

影像處理與電腦視覺 作業四

統計所 李權恩 R26094022

這次嘗試透過作業來學習 C++，因為沒有學過 C/C++ 以及資料結構相關的課程，所以能做到的部分不多，多半是參考別人放在 github 上的作品來完成，以下是我主要使用的參考資料：

編碼器 (encoder)：

<https://github.com/stbrumme/toojpeg>

<https://koushtav.me/jpeg/tutorial/2017/11/25/lets-write-a-simple-jpeg-library-part-1/>

解碼器 (decoder)：

<https://koushtav.me/jpeg/tutorial/c++/decoder/2019/03/02/lets-write-a-simple-jpeg-library-part-2/>

我的主程式如下：

<https://github.com/Quan-En/DIPCV/tree/main/assignment4>

使用方法如下：

```
>cd my_pyth\assignment4
```

```
>Encoder\x64\Debug\Encoder.exe <source file path with full name>
```

Methodology

1. 色域空間 (color spaces)

在一般影像上有需多不同方式來表達色彩空間，最常見的如 RGB 三原色，也有將影像亮度與影像色度區分開的表達方式如 YCbCr。其中 Y 代表流明 (luminance)，Cb 和 Cr 代表色度 (chrominance)，為藍色和紅色的濃度偏移量。在有了亮度與色度的區別後，由於人眼構造對於亮度變化較敏感並對色度變化較遲鈍，因此在 jpeg 的標準下，可以透過轉換色彩空間的方式搭配降採樣 (down-sampling) 的方式來減少需要被壓縮的資訊。轉換方式主要透過下列公式：

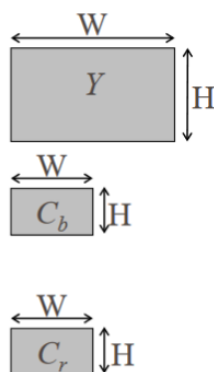
$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299000 & 0.587000 & 0.114000 \\ -0.168736 & -0.331264 & 0.500002 \\ 0.500000 & -0.418688 & -0.081312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix}$$

其反轉換如下：

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.0 & 1.0 & 1.40210 \\ 1.0 & -0.34414 & -0.71414 \\ 1.0 & 1.77180 & 0.0 \end{bmatrix} \begin{bmatrix} Y \\ Cb - 128 \\ Cr - 128 \end{bmatrix}$$

2. 色度降採樣 (chrominance down-sampling)

在色彩空間中提及人眼構造對於色度的變化較遲鈍，因此在圖片轉換到 YCbCr 的空間時，可以在對色度進行降採樣，本篇實作的降採樣方式為 4:2:0，其含義為在 Y 空間全保留並在 Cb, Cr 空間中，只取原影像長寬的一半。如下圖所示：



3. 離散餘弦變換 (Discrete Cosine Transform; DCT)

DCT 是與傅立葉變換相關的一種變換，類似於離散傅立葉變換，但是只使用實數，對於人眼而言，大部分的圖片在沒有高頻頻率的情況下還是能大抵分辨出影像的原樣，因此透過轉換到頻率域再搭配量化 (quantization) 的動作，轉換的結果會得到較多的 0，如此一來又可進一步提升壓縮的能力。操作上將轉換過色彩空間的影像再切分成 8×8 的塊 (block) 再各自進行 DCT。轉換公式如下：

$$F(u, v) = \sum_{m=0}^{8-1} \sum_{n=0}^{8-1} f(m, n) \alpha(u) \alpha(v) \cos\left(\frac{\pi(2m+1)u}{2 \times 8}\right) \cos\left(\frac{\pi(2n+1)v}{2 \times 8}\right)$$

反轉換公式如下：

$$f(x, y) = \sum_{u=0}^{8-1} \sum_{v=0}^{8-1} F(u, v) \alpha(u) \alpha(v) \cos\left(\frac{\pi(2x+1)u}{2 \times 8}\right) \cos\left(\frac{\pi(2y+1)v}{2 \times 8}\right)$$

