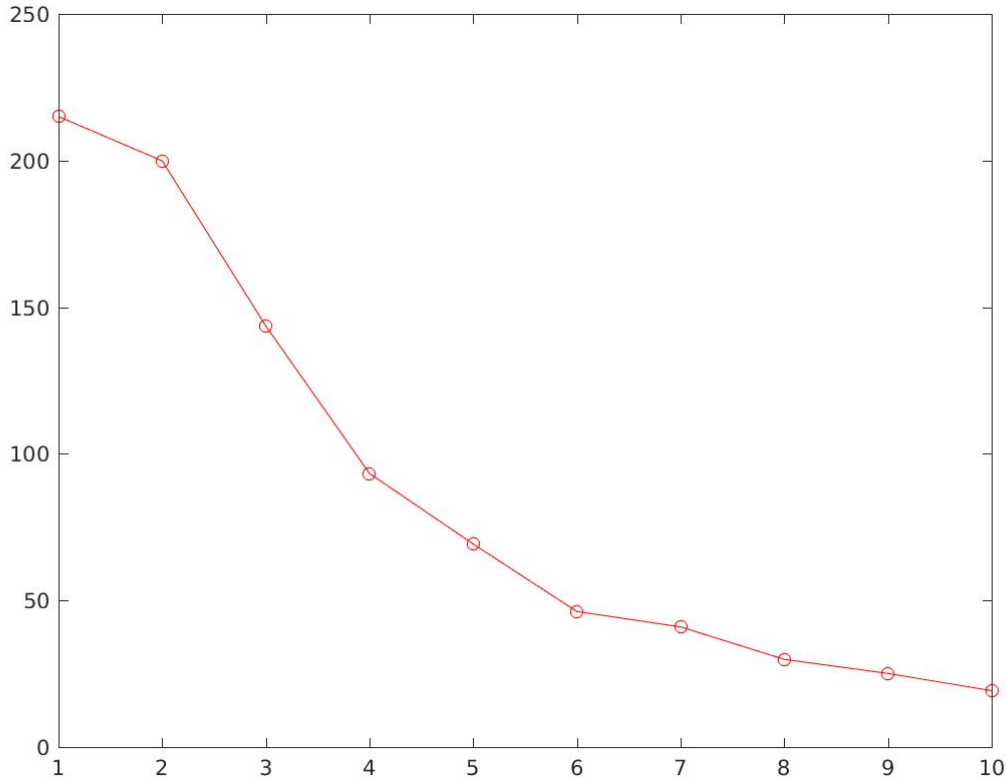


Figure 1: Plot of Largest 10 eigenValues vs components

5.2 : Part B

We have been given 16 images to train, with each image having $80 \times 80 \times 3 = 19200$ variables. Let us call each of these images an instance. Let us denote a general variable as follows:

$$X_i^j = i^{th} \text{ variable of the } j^{th} \text{ image instance} \quad \forall i \leq 19200, j \leq 16$$

Let our eigenvectors corresponding to the top 4 eigenvalues be $\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3$ and \mathbf{V}_4 . Let the mean Vector over all fruits be $\boldsymbol{\mu}$ and let \mathbf{X}^i be a 19200×1 vector of the i th image instance.

Note that all four eigenvectors have zero inner product with respect to each other, since the Covariance matrix is SPD (Symetric Positive Definite)

Hence $\forall i \leq 16, a_i \in R^+$

$$\mathbf{X}^i = \boldsymbol{\mu} + a_1 \mathbf{V}_1 + a_2 \mathbf{V}_2 + a_3 \mathbf{V}_3 + a_4 \mathbf{V}_4$$

