CS215 Assignment2 Problem 6

Parth Pujari and Anish Kulkarni

October 2022

1 PCA on Fruits

A data set consisting of images of fruits of size 80×80 pixels with 3 color channels red (R), green (G), and blue(B), i.e., a $80 \times 80 \times 3$ array is given.
The code first resizes the given images into arrays of length 19200
Then we apply PCA on the 16 images.
The covariance matrix found has eigenvalues depicted in the image shown below.

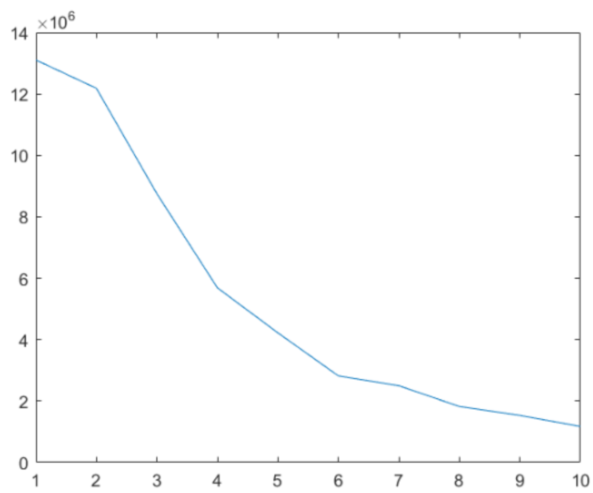


Figure 1: Eigen Values

Further, the mean and the four largest eigen vectors (**Scaled to the mean**) correspond to the following fruit:

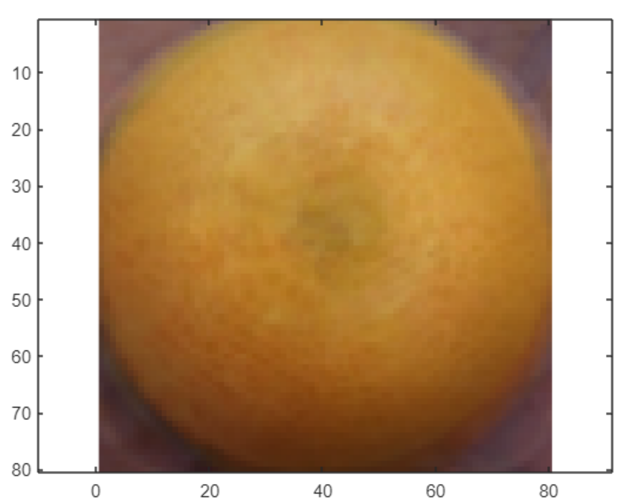
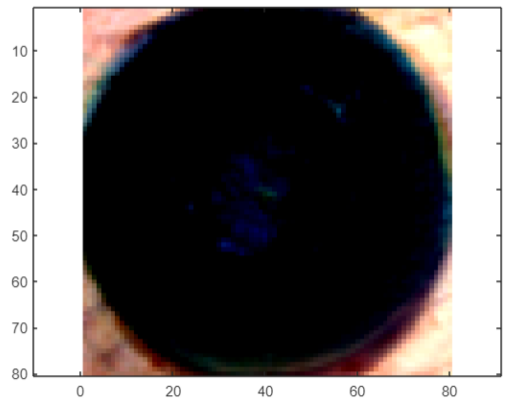


Figure 2: Mean Fruit

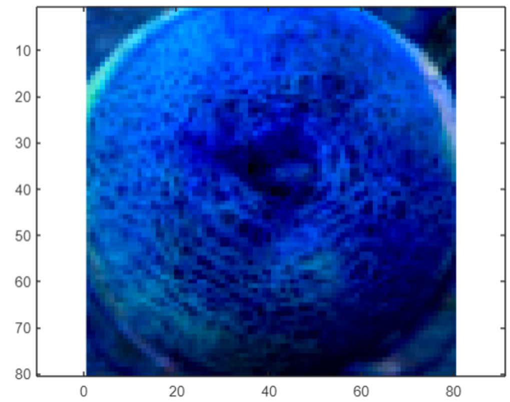
2 Finding the Closest Image Representation

The top four eigenvectors found from the Principal Component Analysis are to be used to generate images having the closest representation to the given 16 images.

