# Question 6

Avadhoot Jadhav - 210050027
Hrishikesh Jedhe Deshmukh - 210050073

October 2022

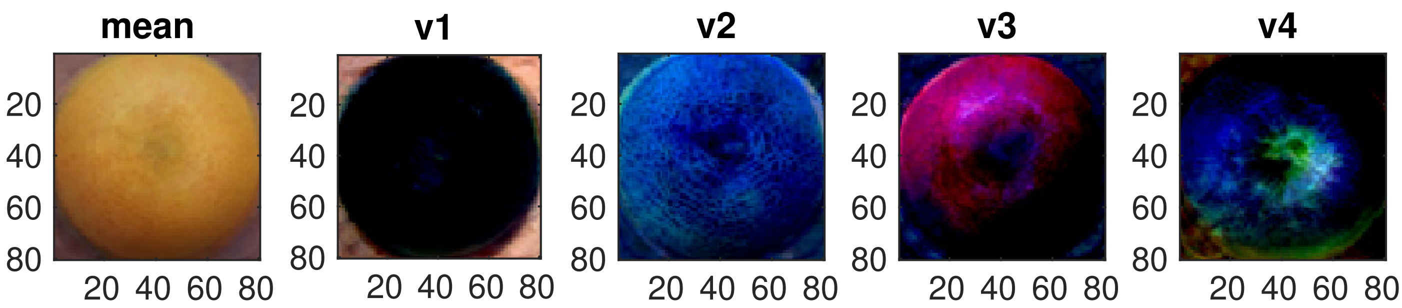
## 1 Computing mean and covariance

The process is similar as done in previous questions to compute sample mean and sample covariance.

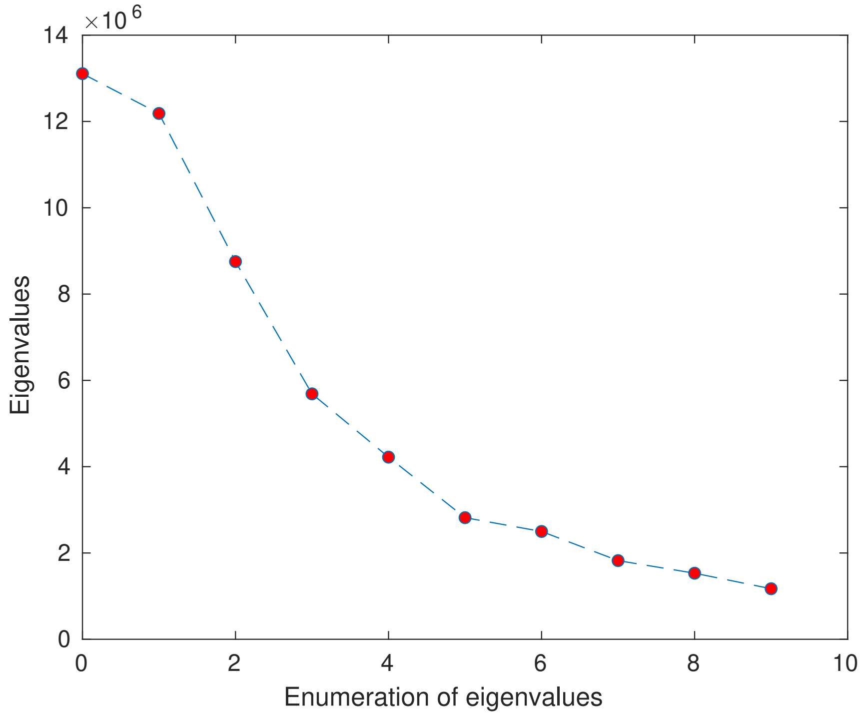
$$\mu = \frac{\sum_{i=1}^{n} X_i}{n}$$

$$C = \frac{\sum_{i=1}^{n}(X_i - \mu)(X_i - \mu)^T}{n}$$

**Mean and Top 4 principle eigenvectors:**




**Plot of top 10 Eigenvalues**



## 2 Finding closest representation

Let, $p = \mu + a_1 v_1 + a_2 v_2 + a_3 v_3 + a_4 v_4$ be the closest representation where $v_1$, $v_2$, ....... $v_{19200}$ are orthonormal eigenvectors of covariance matrix $C$.