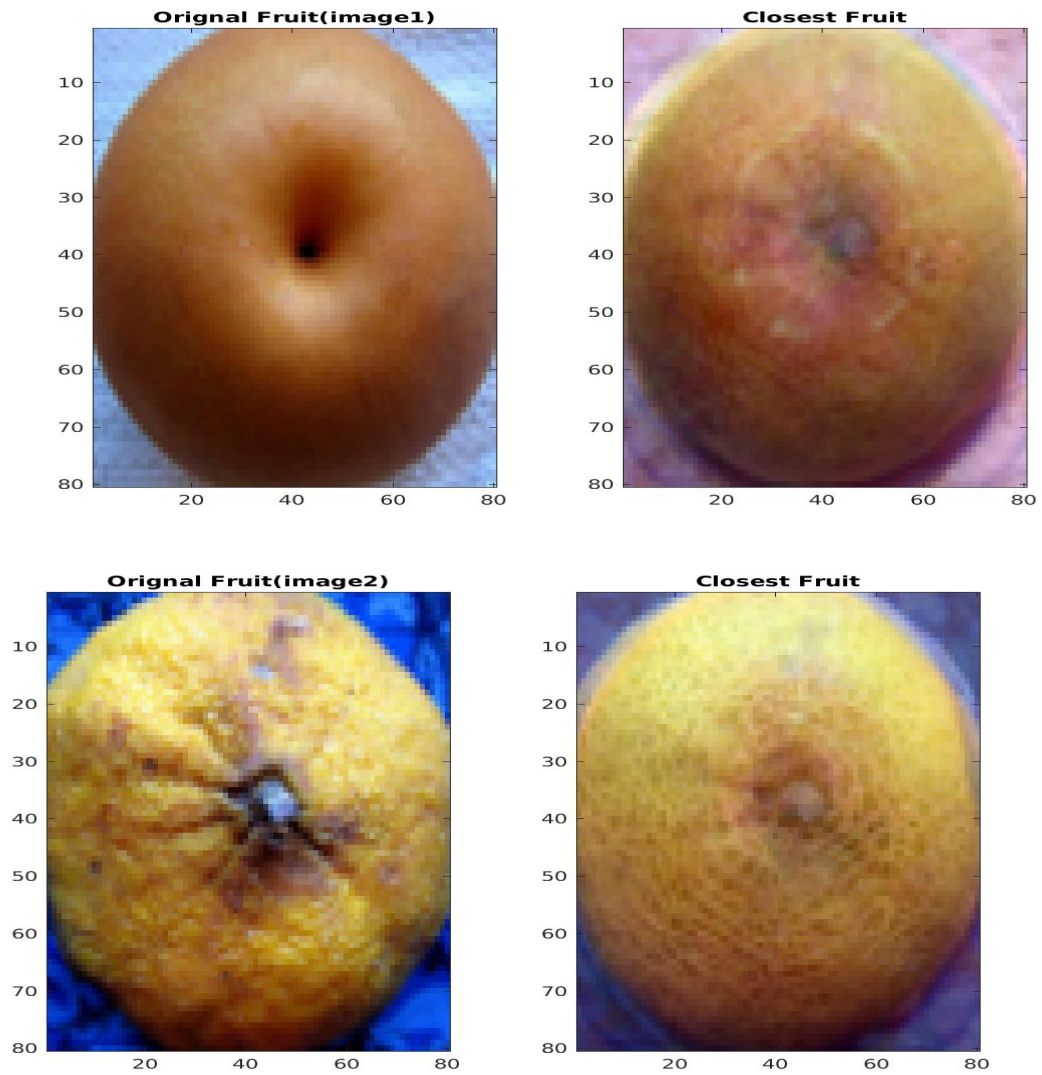
We are keeping the coefficient of $\boldsymbol{\mu}$ as 1, since it is the mean matrix. Multiplying the equation with the transpose of each eigenvector, since their inner product