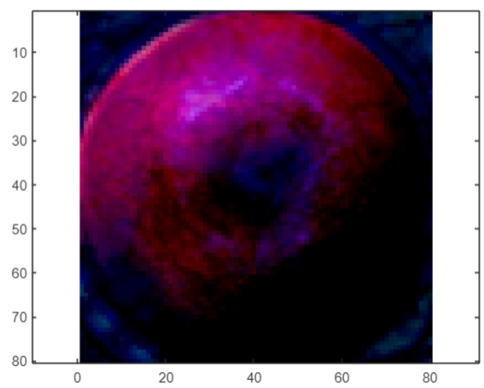
To do so we use the following theorem:



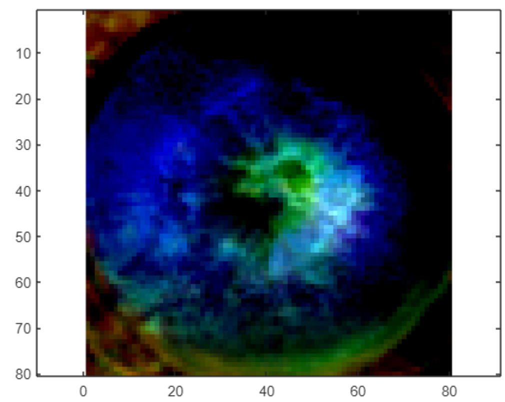
(a) Fruit for Eigen Value 1



(b) Fruit for Eigen Value2



(a) Fruit for Eigen Value 3



(b) Fruit for Eigen Value4

Given a set of linearly independent vectors $\vec{v_1}, \vec{v_2}, ...\vec{v_n}$ and a vector $\vec{b}$, the vector $\vec{u}$, which is a linear combination of the given vectors, such that $||\vec{u} - \vec{b}||_2$ is minimised is:

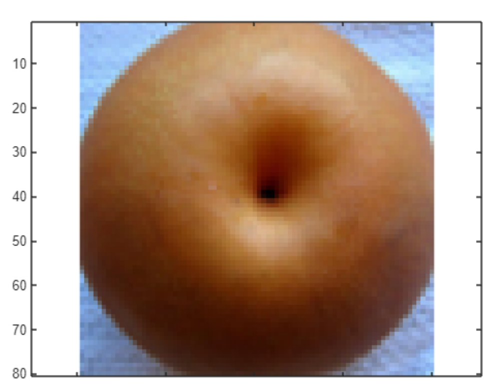
$$\vec{u} = (\hat{u_1}.\vec{b})\vec{v_1} + (\hat{u_2}.\vec{b})\vec{v_2} + ... + (\hat{u_n}.\vec{b})\vec{v_n}$$

where,

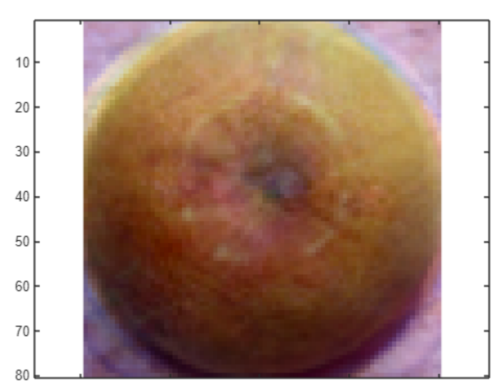
$$\hat{u_1}, \hat{u_2}, ...\hat{u_n}$$

are the unit vectors corresponding to $\vec{v_1}, \vec{v_2}...\vec{v_n}$
Here, the vectors $\vec{v_1}, \vec{v_2}$, etc.. correspond to the 4 eigen vectors and $b \vec{-} \mu$ (where μ is the mean) corresponds to the vector shaped image which we have to replicate.
Adding the mean to the generated vector gives us the required vector.