$v_i \cdot v_i = 1 \quad \forall i \in [1, 19200]$ and
$v_i \cdot v_j = 0 \quad \forall i \neq j \quad$ s.t $i, j \in [1, 19200]$

Now, we have to minimize the Frobenius norm of $p - x$, where $x$ represents the image vector.

$$||p - x||_{Fro} = |p - x|$$

Let's take dot product of $p - x$ with all the orthonormal eigenvectors.

$$\begin{aligned}
(p - x) \cdot v_1 &= (\mu - x) \cdot v_1 + a_1 \\
(p - x) \cdot v_2 &= (\mu - x) \cdot v_2 + a_2 \\
(p - x) \cdot v_3 &= (\mu - x) \cdot v_3 + a_3 \\
(p - x) \cdot v_4 &= (\mu - x) \cdot v_4 + a_4 \\
(p - x) \cdot v_5 &= (\mu - x) \cdot v_5 \\
(p - x) \cdot v_6 &= (\mu - x) \cdot v_6 \\
&. \\
&. \\
&. \\
(p - x) \cdot v_{19200} &= (\mu - x) \cdot v_{19200}
\end{aligned}$$

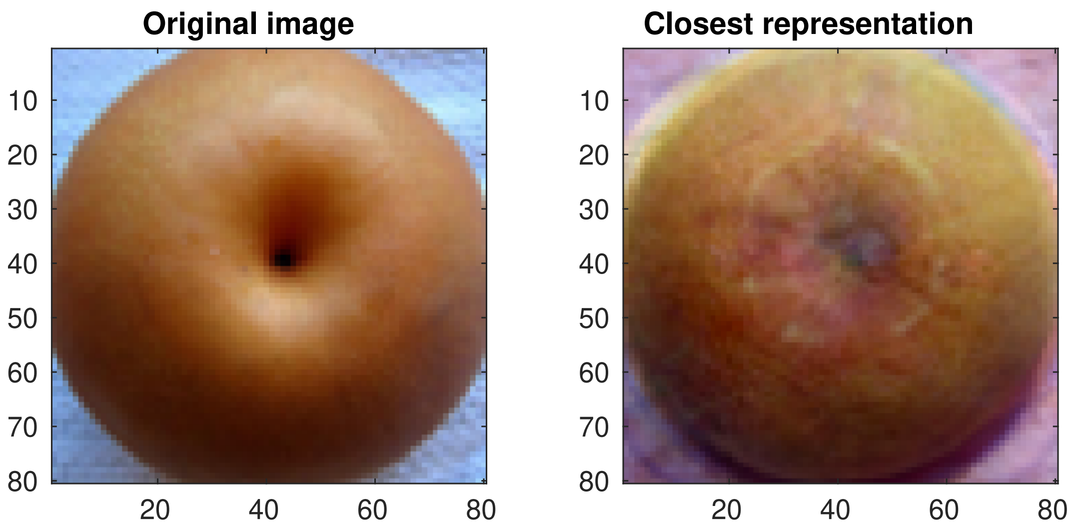
As $|p - x| = \sqrt{\sum_{i=1}^{19200} [(p - x).v_i]^2} = ||p - x||_{Fro}$, we need to minimize $\sum_{i=1}^{19200} [(p - x).v_i]^2$.

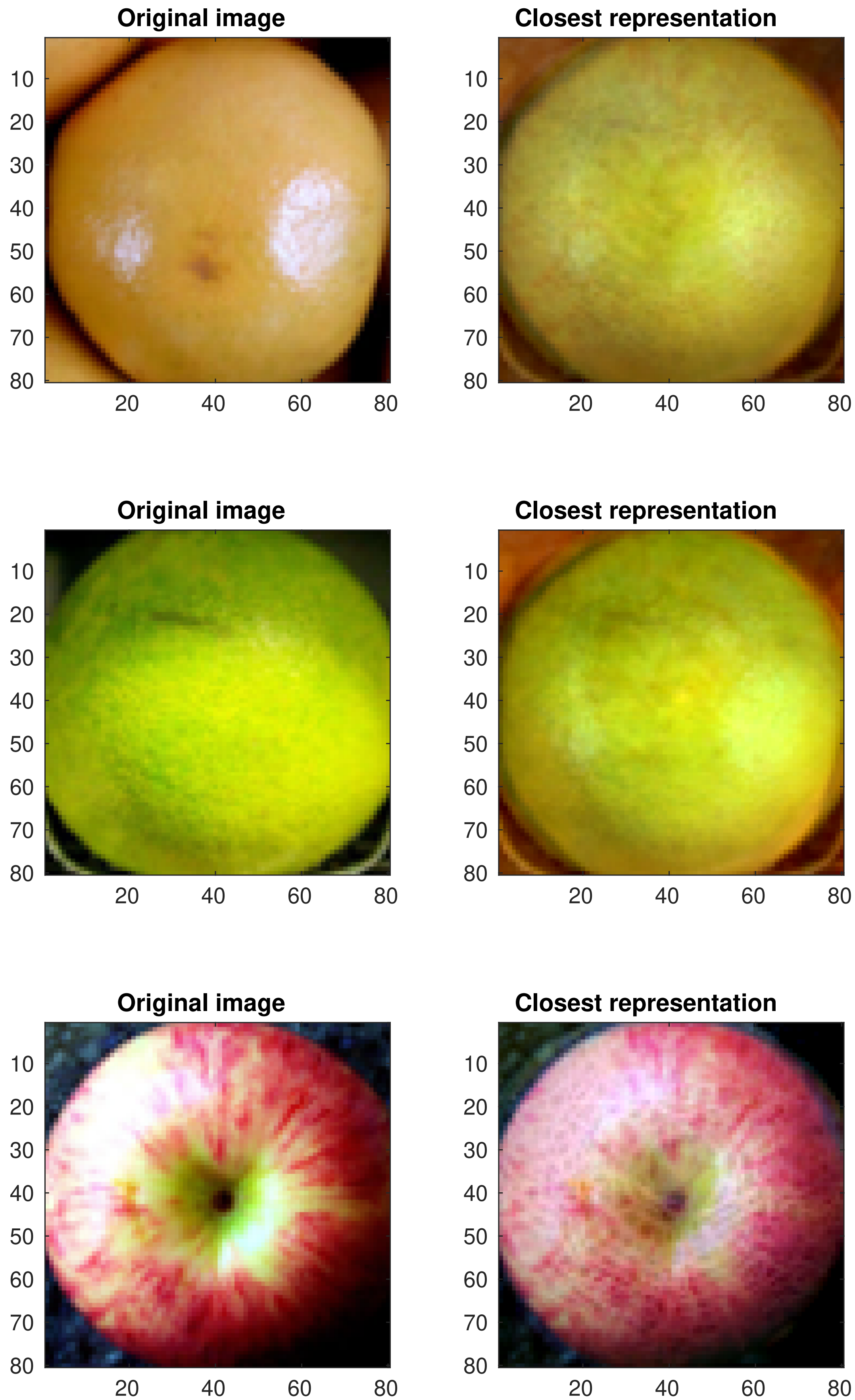
We need to minimize the sum of squares of RHS in the above equations.
Notice that beyond the 4th equation from top, RHS terms are constant and thus cannot be minimized. But we can minimize the first 4 RHS terms by setting them individually to 0.