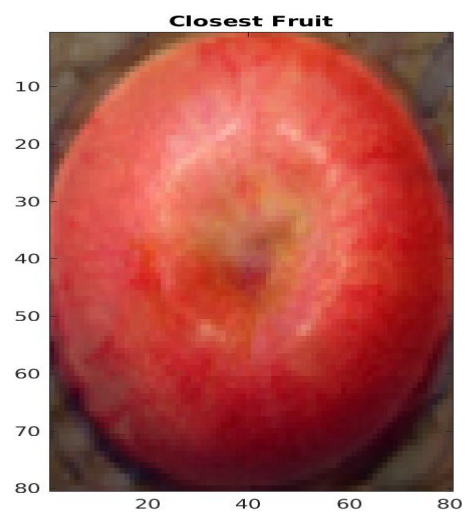
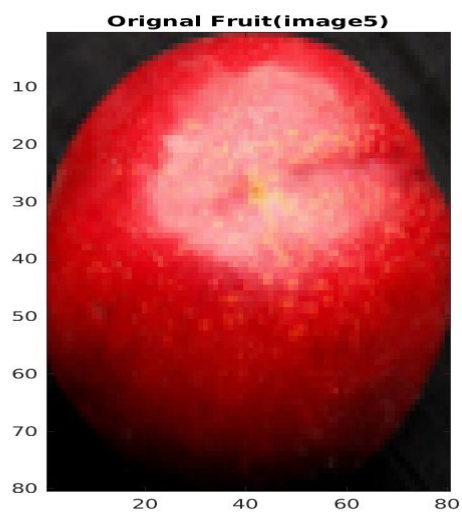
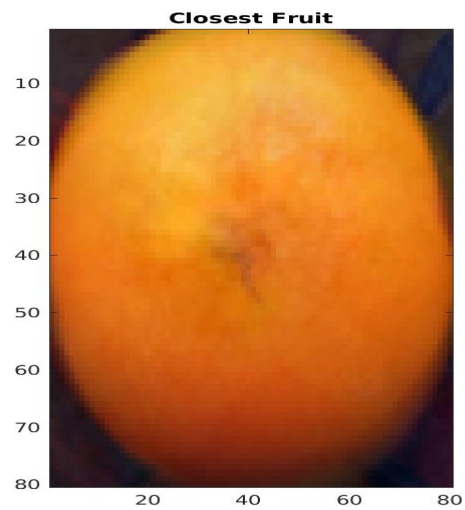
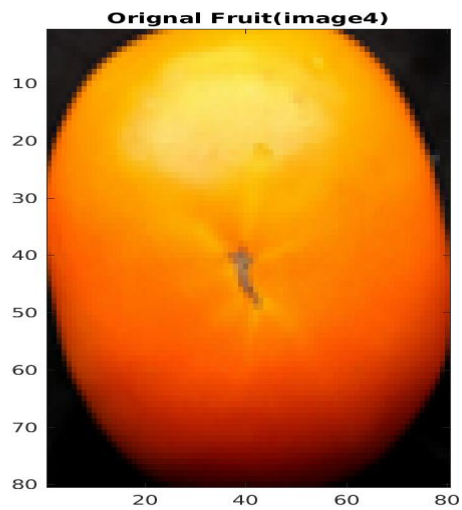
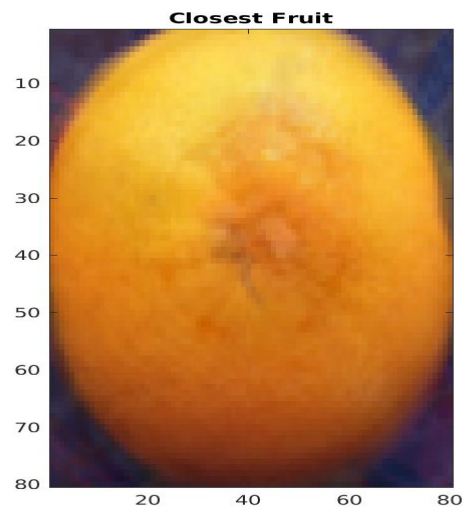
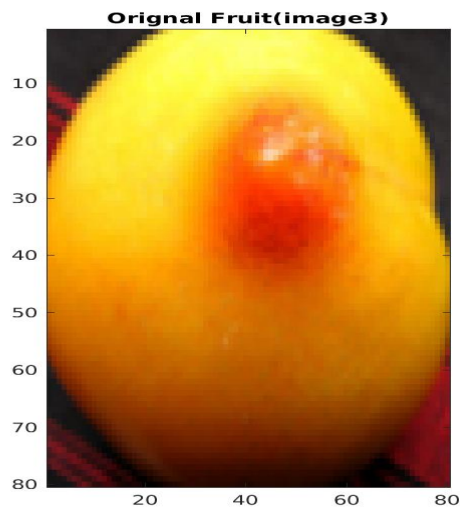
is zero.