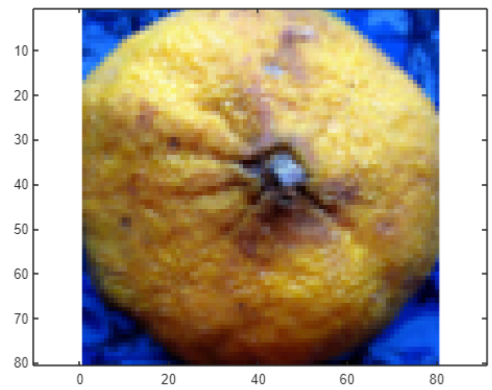
The following are the images (actual and generated) side by side



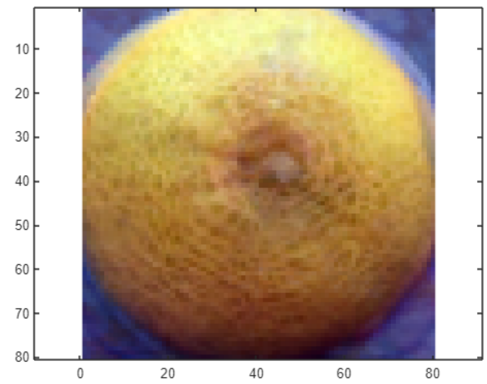
(a) Actual Image



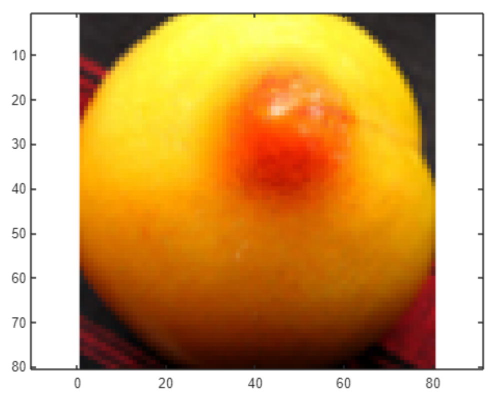
(b) Generated



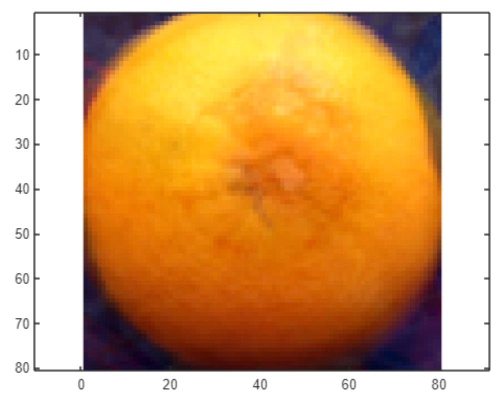
(a) Actual Image



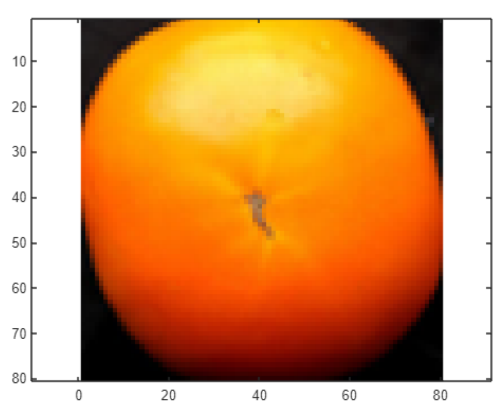
(b) Generated



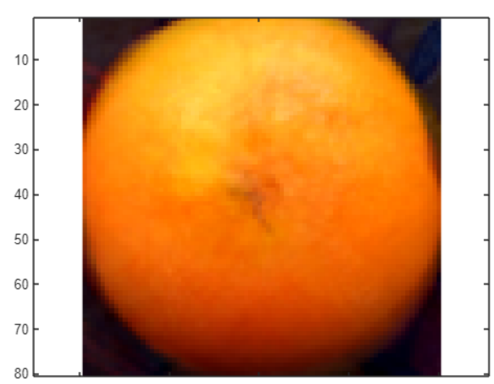
(a) Actual Image



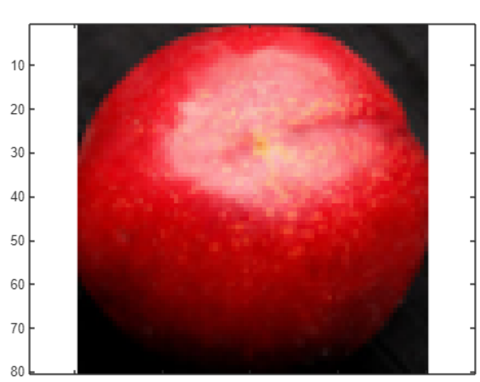
(b) Generated



(a) Actual Image

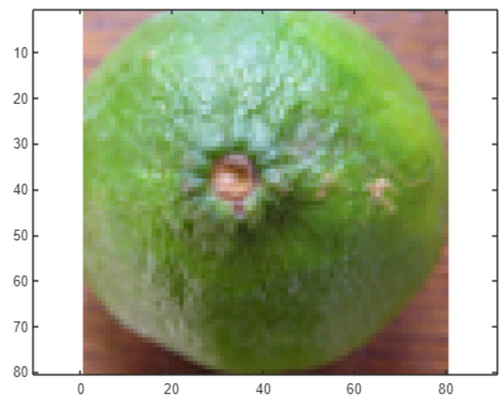