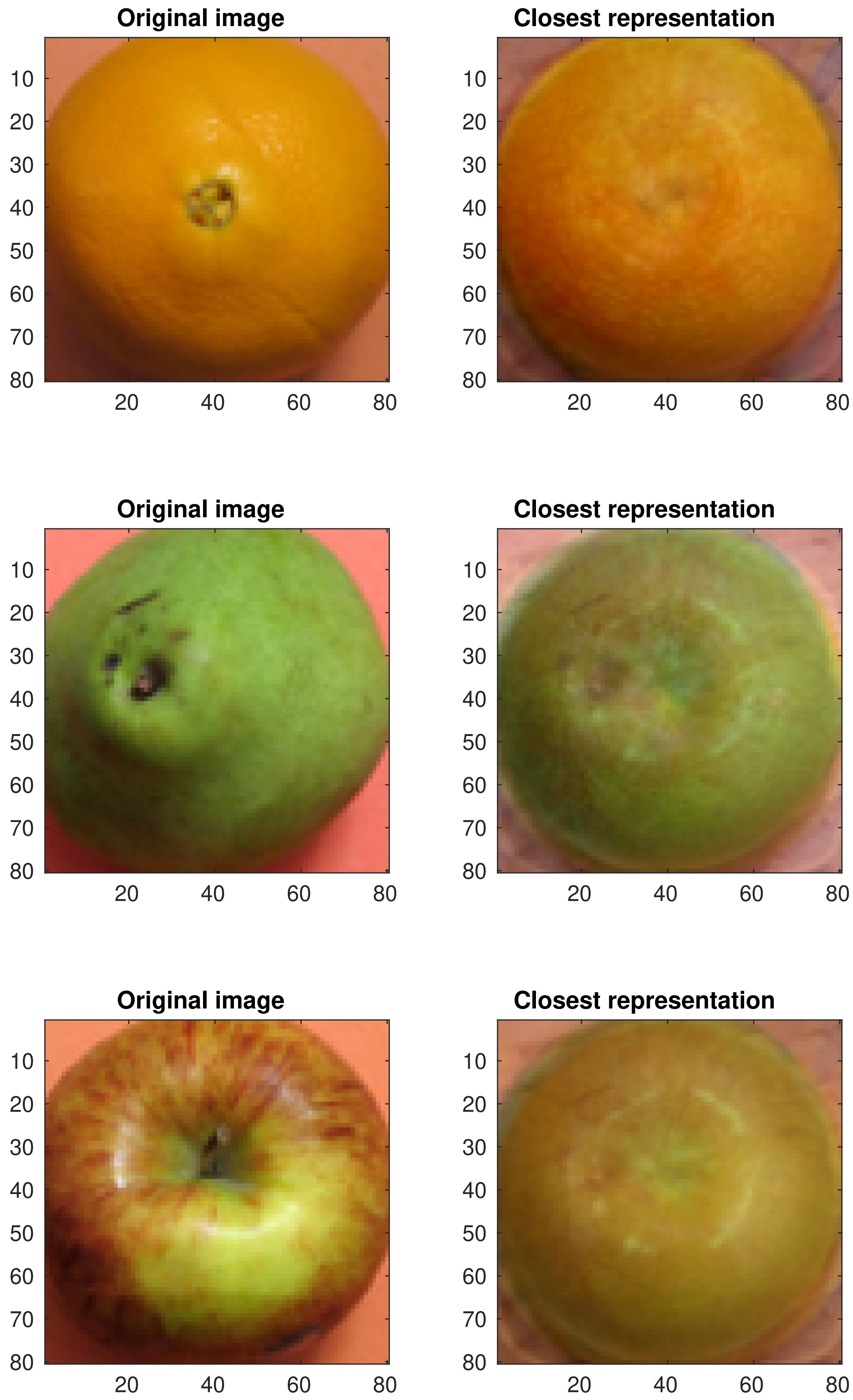
$$\begin{aligned}
(\mu - x) \cdot v_1 + a_1 = 0 &\implies a_1 = -(\mu - x) \cdot v_1 \\
(\mu - x) \cdot v_2 + a_2 = 0 &\implies a_2 = -(\mu - x) \cdot v_2 \\
(\mu - x) \cdot v_3 + a_3 = 0 &\implies a_3 = -(\mu - x) \cdot v_3 \\
(\mu - x) \cdot v_4 + a_4 = 0 &\implies a_4 = -(\mu - x) \cdot v_4
\end{aligned}$$