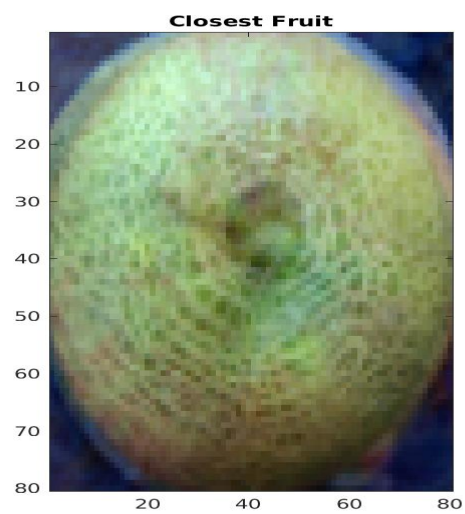
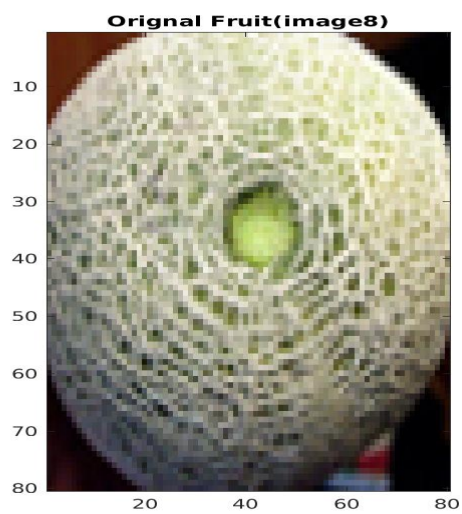
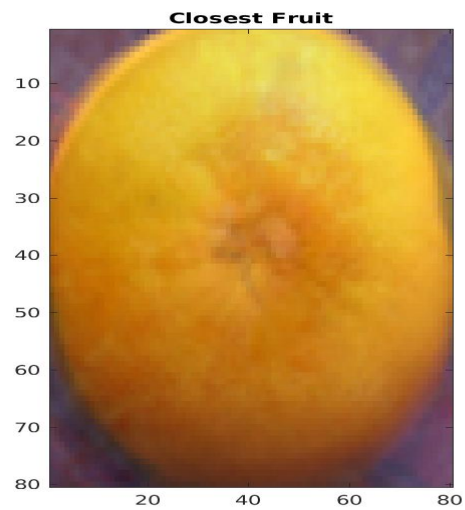
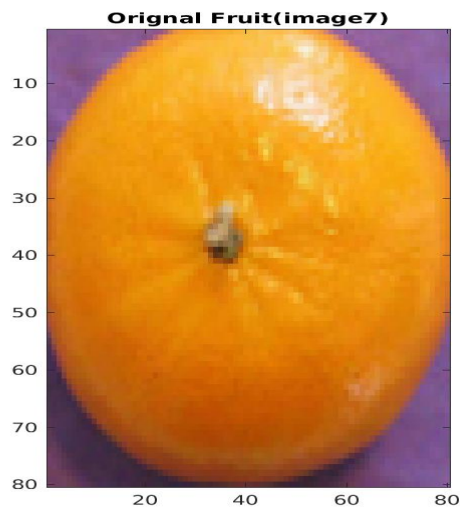
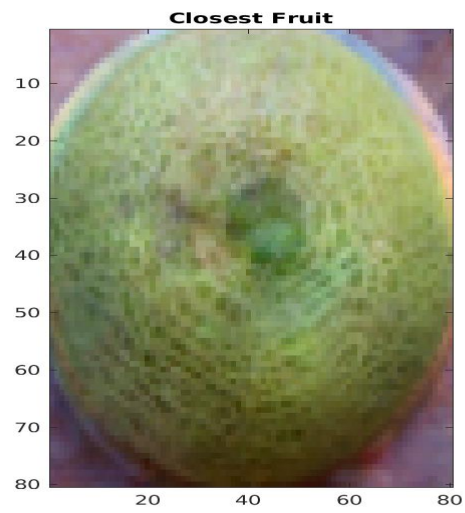
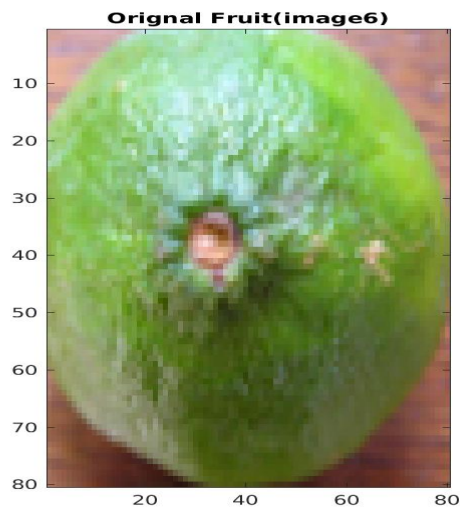
$$\begin{aligned}
 X^i &= \mu + \sum_{i=0}^4 a_i V_i \\
 X^i \cdot V_j^T &= \mu \cdot V_j^T + \sum_{i=0}^4 a_i V_i \cdot V_j^T \\
 X^i \cdot V_j^T &= \mu \cdot V_j^T + \sum_{i \neq j} a_i V_i \cdot V_j^T + a_j \langle V_j, V_j \rangle \\
 a_j &= \frac{X^i \cdot V_j^T - \mu \cdot V_j^T}{\langle V_j, V_j \rangle} \quad \forall j \in \{1, 2, 3, 4\}
 \end{aligned}$$

