Manoj Kumar Selvaraj

Professor Anjum Chida

CS 6375.001

11 April 2021

Assignment 5: K-Means Clustering

Koala Image Files

Image	Size
Original	Total Pixels: 786,432 | 18,874,368 Bits
K Value = 2	786,480 Bits
K Value = 5	2,359,416 Bits
K Value = 10	3,145,968 Bits
K Value = 15	3,146,088 Bits
K Value = 20	3,932,640 Bits

Original Image

K Value = 2

K Value = 5

K Value = 10

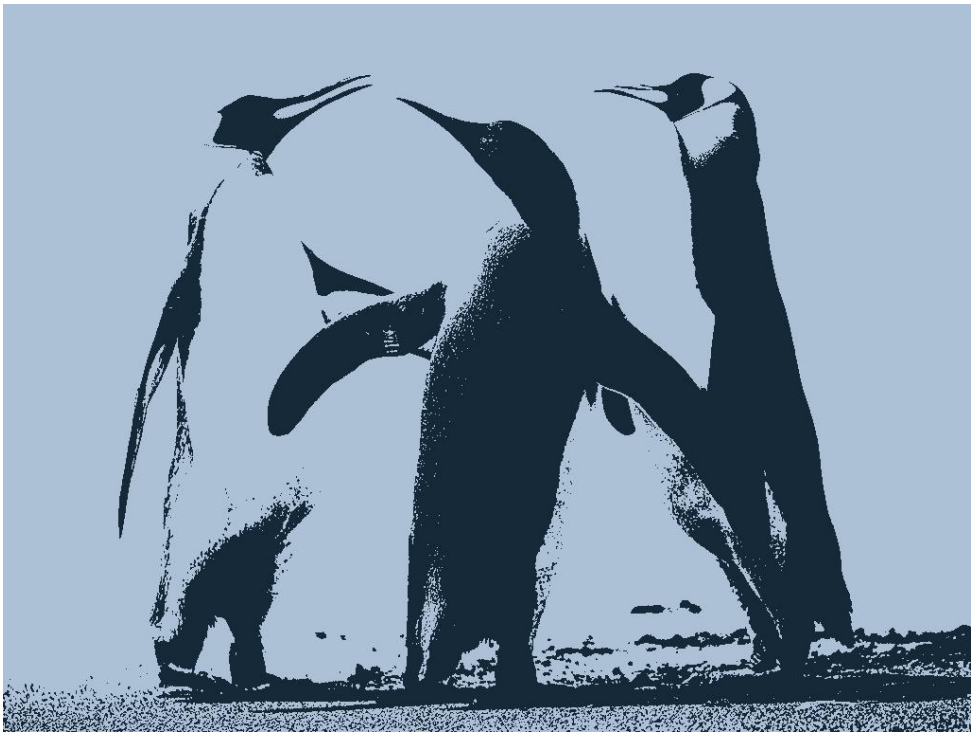
K Value = 15

K Value = 20

Penguins Image Files

Image	Size
Original	Total Pixels: 786,432 | 18,874,368 Bits
K Value = 2	786,480 Bits
K Value = 5	2,359,416 Bits
K Value = 10	3,145,968 Bits
K Value = 15	3,146,088 Bits
K Value = 20	3,932,640 Bits

Original Image



K Value = 2

K Value = 5

K Value = 10

K Value = 15

K Value = 20

Discussion

K-Means Image Compression Size Formula: (Pixels((ceil(log(K)/log(2)))) + (K(24))

There is a tradeoff between image quality and degree of compression. The K value essentially tries to represent the picture in terms of the number of colors present in the picture. A low K value will provide a high degree of compression; however, the image quality will be low. For example, a K value of 2 provides great compression but it represents the entire picture in terms of just 2 colors. Thus, the image quality will be low. The higher the K value the closer the representation gets to the true image but the compression amount will be low. In addition, the initial random points chosen make an impact in the convergence process in terms of how quick the means stabilize. Furthermore, the picture itself also seems to make a difference on how fast the algorithm is able to perform the K-Means process. Taking into consideration the tradeoff between image quality and degree of compression, the optimal K value for the koala image would be 10 and the optimal K value for the penguins image would be 15.