(b) Generated

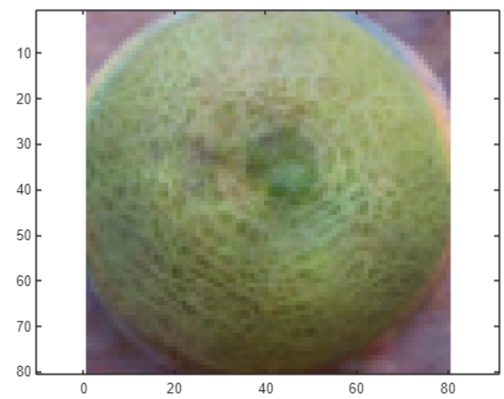


(a) Actual Image

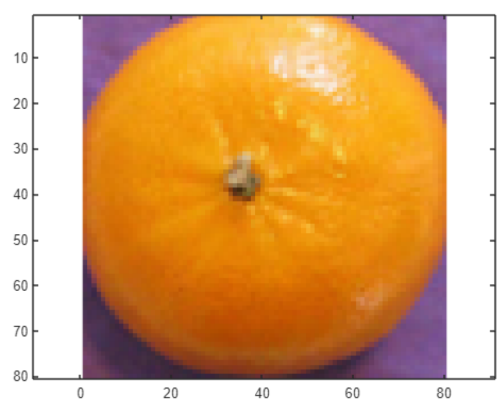
(b) Generated



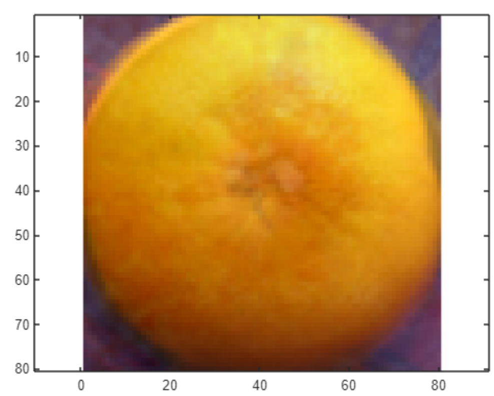
(a) Actual Image



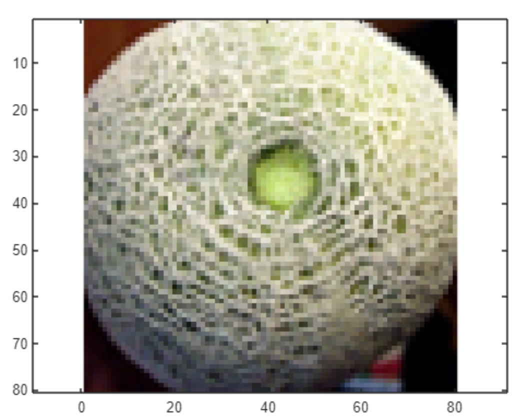
(b) Generated



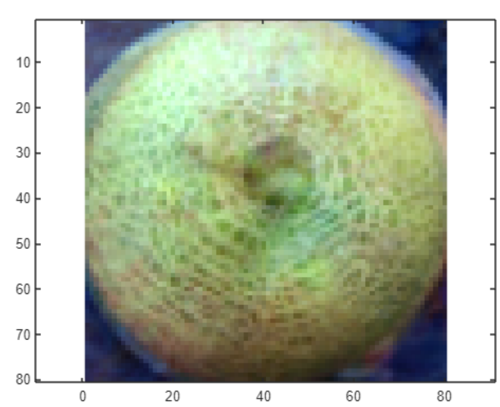
(a) Actual Image



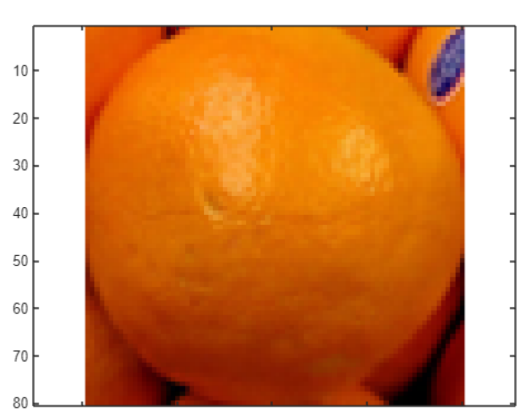
(b) Generated



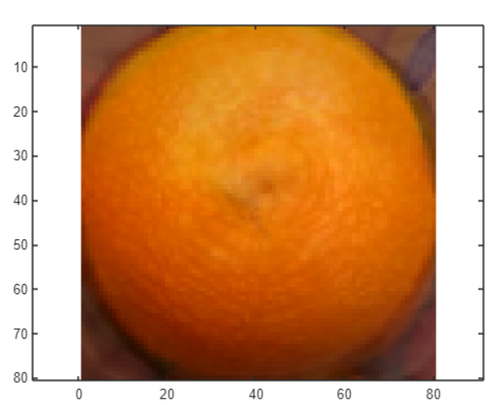
(a) Actual Image



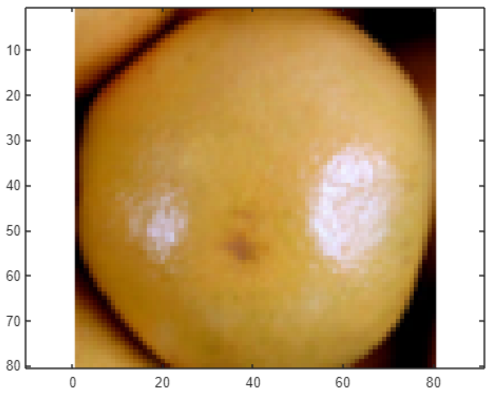
(b) Generated



