ASSIGNMENT 5

K-Means Clustering Compression

1. Koala Image

|  |  |
| --- | --- |
|  | Original Size  780831 bytes |
|  | K = 2  135573 bytes  Compression Ratio = 6.37 |
|  | K =5  176557 bytes  Compression Ratio = 4.42 |
|  | K = 10  171,773 bytes  Compression Ratio = 4.54 |
|  | K = 15  171049 bytes  Compression Ratio = 4.56 |
|  | K = 20  167,422 bytes  Compression Ratio = 4.66 |

1. Penguin Image

|  |  |
| --- | --- |
|  | Original Size  777835 bytes |
|  | K = 2  85161 bytes  Compression Ratio = 9.13 |
|  | K =5  102105 bytes  Compression Ratio = 7.61 |
|  | K = 10  117743 bytes  Compression Ratio = 6.61 |
|  | K = 15  111685 bytes  Compression Ratio = 6.96 |
|  | K = 20  119384 bytes  Compression Ratio = 6.51 |

Is there a tradeoff between image quality and degree of compression? What would be a good value of K for each of the two images?

Certainly, there is a tradeoff between the image quality with the degree of compression. When the K values are less the compression image quality is blurred and not like the original image. K represents the number of the cluster the image should be represented with. When K =2, both the koala image and the penguin image would have been compressed with 2 shades.

The best value of K for Koala image would be K = 15 since the image resembles the original with best compression rate. For the penguin, the best would be K = 10 as the compression ratio between K = 10 and K = 20 are close.

The Number of Iteration is 10.