其中 $\alpha(0) = \frac{1}{2\sqrt{2}}, \alpha(k) = \frac{1}{2}, \text{ for } k > 0$

4. 量化 (Quantization)

在經由 DCT 轉換到頻率域後，可以藉由量化的方式來彌平高頻的資訊，下表為亮度與色度各自的量化表，主要作法為該塊 (block) 對應的位置除以量化表的數值並四捨五入。

Luminance Quantization Table of $Q_{j,k}$ values

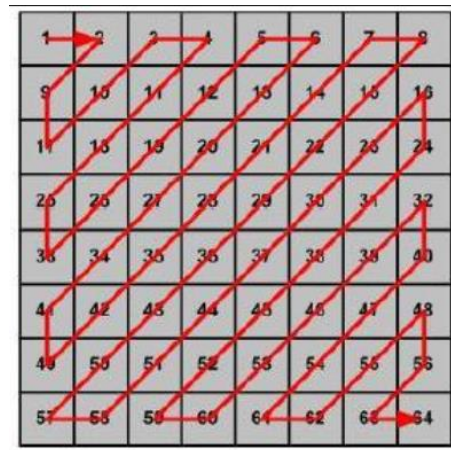
16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Chrominance Quantization Table of $Q_{j,k}$ values

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

5. 曲折掃描 (Zigzag scanning)

曲折掃描是 JPEG 壓縮技術的核心之一，輸入三個 64DCT 係數 DCT-Y, DCT-Cb, DCT-Cr 後，輸出三個曲折掃描向量，目的是將出現頻率低的係數組合在一起，掃描方式如下圖。



6. Differential coding

在差分編碼主要的記錄方式便是存取前後兩期的差，在 JPEG 技術當中，圖片會被切分成許多 8×8 的塊，在每個塊執行完 DCT 後，我們取 DC 來進行差分編碼，公式如下：

$$DC_{i+1} = DC_i + \text{difference}$$

其中第 0 期的 difference 設為 0。

7. Run length coding

遊程編碼，是一種與資料性質無關的無失真資料壓縮技術，基於「使用變動長度的碼來取代連續重複出現的原始資料」來實現壓縮，在 JPEG 的過程中只對 AC 係數進行此編碼，DC 係數由另外的差分編碼執行。

8. Huffman encoding

在有了差分編碼與遊程編碼後，可以根據霍夫曼編碼來將上述所有的結果轉換成 01 的 bitstream。

Category	Values	Bits for the value
1	-1,1	0,1
2	-3,-2,2,3	00,01,10,11
3	-7,-6,-5,-4,4,5,6,7	000,001,010,011,100,101,110,111
4	-15,...,-8,8,...,15	0000,...,0111,1000,...,1111
5	-31,...,-16,16,...,31	00000,...,01111,10000,...,11111
6	-63,...,-32,32,...,63	000000,...,011111,100000,...,111111
7	-127,...,-64,64,...,127	0000000,...,0111111,1000000,...,1111111
8	-255,...,-128,128,...,255	...
9	-511,...,-256,256,...,511	...
10	-1023,...,-512,512,...,1023	...
11	-2047,...,-1024,1024,...,2047	...

Implement result

將 standard_test_image 內的所有影像壓縮成 JPEG 檔後，下表整理了壓縮前後的 PSNR 值。此外所有壓縮成 JPEG 檔的影像皆可直接透過電腦開啟。但在 Decoder 方面並沒有做出正確的結果，輸出的圖片會有扭曲色彩錯誤的狀況。

Image name	color	PSNR	
		4:4:4	4:2:0
Cameraman	gray	46.183	-
House		49.389	
Jetplane		42.689	
Lake		38.540	
Lena_gray_256		39.015	
Lena_gray_512		40.822	
Livingroom		39.444	
Mandrill_gray		42.261	
Peppers_gray		38.195	
Pirate		39.771	
Walkbridge		38.049	

Woman_blonde		39.063	
Woman_darkhair		43.801	
Lena_color_256	color	40.738	34.981
Lena_color_512		40.821	34.777
Mandril_color		36.801	27.387
Peppers_color		38.234	31.665