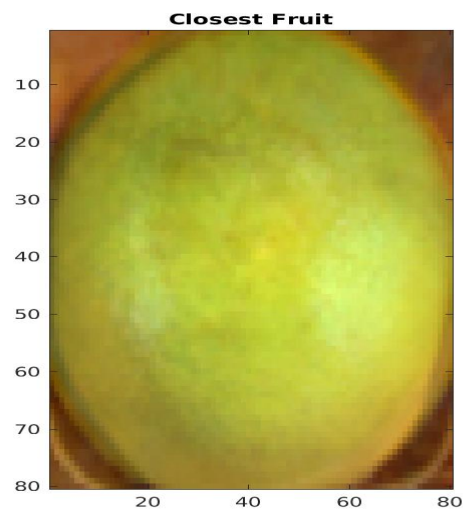
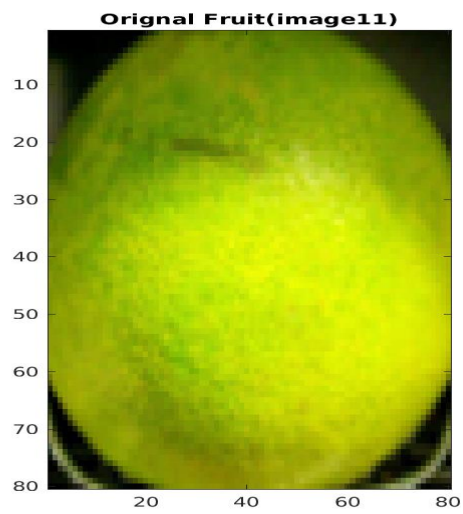
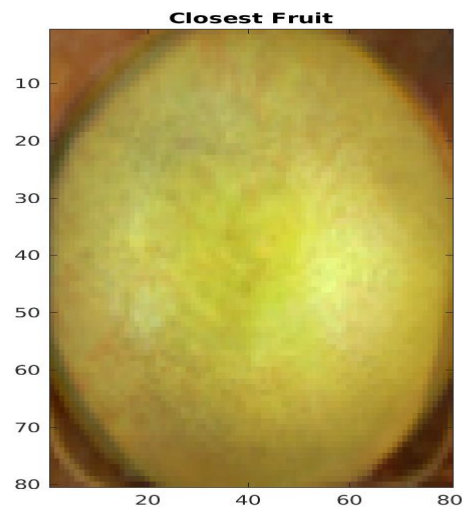
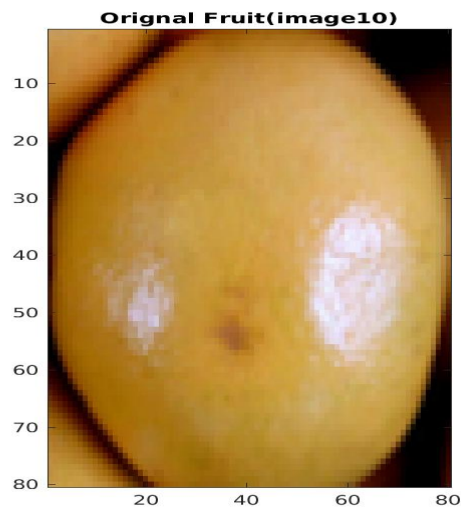
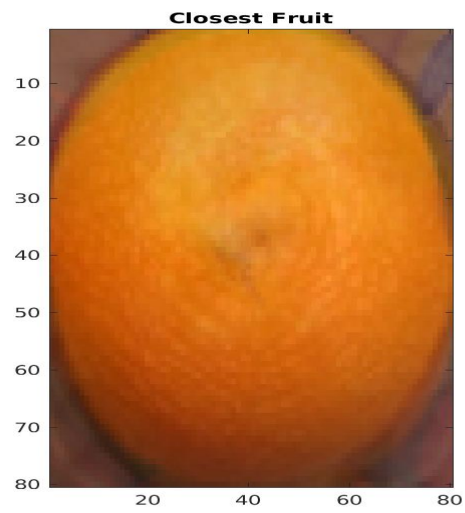
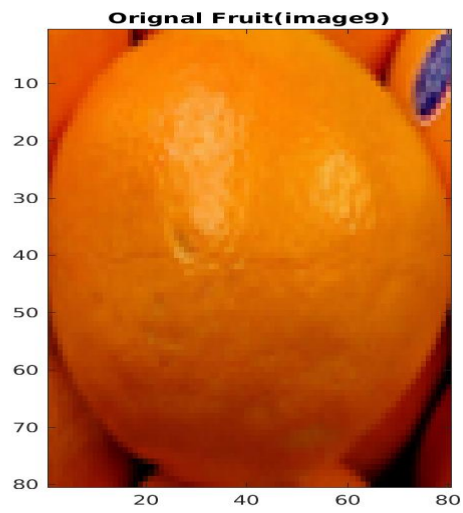
Hence we can find the closest representation of every fruit as a linear combination of the mean and the first four eigen vectors.

