

Image Compression*

石大川 <sdc17@mails.tsinghua.edu.cn>

2020 年 4 月 26 日

目录

1 1D-DCT & 2D-DCT	1
1.1 Principle	1
1.2 Results	2
1.2.1 Comparison between two trials	2
1.2.2 Different Coefficients	3
1.2.3 Different Block Size	3
2 Quantization	5
2.1 Principle	5
2.2 Results	6
2.2.1 With and Without Quantization	6
2.2.2 Different A and Q	6
3 Conclusion	8

1 1D-DCT & 2D-DCT

1.1 Principle

1D-DCT 公式

$$F(u) = c(u) \sum_{i=0}^{N-1} f(i) \cos\left[\frac{(2i+1)\pi}{N} u\right]$$

其中

$$c(u) = \begin{cases} \sqrt{\frac{1}{N}}, & u = 0 \\ \sqrt{\frac{2}{N}}, & u \neq 0 \end{cases}$$

2D-DCT 公式

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

*<https://github.com/sdc17/NaiveIC>

其中

$$c(u) = \begin{cases} \sqrt{\frac{1}{N}}, & u = 0 \\ \sqrt{\frac{2}{N}}, & u \neq 0 \end{cases}$$

$$c(v) = \begin{cases} \sqrt{\frac{1}{N}}, & v = 0 \\ \sqrt{\frac{2}{N}}, & v \neq 0 \end{cases}$$

实际做的时候写成矩阵形式

$$G = \begin{bmatrix} \sqrt{\frac{1}{N}}[1 & 1 & \dots & 1] \\ \sqrt{\frac{2}{N}}[[\cos(\frac{\pi}{2N})] & [\cos(\frac{3\pi}{2N})] & \dots & [\cos(\frac{(2N-1)\pi}{2N})]] \\ \dots & \dots & \dots & \dots \\ \sqrt{\frac{2}{N}}[[\cos(\frac{(N-1)\pi}{2N})] & [\cos(\frac{(N-1)3\pi}{2N})] & \dots & [\cos(\frac{(N-1)(2N-1)\pi}{2N})]] \end{bmatrix}$$