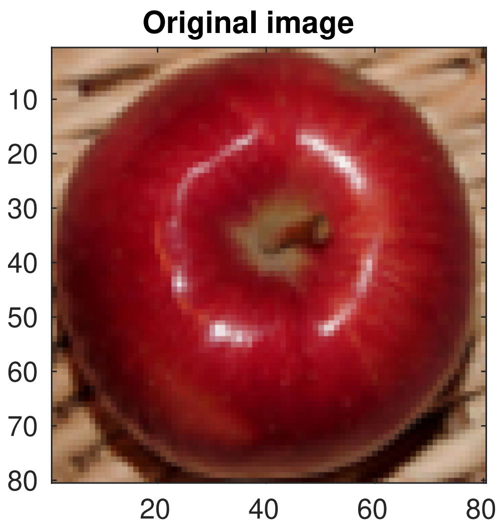
Thus, we can get the closest representation by using the above obtained values in the representation of $p$.

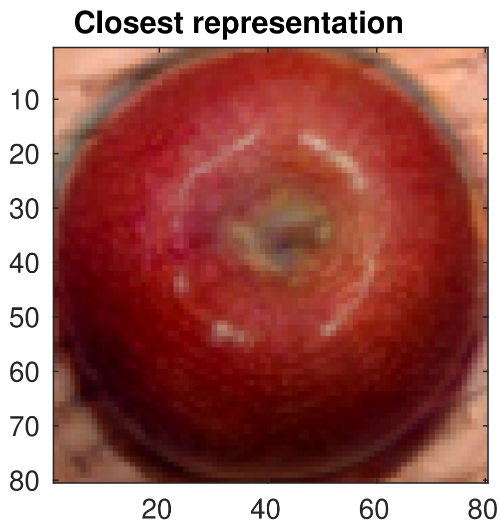
**Original fruit images and closest representation**

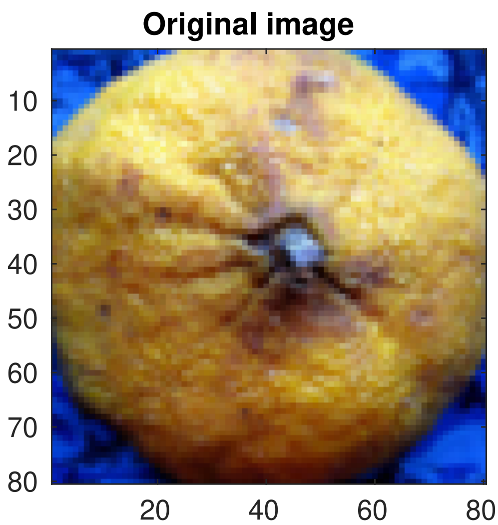


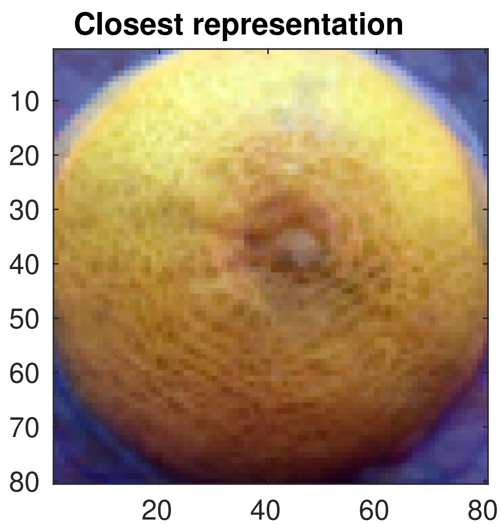

Original image
Closest representation
10
20
30
40
50
60
70
80
20
40
60
80
Original image
Closest representation
Original image
Closest representation