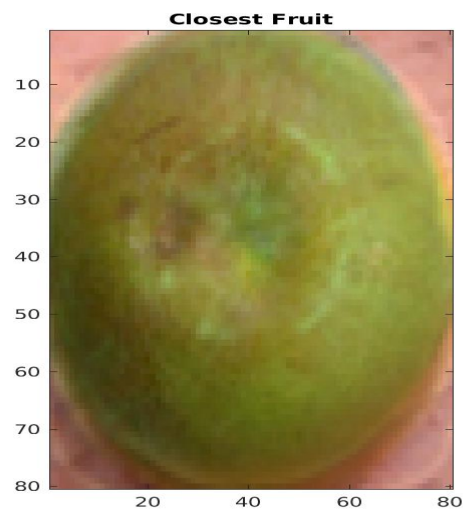
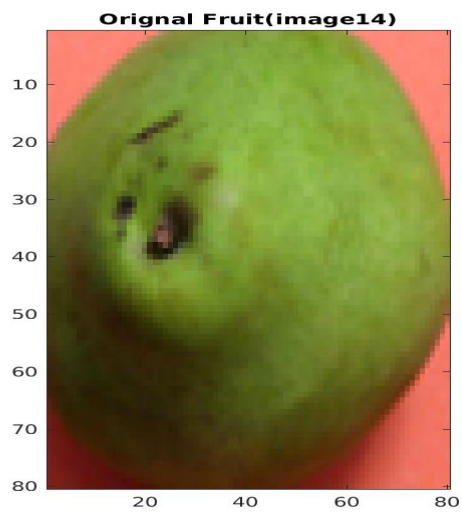
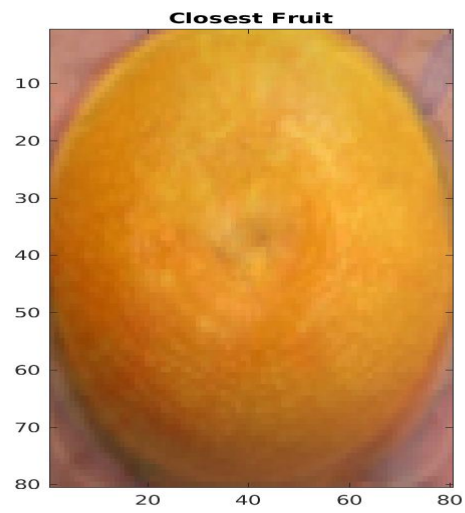
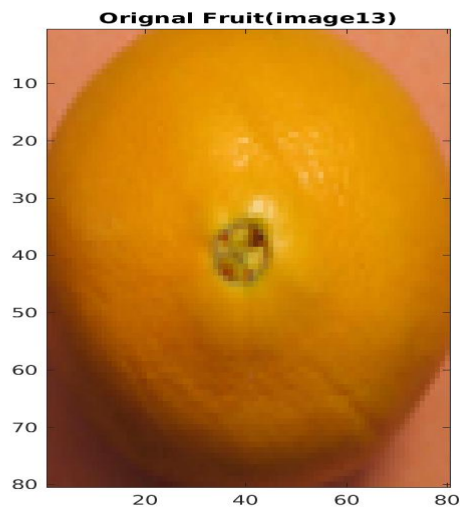
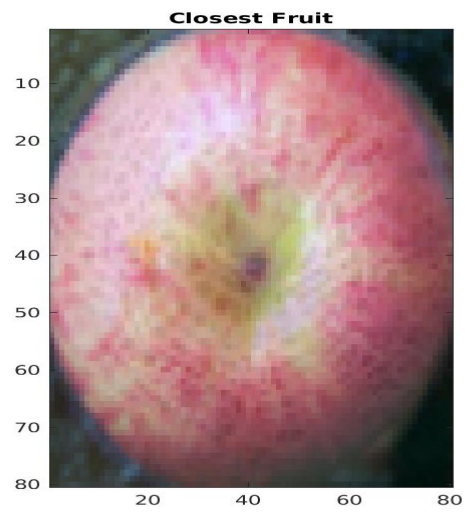
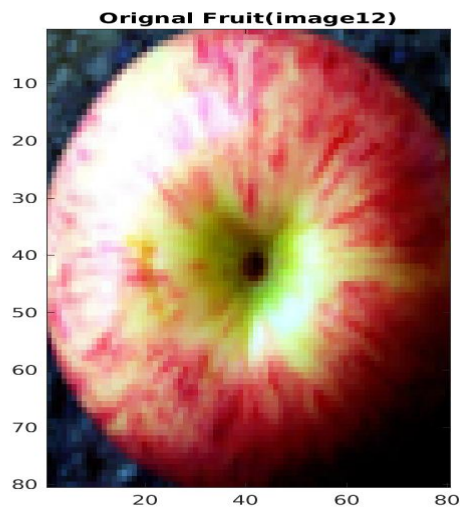
The fruits, along with their closest representation follow:

