Assignment_5 Report

Swapna Kumar

Performed K-means clustering for image compression on two images for different values of K= [2,5,10,15,20].

In the context of image compression, K-means clustering can be used to reduce the number of colors in an image. Each cluster in this case represents a different color. The centroid of each cluster is the new color, and all the pixels in that cluster take on that color. The number of clusters K is equivalent to the number of colors in the compressed image.

Below are my observations from the resulting images.

- As we increase the value of K, the image quality generally improves because there are more centroids to represent different colors.

- However, there is a tradeoff between image quality and degree of compression. As we increase the degree of compression (reduce K), the image quality decreases. This is because we are reducing the number of colors in the image, which can cause loss of detail and distinction between different parts of the image. Higher K values lead to larger file sizes.
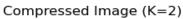
Below are the compressed images for both the datasets:

1. Penguins

Original image:

Compressed images:

Compressed Image (K=5)

Compressed Image (K=10)

Compressed Image (K=15)

Compressed Image (K=20)

2. Koala

Original image:

Compressed images:

Compressed Image (K=2)

Compressed Image (K=5)

Compressed Image (K=10)

Compressed Image (K=15)

Compressed Image (K=20)

From the above results, a good value of K will be:

- For Penguin dataset, K = 15.

- For Koala dataset, K = 10.