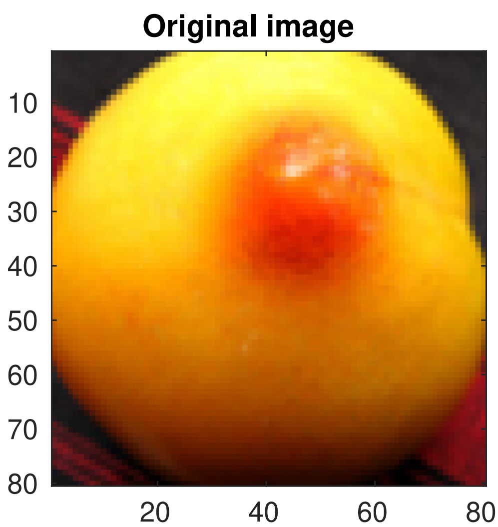

Original image
Closest representation
10
20
30
40
50
60
70
80
20
40
60
80
Original image
Closest representation
Original image
Closest representation

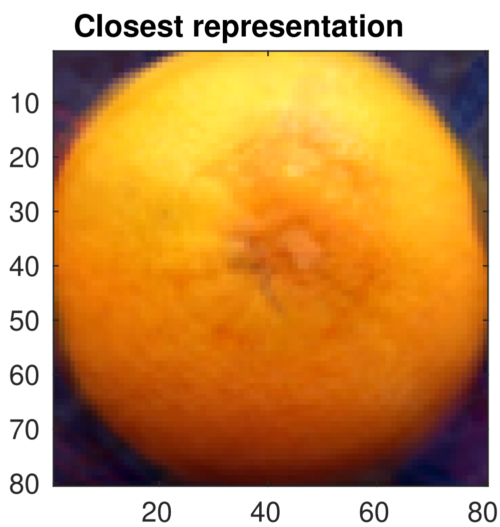

Original image
10 20 30 40 50 60 70 80
20 40 60 80

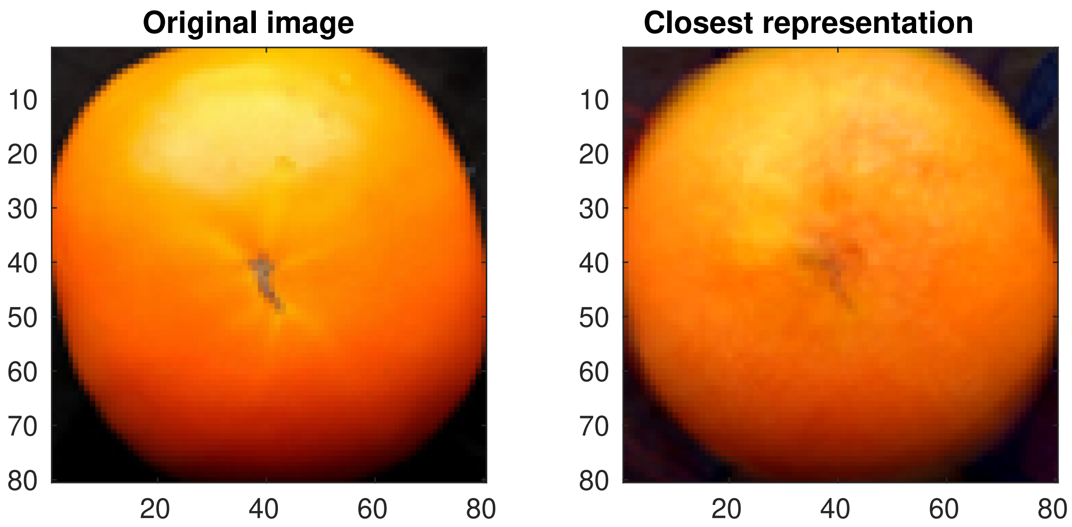

Closest representation
10 20 30 40 50 60 70 80
20 40 60 80

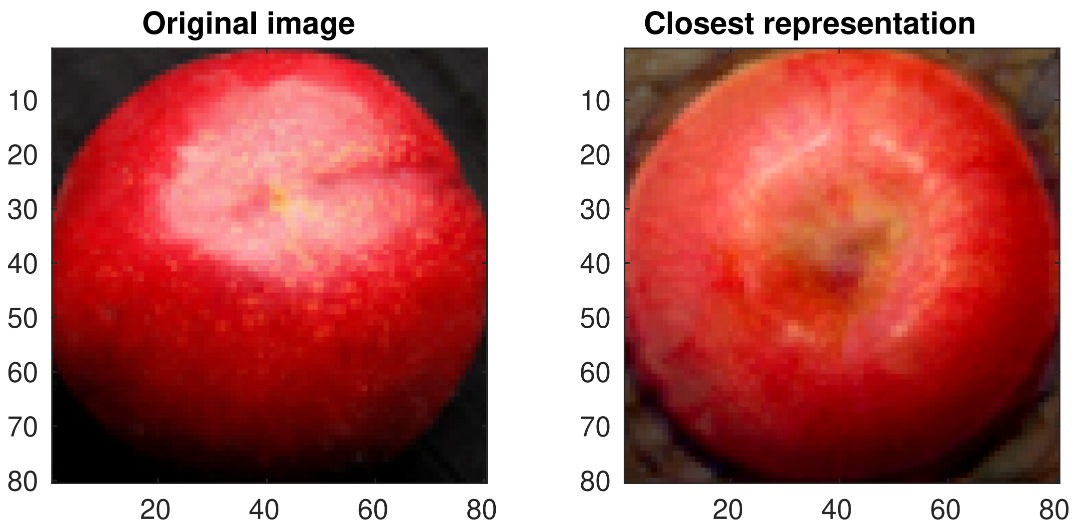

Original image
10 20 30 40 50 60 70 80
20 40 60 80

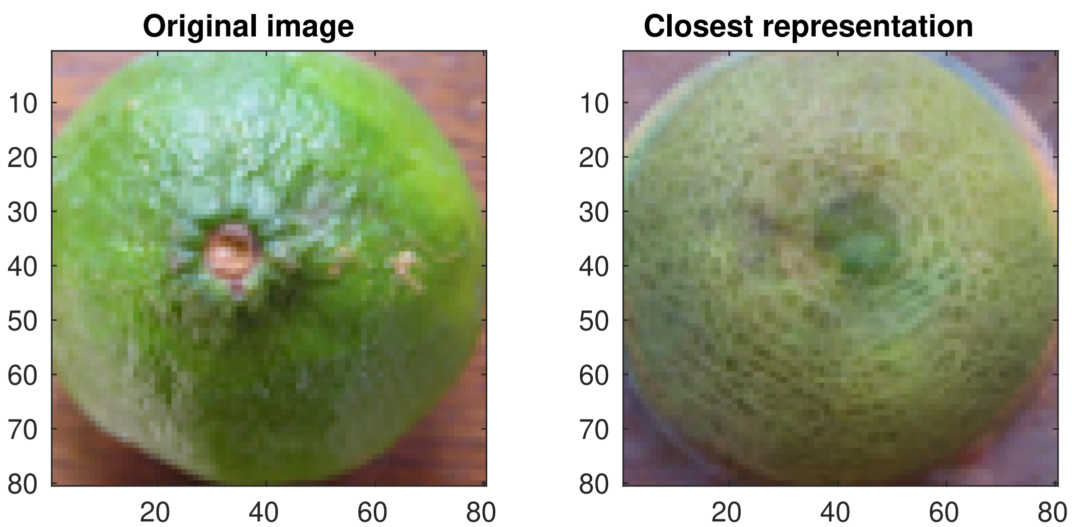

Closest representation
10 20 30 40 50 60 70 80