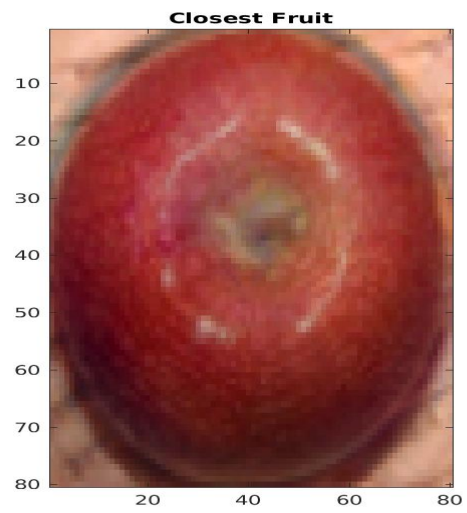
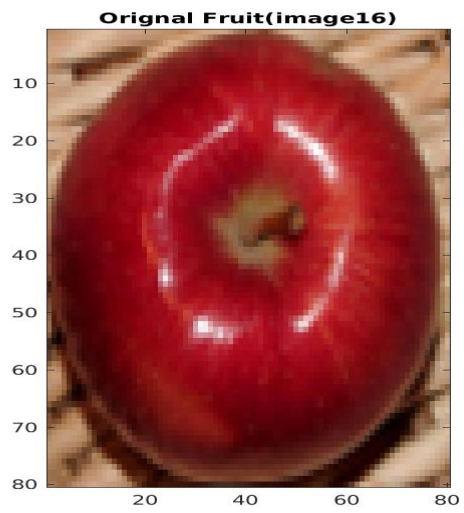
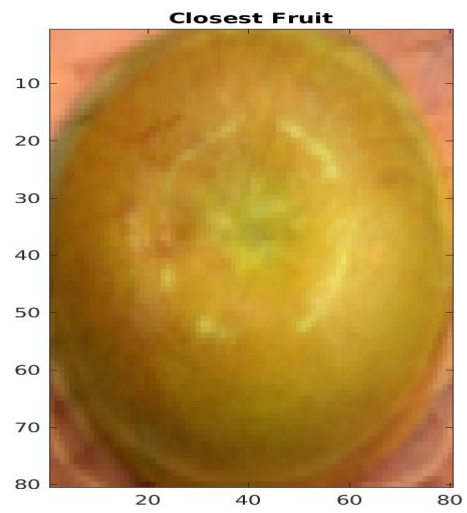
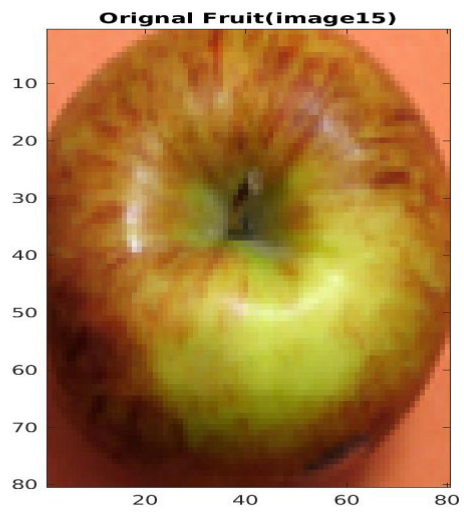












5.3 : Part C

Now, for generating a new representation of a fruit, let us recall the equation that we started with in Part B.

$$\mathbf{X}^i = \boldsymbol{\mu} + a_1 \mathbf{V}_1 + a_2 \mathbf{V}_2 + a_3 \mathbf{V}_3 + a_4 \mathbf{V}_4$$

In this case, we need to draw random values of a_i for generating a new fruit.

We know that the **magnitude** by which any particular feature varies along an eigenvector is **proportional** to it's corresponding **eigenvalue**.