做 1D-DCT 时

$$\mathbf{F} = \mathbf{G}\mathbf{f}$$

2D-DCT 时

$$\mathbf{F} = \mathbf{G}\mathbf{f}\mathbf{G}^T$$

2D-DCT 实际上等价于做两次 1D-DCT

$$\mathbf{F} = \mathbf{G}_{col}[\mathbf{G}_{row}[\mathbf{f}]^T]^T$$

1.2 Results

1.2.1 Comparison between two trials

首先是两个 trials 对比见Figure 1，为了便于观察我对 DCT 的结果取了 log，下同。

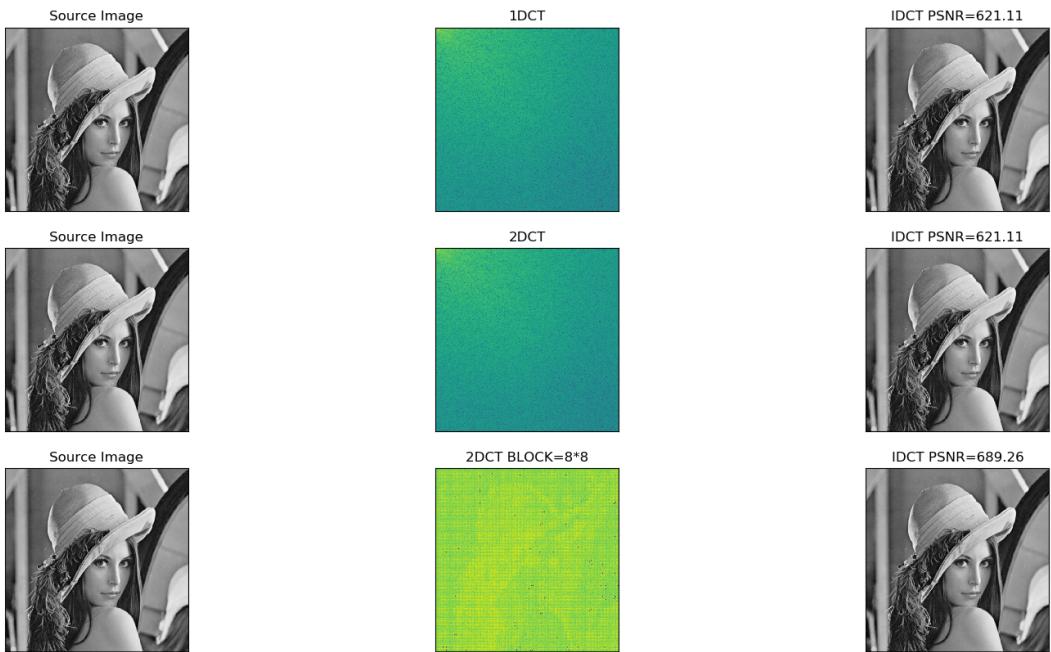


Figure 1: 1D-DCT and 2D-DCT

第一行和第二行分别是在全图上做 1DCT 和 2DCT, PSNR 值相等, 可见在相同大小的图上做 2 次 1DCT 和 1 次 2DCT 是等价的, 第三行是 BLOCK SIZE 为 8×8 的 2DCT, 可见 BLOCK SIZE 的大小这个变量是会影响 PSNR 的。

当保留部分系数, 如 $\frac{1}{4}$ 时, 1DCT 先在行方向上保留 $\frac{1}{2}$, 再到列方向上保留 $\frac{1}{2}$ 见 Figure 2

2



Figure 2: 1D-DCT(1/4)

和 2DCT 直接保留 $\frac{1}{4}$ 的对比见 Figure 3

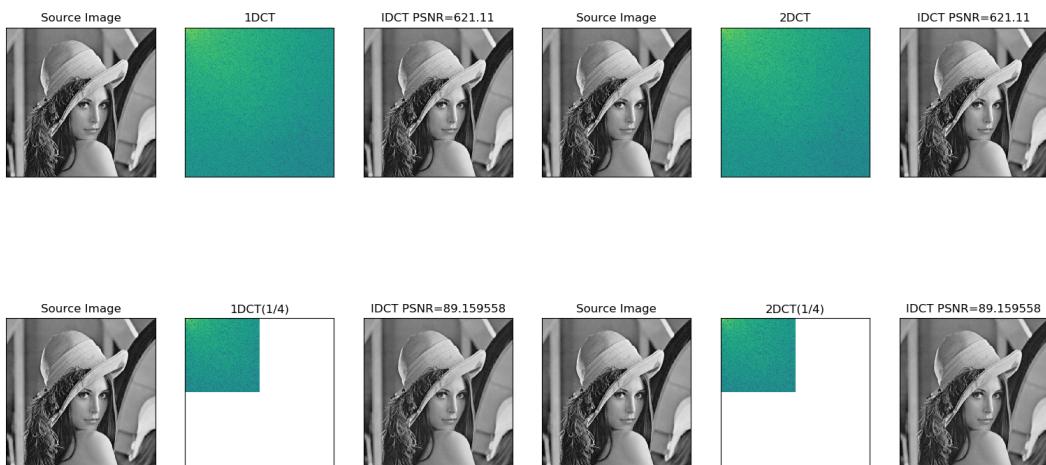


Figure 3: 1D-DCT(1/4) and 2D-DCT(1/4)

二者最终的 PSNR 在 8 位有效数字的精度下是相等的, 可以认为 1D-DCT 和 2D-DCT 在 8 位有效数字的精度下, 保留部分系数的结果也是等价的。

1.2.2 Different Coefficients

这里用 BLOCK SIZE 为 8×8 的 2DCT 来进行实验, 见 Figure 4

可见随着保留下来的系数的减少, PSNR 下降了且 DCT 的结果在观感上变得更稀疏了, IDCT 的结果在观感上变得更加模糊了, 这是符合逻辑的。

1.2.3 Different Block Size

这里用保留全部系数的 2DCT 来进行实验, 见 Figure 5

有一个实现上的细节是当原图边长不能整除 Block 边长时可以将原图 padding 到刚好整除, 做完 DCT 和 IDCT 后再裁剪回原来的大小即可。

相应的折线图见 Figure 6

可见总体上随着 Block Size 的增大, PSNR 是会下降的。特别地, 当 Block Size 为 1×1 时做的是 pixelwise 的 DCT, 此时 MSE 为 0, PSNR 为 inf, 当 Block Size 为原图尺寸时,