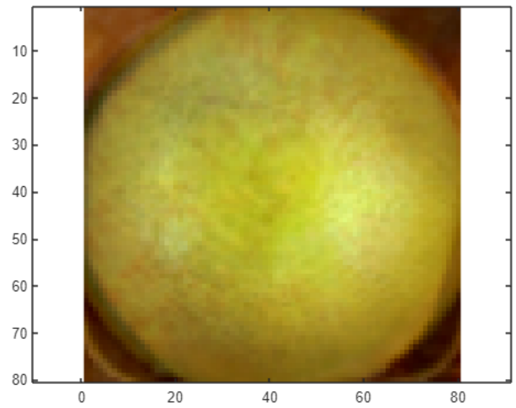
(a) Actual Image



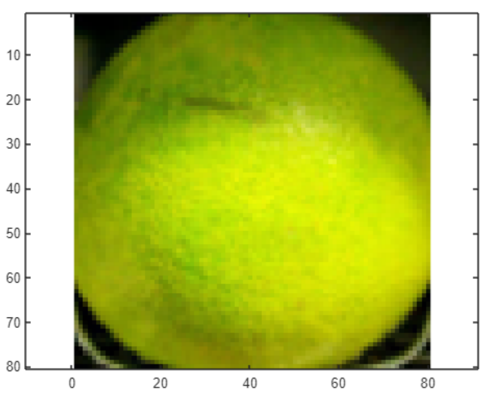
(b) Generated



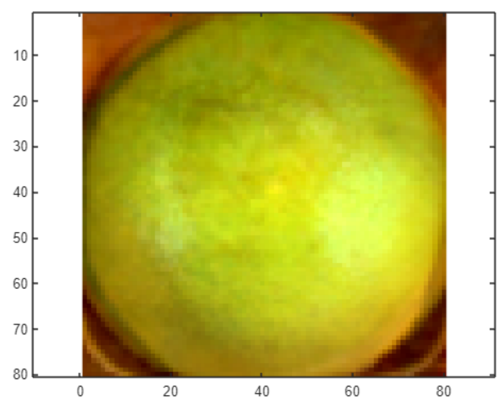
(a) Actual Image



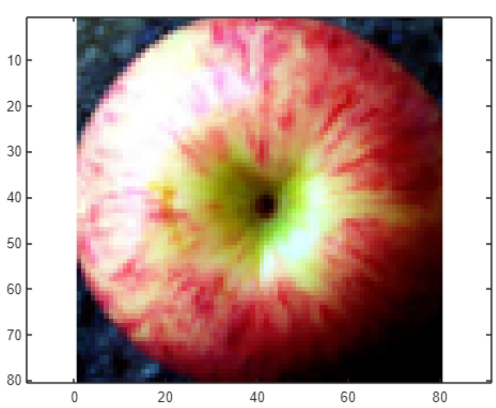
(b) Generated



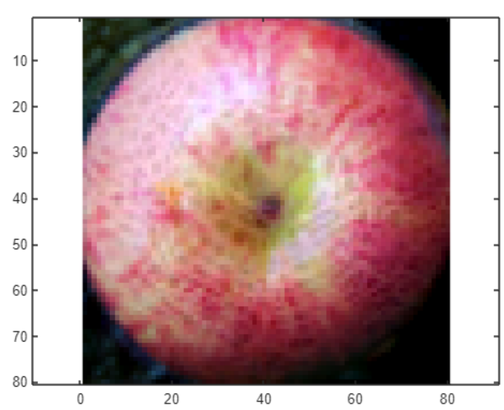
(a) Actual Image



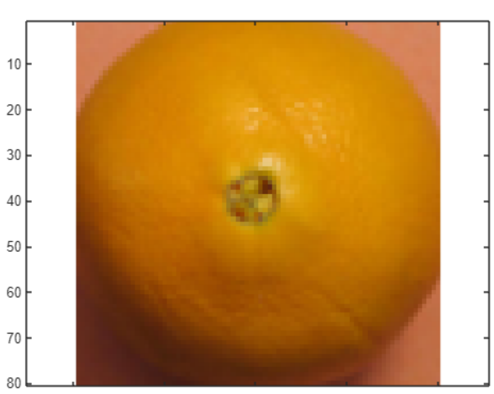
(b) Generated



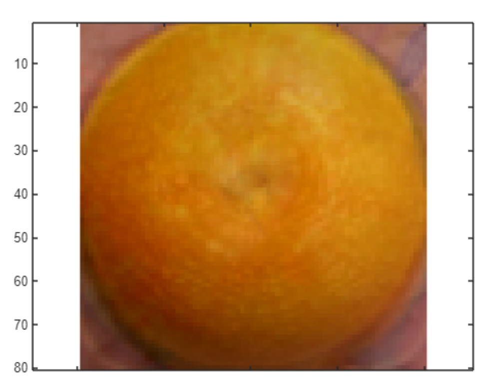
(a) Actual Image



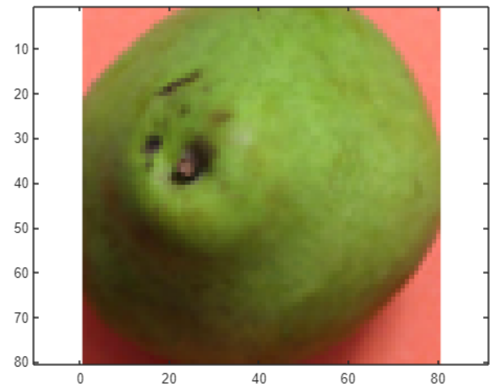
(b) Generated



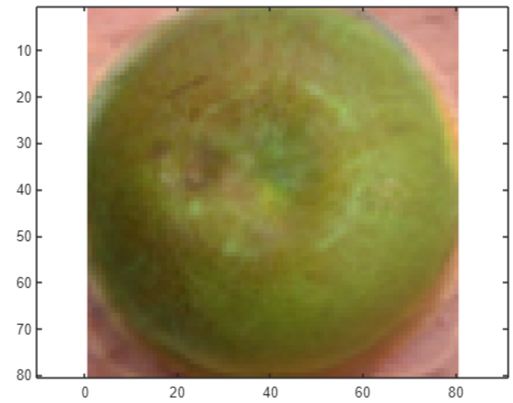
(a) Actual Image



