**Program 3**

A081605 清華研 張睿

**Outline**

* Test images
* Comparison of different block sizes
* Information packing ability
* Comparison of different quantization schemes with qualitative comparison
  + Scheme 1: different k
  + Scheme 2: different k
  + Scheme 1 vs scheme 2 (with the same k)
  + Scheme 3: different number of bits
* Code listing

**Test images**

* Image 1: more spatial details

A train crossing a bridge over a body of water

Description automatically generated

* Image 2: less spatial details

A picture containing water, outdoor, rain, large

Description automatically generated

**Comparison of different block sizes**

* With 1st quantization scheme: (remaining the left-upper triangle of the coefficients map, i.e. remaining the first half of the coefficients)
  + Image 1: DCT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Block size | 4 | 8 | 16 | 32 |
| eRMS | 146.2 | 171.0 | 183.3 | 189.1 |
| SNR | 121.6 | 103.8 | 96.7 | 93.8 |

* + Image 1: WHT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Block size | 4 | 8 | 16 | 32 |
| eRMS | 152.7 | 185.1 | 202.4 | 212.0 |
| SNR | 116.4 | 95.8 | 87.6 | 83.5 |

* + Image 1: DFT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Block size | 4 | 8 | 16 | 32 |
| eRMS | 138.1 | 171.5 | 215.6 | 284.7 |
| SNR | 128.1 | 102.4 | 80.8 | 60.6 |

* + For all the three transforms, smaller sub-image sizes perform better on reconstruction.
  + To find the reason, I did a further test and found that for larger block size, there are less “first-half largest coefficients” falls exactly in the upper-left triangle of the coefficient map. For block sizes 4, 8, 16, 32, the ratios are 38.1%, 33.8%, 31.7%, 30.5%, respectively. (test with DCT)
  + Image 2: DCT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Block size | 4 | 8 | 16 | 32 |
| eRMS | 6.4 | 7.6 | 7.7 | 8.3 |
| SNR | 1233.3 | 1049.5 | 1034.2 | 958.0 |

* + Image 2: WHT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Block size | 4 | 8 | 16 | 32 |
| eRMS | 6.7 | 8.0 | 8.8 | 9.1 |
| SNR | 1194.1 | 998.6 | 905.2 | 871.2 |

* + Image 2: DFT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Block size | 4 | 8 | 16 | 32 |
| eRMS | 5.3 | 6.5 | 9.3 | 13.5 |
| SNR | 1491.1 | 1225.2 | 850.4 | 587.0 |

* + Same as image 1: smaller sub-image sizes perform better on reconstruction.
  + The ratio mentioned above for image 2 are 35.5%, 31.2%, 28.9%, 27.8%, respectively.
  + The reconstruction error is much lower than image 1 possibly because the spatial details are much less in the image 2.
* With 2nd quantization scheme: (keep the first-half largest coefficients for each sub-images)
  + Image 1: DCT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Block size | 4 | 8 | 16 | 32 |
| eRMS | 52.0 | 38.2 | 35.4 | 35.9 |
| SNR | 343.7 | 467.3 | 505.5 | 497.6 |

* + Image 1: WHT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Block size | 4 | 8 | 16 | 32 |
| eRMS | 59.4 | 45.4 | 41.2 | 42.6 |
| SNR | 300.6 | 393.3 | 433.9 | 420.0 |

* + Image 1: DFT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Block size | 4 | 8 | 16 | 32 |
| eRMS | 98.1 | 83.4 | 76.0 | 75.2 |
| SNR | 181.6 | 213.7 | 234.7 | 237.2 |

* + Opposite from the result of scheme 1, the trend is larger blocks lead to smaller reconstruction errors.
  + Image 2: DCT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Block size | 4 | 8 | 16 | 32 |
| eRMS | 3.0 | 2.0 | 1.6 | 1.5 |
| SNR | 2630.1 | 3986.0 | 4889.4 | 5323.5 |

* + Image 2: WHT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Block size | 4 | 8 | 16 | 32 |
| eRMS | 4.5 | 3.9 | 2.7 | 2.1 |
| SNR | 1763.9 | 2037.2 | 2879.2 | 3683.7 |

* + Image 2: DFT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Block size | 4 | 8 | 16 | 32 |
| eRMS | 6.3 | 5.4 | 4.5 | 3.7 |
| SNR | 1257.4 | 1470.2 | 1743.0 | 2160.1 |

* + The trend is same as image 1, but the reconstruction errors are much smaller.

**Information packing ability**

* I follow the mean square error with assumption in the textbook to compare the information packing ability of the different transforms.

**A close up of a sign

Description automatically generated**

where

A screenshot of a cell phone

Description automatically generated

is the coefficient masking function and SigmaT(u,v) ^2 is the variance of the coefficient at transform location (u,v). The final simplification is based on the orthonormal nature of the basis images and the assumption that the pixels of G are generated by a random process with zero mean and known covariance.

