Group 2
A. Christenbury, G. Menara
Homework 3

Step 1: Initially we took three photos and we produced 5 photos for each one, using the k-means algorithm for 5 different values (k = {1,2,5,10,20}). The resulting images are shown below.

Original Photo 1   Original Photo 2   Original Photo 3

Photo 1:

Original Photo

K = 1

K = 2

K = 5

K = 10

K = 20

Photo 2:

Original Photo

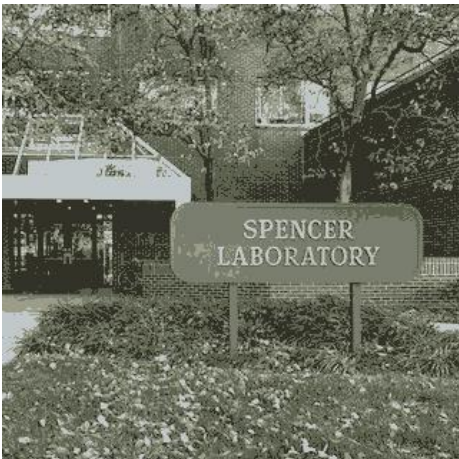K = 1

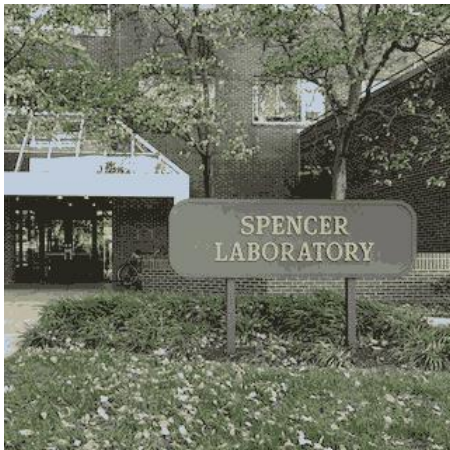K = 2

K = 5

K = 10

K = 20

Photo 3:

Original Photo

K = 1

K = 2

K = 5

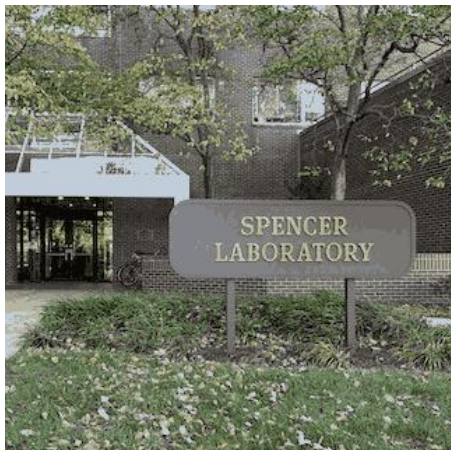K = 10

K = 20

<u>Step 2:</u>

The following table presents the size of each photo in kB.

	Original	K = 1	K = 2	K = 5	K = 10	K = 20
Photo 1	42.1	2.2	23.2	27.4	26.4	26.7
Photo 2	35.6	2.2	12.2	20.6	19.5	20.1
Photo 3	57.9	2.2	32.2	33.9	34	34.1

We observed that as k increases, the produced photos become more similar to the original one, which was expected, since the number of k corresponds to the number of different colors we have in the picture. Furthermore, as we can see from the table above showing the sizes of the each photo we can make the following conclusions:

•   Initially, for k = 1, we have the same size and it is significantly less than the size of the original ones.

•   For k=2 and k=5 the size increases as k increases in all the three photos.

•   For k>5 we observe that the size converges to a number which is approximately half of the size of the original images.

The produced photos which are the best for all instances are:

1)   As we can see for the photo 1, the produced photos for k = 5,10,20 are pretty similar but the photo of k=20 has the smallest size. That is why we would choose k=20.

2)   For photo 2, the produced photos using k=10,20 are pretty close and since the size of k=10 picture is smaller, we would choose k=10.

3)   For photo 3, the produced photos using k=5,10,20 are similar and therefore we would choose k=5 because is the photo with the smallest size.

<u>Step 3:</u>

Since we want a photo that the best k will be k=2, we tried to find a picture with exact 2 colors. For this reason we chose the UD logo. As we can see from the images below when we chose k = {5,10,20} we do not get an image which is similar to the original one. Therefore, we conclude that the best k in this case is k=2.

Original Extra Photo   K =1   K =2

K =5   K =10   K =20