20 40 60 80

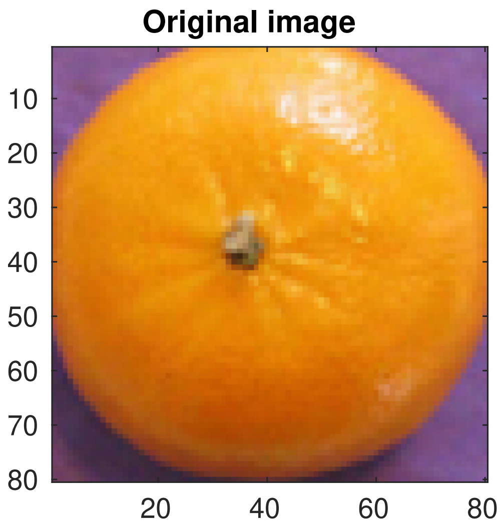

Original image
10 20 30 40 50 60 70 80
20 40 60 80

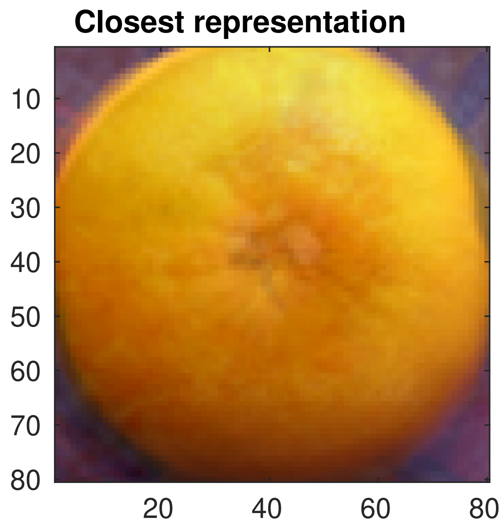

Closest representation
10 20 30 40 50 60 70 80
20 40 60 80

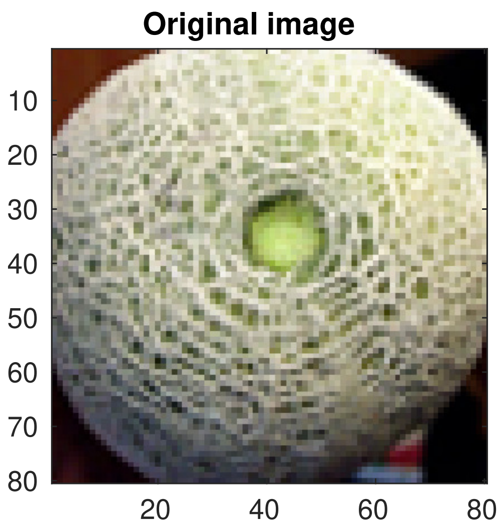

Original image
Closest representation
10 20 30 40 50 60 70 80
20 40 60 80

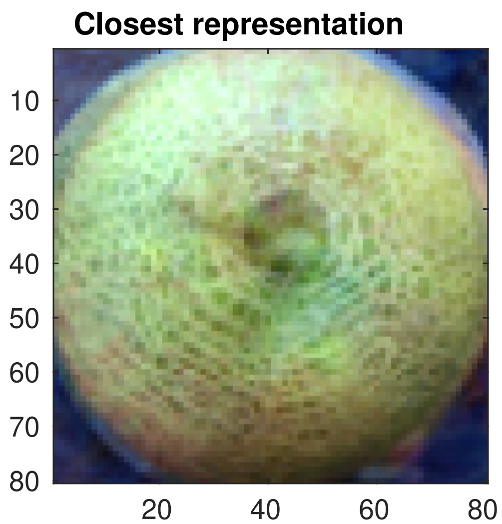

Original image
Closest representation
10 20 30 40 50 60 70 80
20 40 60 80

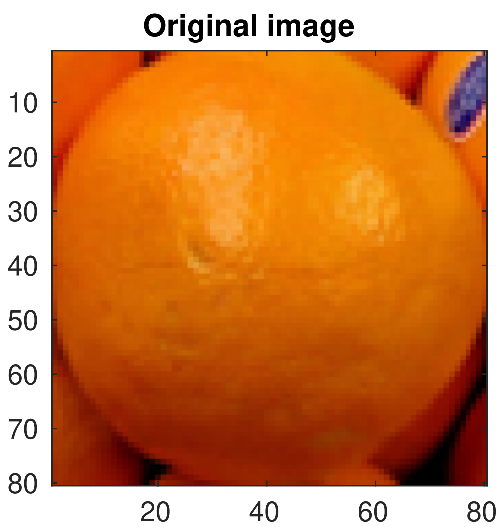

Original image
Closest representation
10 20 30 40 50 60 70 80
20 40 60 80