* My setting: block size=8, keep only the first 32 coefficients for quantization.
* The ems results show that DCT has the best information packing ability.
  + Image 1, DCT: 10941.5
  + Image 1, WHT: 11842.8
  + Image 1, DFT: 16691.3
  + Image 2, DCT: 488.0
  + Image 2, WHT: 512.9
  + Image 2, DFT: 672.8

**Comparison of different quantization schemes**

* Scheme 1: keep only the first k coefficients (block size=8; k follows the zig-zag fashion)
  + Image 1: DCT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| k | 6 | 15 | 32 | 54 |
| eRMS | 600.4 | 402.7 | 171.0 | 50.8 |
| SNR | 28.8 | 43.5 | 103.8 | 351.9 |

* + Image 1: WHT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| k | 6 | 15 | 32 | 54 |
| eRMS | 663.1 | 454.0 | 185.1 | 56.4 |
| SNR | 26.0 | 38.5 | 95.8 | 316.8 |

* + Image 1: DFT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| k | 6 | 15 | 32 | 54 |
| eRMS | 653.2 | 489.8 | 171.5 | 45.5 |
| SNR | 23.0 | 34.9 | 102.4 | 390.7 |

* + As expected, larger k lead to smaller reconstruction error.
  + Image 1: Qualitative comparison of DCT: k = 32 is better enough.

A train crossing a bridge over a body of water

Description automatically generated A bridge over a body of water

Description automatically generated A train crossing a bridge over a body of water

Description automatically generated A train crossing a bridge over a body of water

Description automatically generated A train crossing a bridge over a body of water

Description automatically generated

Original k = 6 k = 15 k = 32 k = 54

* + Image 2: Qualitative comparison of DCT: the differences in the sky part are subtle in all settings.

A picture containing water, outdoor, rain, large

Description automatically generated A picture containing water, sky, large, star

Description automatically generated A picture containing water, outdoor, large, body

Description automatically generated A picture containing water, outdoor, nature, skiing

Description automatically generated A picture containing water, outdoor, nature, large

Description automatically generated

Original k = 6 k = 15 k = 32 k = 54

* Scheme 2: keep only the coefficients of the k largest coefficients (block size=8)
  + Image 1: DCT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| k | 6 | 15 | 32 | 54 |
| eRMS | 379.9 | 167.3 | 38.2 | 1.2 |
| SNR | 46.2 | 106.1 | 467.3 | 14445.6 |

* + Image 1: WHT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| k | 6 | 15 | 32 | 54 |
| eRMS | 416.5 | 189.9 | 45.4 | 1.7 |
| SNR | 42.0 | 93.4 | 393.3 | 10051.3 |

* + Image 1: DFT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| k | 6 | 15 | 32 | 54 |
| eRMS | 470.7 | 251.3 | 83.4 | 7.2 |
| SNR | 37.0 | 70.3 | 213.7 | 2471.6 |

* + As expected, larger k lead to smaller reconstruction error.
  + Image 1: Qualitative comparison of DCT: similar with scheme 1, k = 32 is good enough.

A train crossing a bridge over a body of water

Description automatically generated A train crossing a bridge over a body of water

Description automatically generated A train crossing a bridge over a body of water

Description automatically generated A train crossing a bridge over a body of water

Description automatically generated A train crossing a bridge over a body of water

Description automatically generated

Original k = 6 k = 15 k = 32 k = 54

* Scheme 1 vs scheme 2
  + As expected, scheme 2 has a much better performance than scheme 1 due to remaining the basis blocks with more importance. However, the computation is heavier due to like sorting.
  + Image 1: Qualitative comparison of DCT: with the same k, scheme 2 is better

A train crossing a bridge over a body of water

Description automatically generated A train crossing a bridge over a body of water

Description automatically generated

k = 15, scheme 1 k = 15, scheme 2

* Scheme 3: distribute different number of bits to different coefficients (comparing different number of total bits used) (block size=8)
  + Image 1: DCT

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| total | 16 | 64 | 128 | 512 | 1024 |
| eRMS | 967.3 | 759.4 | 600.2 | 394.0 | 299.2 |
| SNR | 17.4 | 22.3 | 25.1 | 44.0 | 57.7 |

* + Image 1: WHT

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| total | 16 | 64 | 128 | 512 | 1024 |
| eRMS | 1063.9 | 797.4 | 690.7 | 381.0 | 289.7 |
| SNR | 15.7 | 21.3 | 24.7 | 45.5 | 59.6 |

* + Image 1: DFT

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| total | 16 | 64 | 128 | 512 | 1024 |
| eRMS | 509.2 | 586.5 | 510.1 | 425.2 | 352.8 |
| SNR | 33.9 | 29.2 | 33.7 | 40.5 | 48.9 |

* + As expected, more bits lead to better performance.
  + Image 1: Qualitative comparison of DCT

A picture containing photo, building, sitting, bridge

Description automatically generatedA bridge over a body of water

Description automatically generatedA black and white photo of a cage

Description automatically generatedA bridge over a body of water

Description automatically generatedA bridge over a body of water

Description automatically generated

k = 16 k = 64 k = 128. k = 512 k = 1024

* + Image 2: Qualitative comparison of DCT

A picture containing white

Description automatically generated  A close up of an animal

Description automatically generated A picture containing large, water, white

Description automatically generated A picture containing water, large, white

Description automatically generated

k = 16. k = 64. K = 128 k = 512 k = 1024