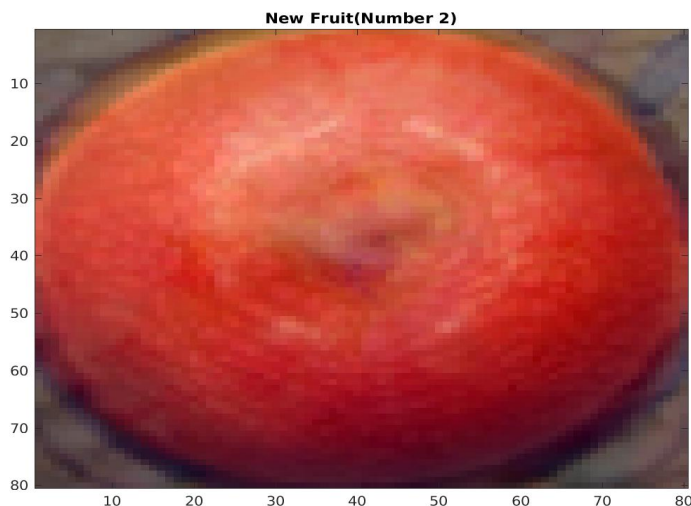
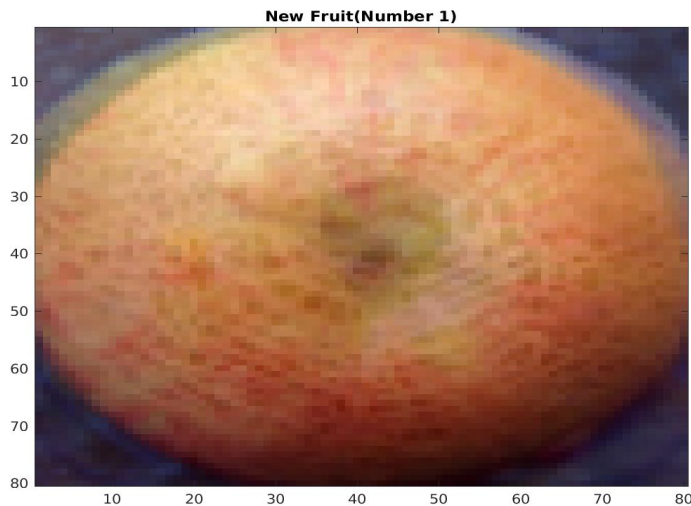
Since we are taking the linear combinations of the eigenvectors into the mean to form another image, it would be most suitable if all the a_i would be the **corresponding eigenvalues**

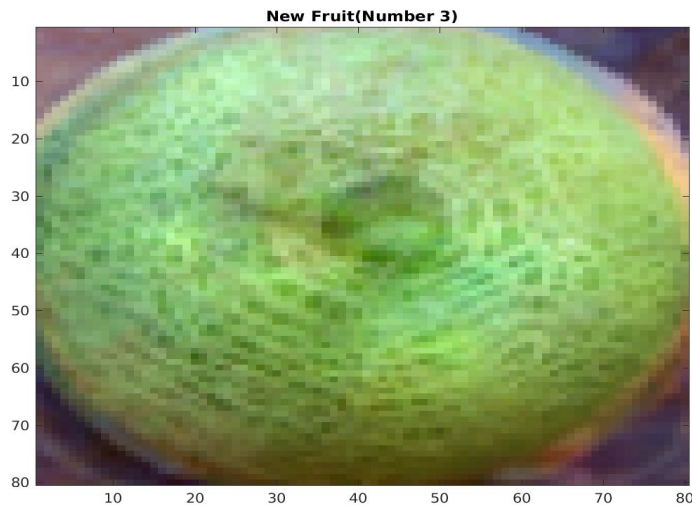
Hence, we have:

$$a_i \in G(\mu = 0, \sigma^2 = EigenValue)$$

This was the algorithm that I used to generate new instances of fruits.

The plots for the new generated images are as follows.:





5.4 : Usage of Code

The following are the instructions for the usage of the code:

- Load the code present in 'submission/code/q5/q5.m' .
- In the same directory are functions implemented like myMean, myCov which return the mean and covariance of appropriate matrices. Functions like myScale, myeigs scale the data and return sorted eigenvalues and eigenvectors, respectively.
- Simply run the code in 'q5.m' and this will automatically create the required plots. (Requires a lot of time)
- Lines 39, 44, 63, 81 (commented by default) have a code to save jpg files of the respective plots. Comment/Uncomment these lines appropriately according to need.
- The results.mat present in 'results/mat' contains all the required results for the partA stored:

Mean Vector: Stored as a 19200x1 vector by the name MeanFruit

Covariance Matrix: Was not be able to store and upload this file to moodle, since it's a 19200x19200 matrix, and has a whooping size of 2.6 GB.

Largest 10 EigenVectors: Stored as a 19200x10 matrix by the name EigenVectors, with each column as a eigenvector.

Corresponding Eigenvalues: Stored as a 10x1 vector by the name EigenValues, where the ith cell has the ith largest eigenvalue 'i'