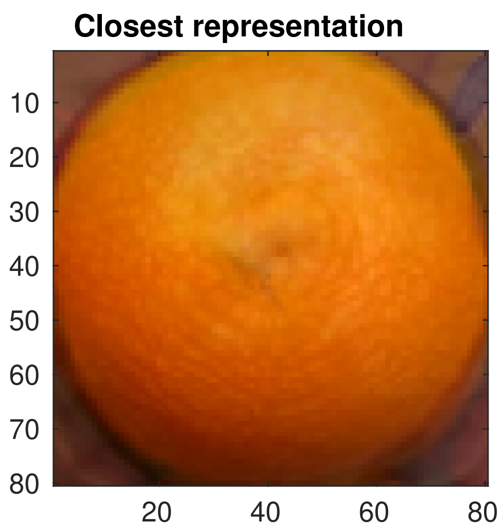

Original image
10 20 30 40 50 60 70 80
20 40 60 80


Closest representation
10 20 30 40 50 60 70 80
20 40 60 80


Original image
10 20 30 40 50 60 70 80
20 40 60 80


Closest representation
10 20 30 40 50 60 70 80
20 40 60 80


Original image
10 20 30 40 50 60 70 80
20 40 60 80


Closest representation
10 20 30 40 50 60 70 80
20 40 60 80

## 3 Generating new representative images

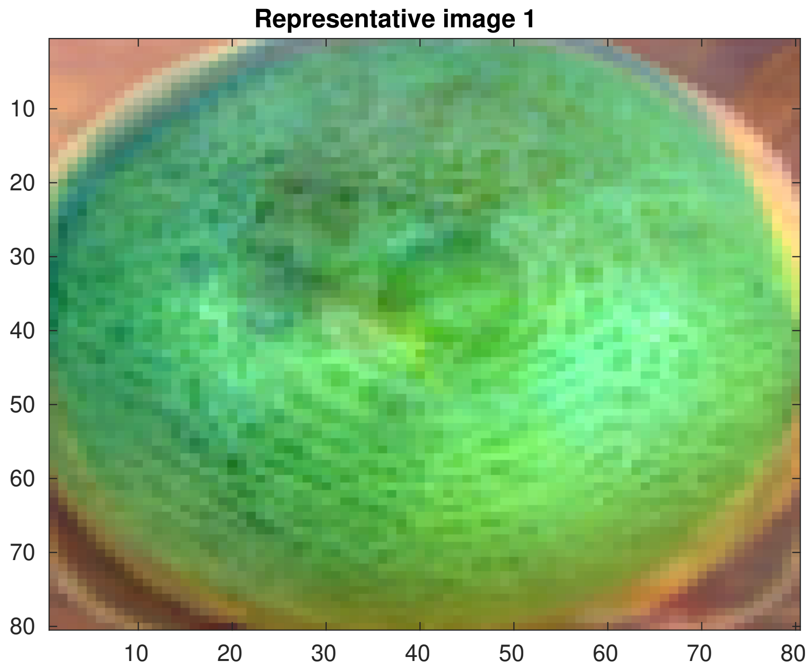
The idea behind creating new, representative images of fruits is to perturb the values around mean. We