Simple JPEG-compressor

1. Introduction

Image compression is a type of data compression applied to digital images, the objective of image compression is to reduce the redundancy of the image and to store or transmit in an efficient form.

2. Method

The main goal of such system is to reduce the storage quantity as much as possible, and the decoded image displayed in the monitor can be similar to the original image as much as can be.

2.1 Discrete Cosine Transform

The objective of this step is to divide the gray value of the image into many 8×8 blocks. The mathematical definition of the Forward DCT and the Inverse DCT are as follows:

Forward DCT

$$F(\mathrm{u},\mathrm{v}) = \frac{2}{N} C(u)C(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cos\left[\frac{\pi(2x+1)u}{2N}\right] \cos\left[\frac{\pi(2y+1)v}{2N}\right]$$

for $\mathrm{u},\mathrm{v} = 0, \ldots, N-1$;

where $N = 8$ and $C(k) = \begin{cases} \frac{1}{\sqrt{2}} & for\ \ k = 0 \\ 1 & otherwise \end{cases}$

Inverse DCT

$$f(x,y) = \frac{2}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u)C(v)F(u,v) \cos\left[\frac{\pi(2x+1)u}{2N}\right] \cos\left[\frac{\pi(2y+1)v}{2N}\right]$$

for $\mathrm{u},\mathrm{v} = 0, \ldots, N-1$; where $N = 8$

The $f(x,y)$ is the value of each pixel in the selected 8×8 block, and the $F(u,v)$ is the DCT coefficient after transformation. The transformation of the 8×8 block is also a 8×8 block composed of $F(u,v)$.

2.2 Quantizer

Quantization is the step where we actually throw away data. The DCT is a lossless procedure. The data can be precisely recovered through the IDCT. During quantization every coefficients in the 8×8 DCT matrix is divided by a corresponding quantization value. The quantized coefficient is defined as:

$$F(u,v)_{Quantization} = \text{round}(\frac{F(u,v)}{Q(u,v)})$$

and the reversed coefficient is defined as:

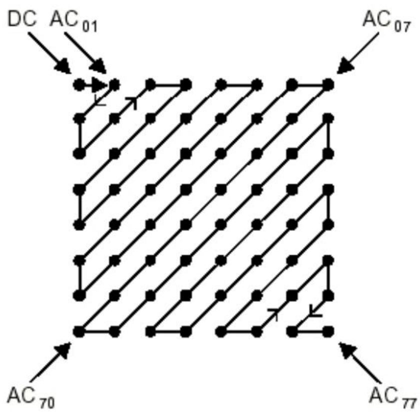
$$F(u,v)_{De-quantization} = F(u,v)_{Quantization} \times Q(u,v)$$

where $Q_{8\times 8}(u,v)$ is defined as:

$$Q_Y = \begin{pmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{pmatrix}$$

2.3 Zigzag coding

After quantization, the DC coefficient is treated separately from the 63 AC coefficients. The DC coefficient is a measure of the average value of the original 64 image samples. Because there is usually strong correlation between the DC coefficients of adjacent 8×8 blocks, the quantized DC coefficient is encoded as the difference from the DC term of the previous block. This special treatment is worthwhile, as DC coefficients frequently contain a significant fraction of the total image energy. The other 63 entries are the AC components. They are treated separately from the DC coefficients in the entropy coding process.



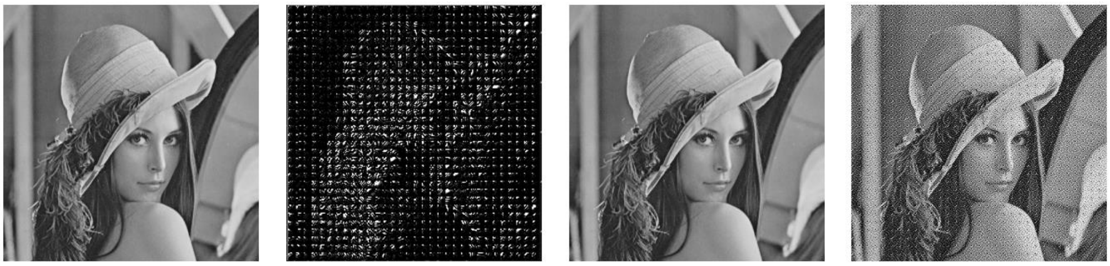
2.4 Run-length-encoding

After quantization and zigzag scanning, we obtain the one-dimensional vectors with a lot of consecutive zeroes. We can make use of this property and apply run-length encoding, which is variable length coding.

The notation (L,F) means that there are L zeros in front of F, and EOB (End of Block) is a special coded value means that the rest elements are all zeros.

3. Results and Evaluation

3.1 Result



(a) (b) (c) (d)

Figure 1 – (a) is the input image, (b) is the result of performing 8x8 block-based DCT, (c) is the result of lossless reconstruction (i.e. being quantized to fraction), (d) is the result of lossy reconstruction (i.e. being quantized to integer).

3.2 Evaluation

In order to evaluate the performance of the image compression coding, it is necessary to define a measurement that can estimate the difference between the original image and the decoded image

3.2.1 Signal-to-Noise Ratio

An image with low noise would tend to have a high SNR and, conversely, the same image with a higher level of noise would have a lower SNR.

$$SNR = \frac{\sum_{y=0}^{H-1}\sum_{x=0}^{W-1} f'(x,y)^2}{\sum_{y=0}^{H-1}\sum_{x=0}^{W-1}[f(x,y) - f'(x,y)]^2}$$

The SNR between Figure. 1 (a) and (d) is shown below.

SNR: 10.301072

3.2.2 Root Mean Square Error

The mean square error given below can be approximated in terms of a summation involving the original and restored images.

$$RMSE = \sqrt{\frac{\sum_{y=0}^{H-1} \sum_{x=0}^{W-1} [f(x,y) - f'(x,y)]^2}{HW}}$$

The RMSE between Figure. 1 (a) and (d) is shown below.

```
RMSE:  42.047250962180186
```

4. Reference

[1] https://github.com/getsanjeev/compression-DCT/blob/master/zigzag.py