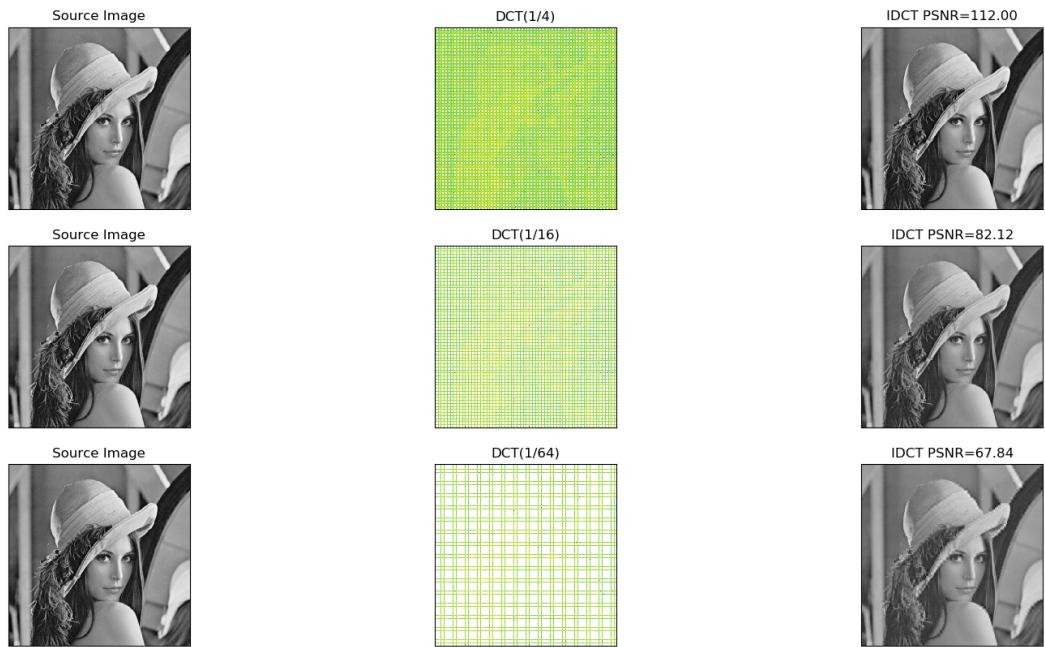


Figure 4: Different Coefficients

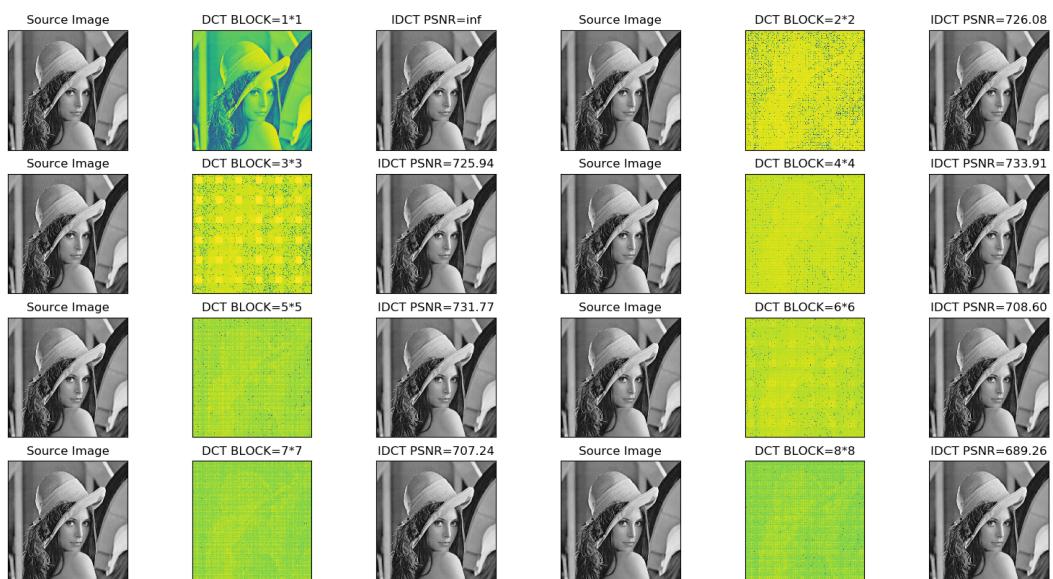


Figure 5: Different Coefficients

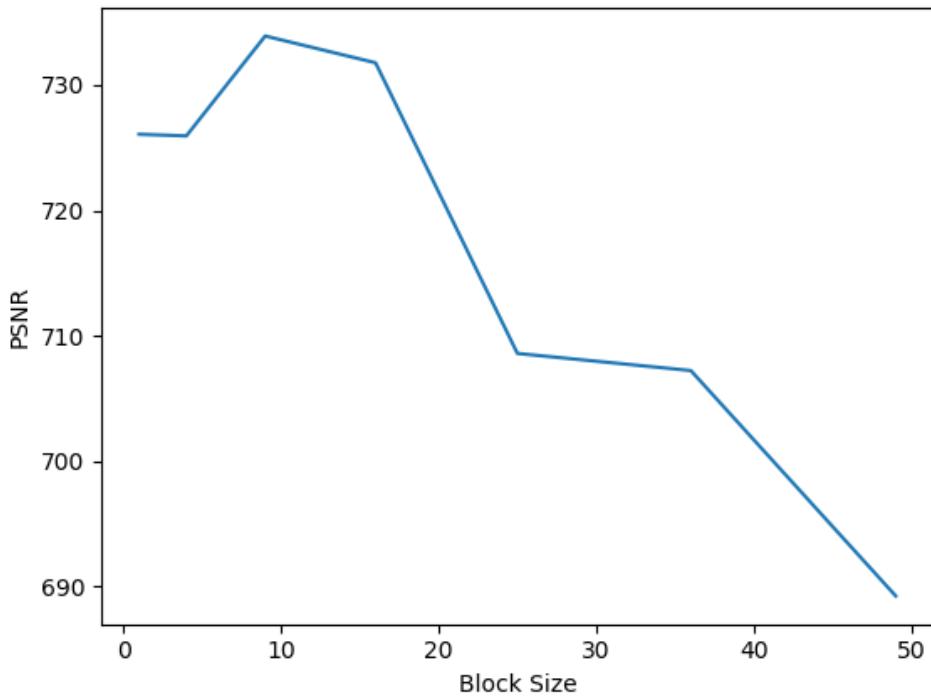


Figure 6: Different Coefficients

也即Figure 1中对全图做的2DCT, PSNR取得最小(或接近最小),这些结果是符合逻辑的。

2 Quantization

2.1 Principle

公式很简洁, Quantization 时

$$C_q(u, v) = \text{Round}\left[\frac{C(u, v)}{Q(u, v)}\right]$$

Dequantization 时乘上 $Q(u, v)$ 即可

本次实验一共给了 3 种 $Q(u, v)$, 见Figure 7

Quantization Table: Luminance							
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

Quantization Table: Luminance							
2	1	1	2	3	5	6	7
1	1	2	3	4	8	9	8
2	2	2	3	6	8	10	8
2	2	3	4	7	12	11	9
3	3	8	11	10	16	15	11
3	5	8	10	12	15	16	13
7	10	11	12	15	17	17	14
14	13	13	15	15	14	14	14

Quantization Table: Luminance							
2	1	1	2	3	5	6	7
1	1	2	2	3	7	7	7
2	2	2	3	5	7	8	7
2	2	3	3	6	10	10	7
2	3	4	7	8	13	12	9
3	4	7	8	10	12	14	11
6	8	9	10	12	15	14	12
9	11	11	12	13	12	12	12

(a) Jpeg Standard

(b) Canon IXUS 60 (fine)

(c) Nikon CoolPix L12 (fine)

Figure 7: Different $Q(u, v)$

2.2 Results

2.2.1 With and Without Quantization

先用 Jpeg Standard 对比使用和不使用 Quantization 的结果，见[Figure 8](#)