will perturb around mean using the eigenvectors whose coefficients are drawn from a standard normal distribution but scaled by a factor of standard deviation.

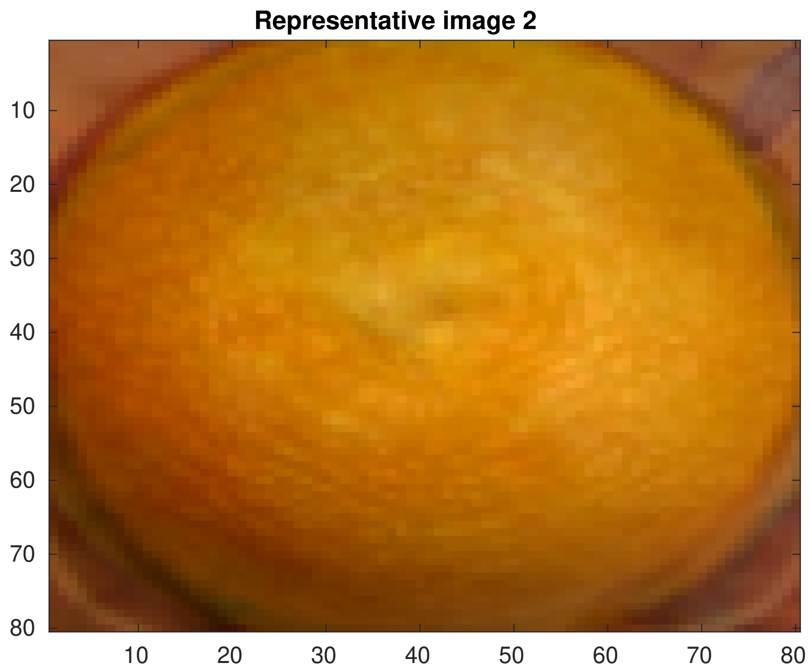
Let $I$ be the new representative image,

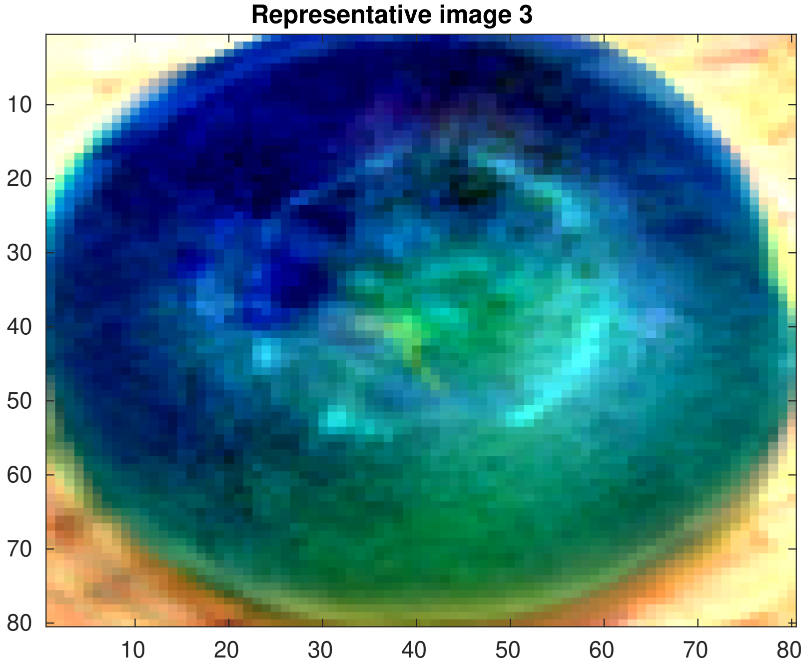
$$I = \mu + a\sqrt{\lambda_1}v_1 + b\sqrt{\lambda_2}v_2 + c\sqrt{\lambda_3}v_3 + d\sqrt{\lambda_4}v_4$$

where $a$, $b$, $c$ and $d$ are drawn from a standard normal distribution.

**New images generated:**




Representative image 3
10
20
30
40
50
60
70
80
10
20
30
40
50
60
70
80