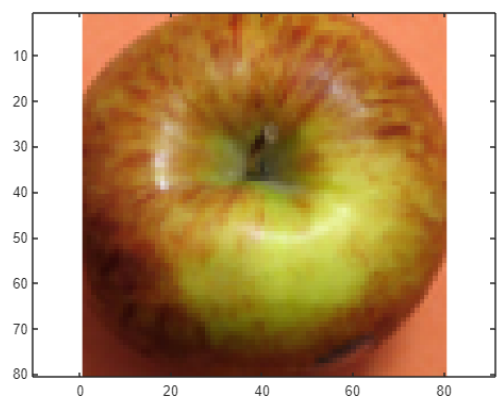
(b) Generated



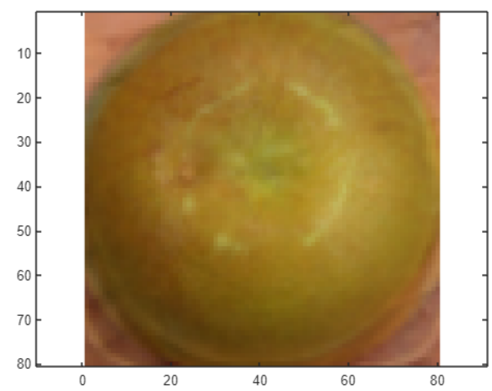
(a) Actual Image



(b) Generated

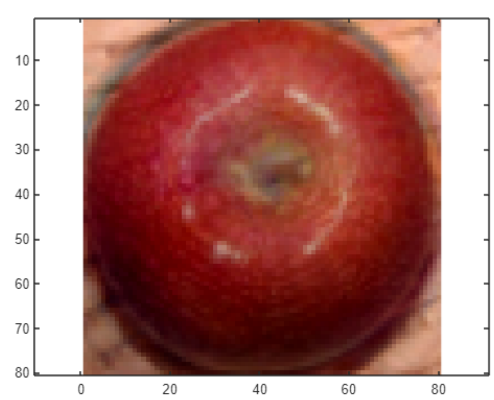


(a) Actual Image



(b) Generated

(a) Actual Image



(b) Generated

3 Generating Random Images

To generate random images that are representatives of the given data set we note that every image is closest to the linear combination of the four eigen vectors as given above. However, to be a representative of the given data set, we should produce linear combinations of the same four eigenvectors.
To make sure the images are distinct we make sure that the coefficients used are not the same as those used in finding any of the representative images.
This algorithm produces distinct fruits, the images of which are after the comparisons.

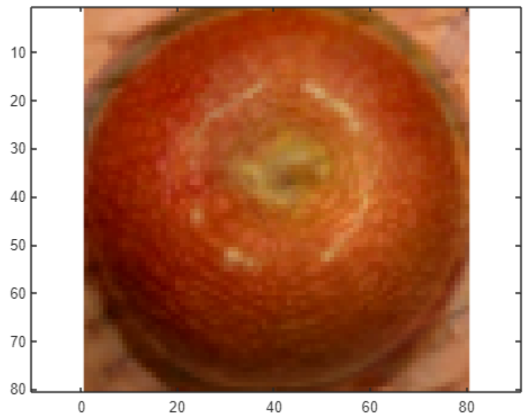


Figure 21: Random Fruit 1

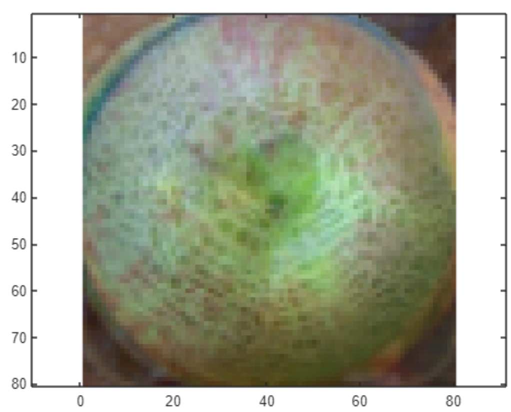


Figure 22: Random Fruit 2

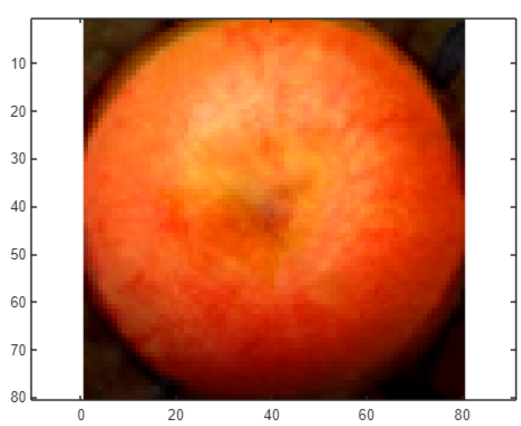


Figure 23: Random Fruit 3