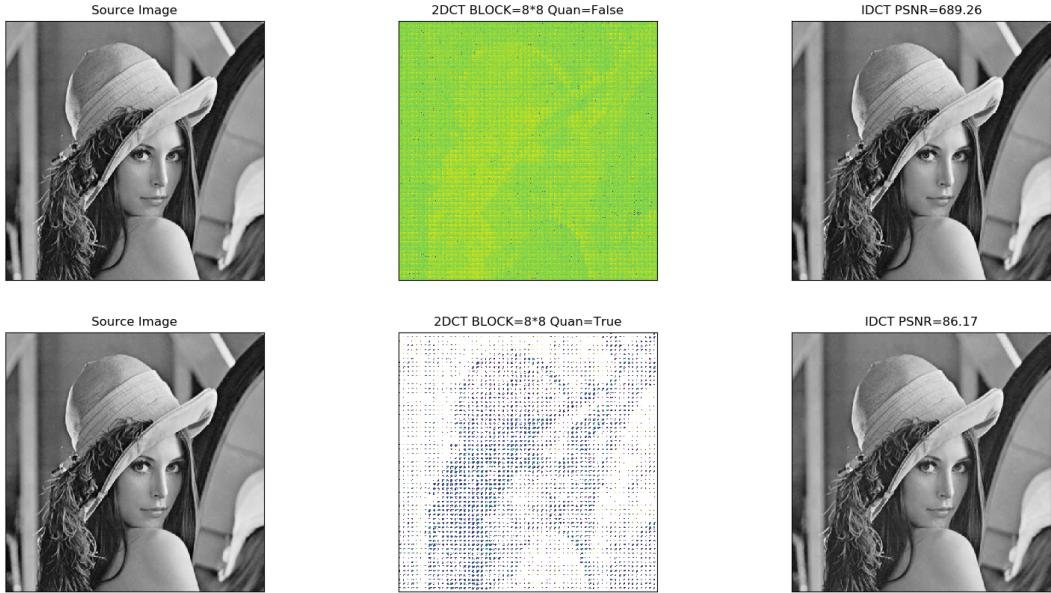


Figure 8: With and Without Quantization

可见 Quantization 后，DCT 明显稀疏了很多，PSNR 也下降了不少，这是符合逻辑的。

2.2.2 Different A and Q

insights 部分的不同 $a^*Q(0.1 < a < 2)$ 和不同 Q 矩阵的两个任务我整合在了一起见[Figure 9](#)，其中 a 取值时步幅为 0.1

此外对于每种 Q，我单独把 a 等于 0.5, 1.0 和 1.5 的结果展示出来见[Figure 10](#)，以便获得更直观的感受。图片的组织方式是各行之间区别在于使用的 Q，从上到下依次是 Standard Jpeg, Canon 和 Nikon，每一行内则包括三组 a 分别等于 0.5, 1.0 和 1.5 的 DCT 和 IDCT 结果

A 单独分析 a 系数，从折线图可见随着 a 的增加 PSNR 是会下降的，从抽取的样例上直观来看，随着 a 的增加（同一行内从左到右），DCT 变得更加稀疏，IDCT 的结果观感上变得更加模糊

Q 单独分析使用的 Q 矩阵，从折线图可知 PSNR 值最大的是 Nikon, Canon 与 Nikon 相差不大，而 Standard Jpeg 则和前面两种差得比较远。从抽取的样例直观上来看（同一列从上到下），可见 Nikon 和 Canon 的 DCT 图明显更稀疏，仔细观察 IDCT 的结果会比 Standard Jpeg 的 IDCT 结果更清晰。

我推测 Nikon, Canon 这两个相机用的 Q 矩阵相比 Standard Jpeg 的 Q 矩阵，会更加注重图片的质量，也即保留更多的 DCT 信息，以求更高的 PSNR，而 Standard Jpeg 则更侧重图片的压缩这一本职功能，所以会保留更少的 DCT 信息，以求更高的压缩率。

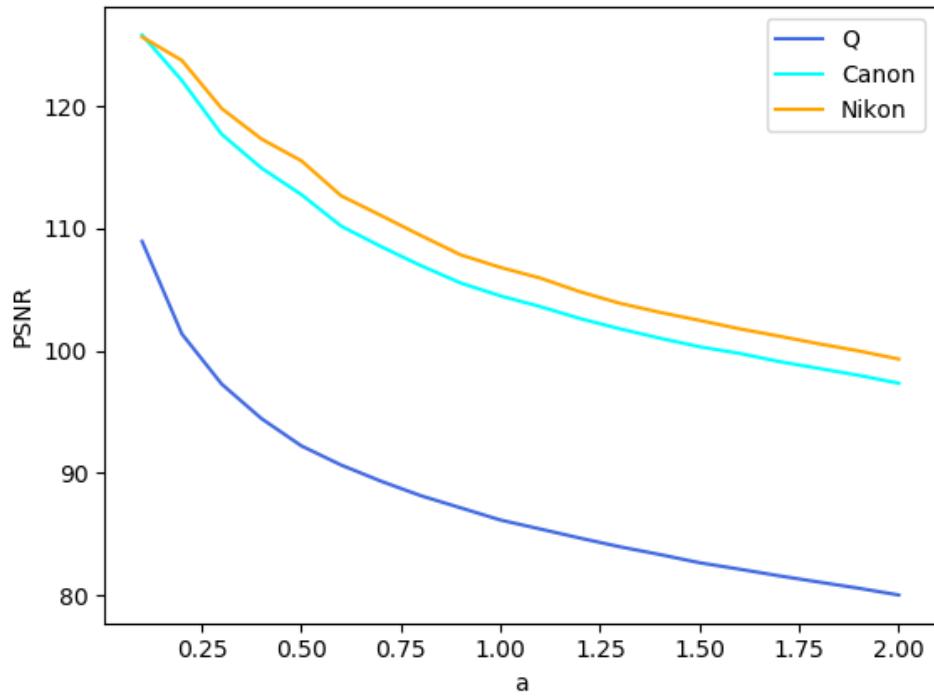


Figure 9: Different A and Q

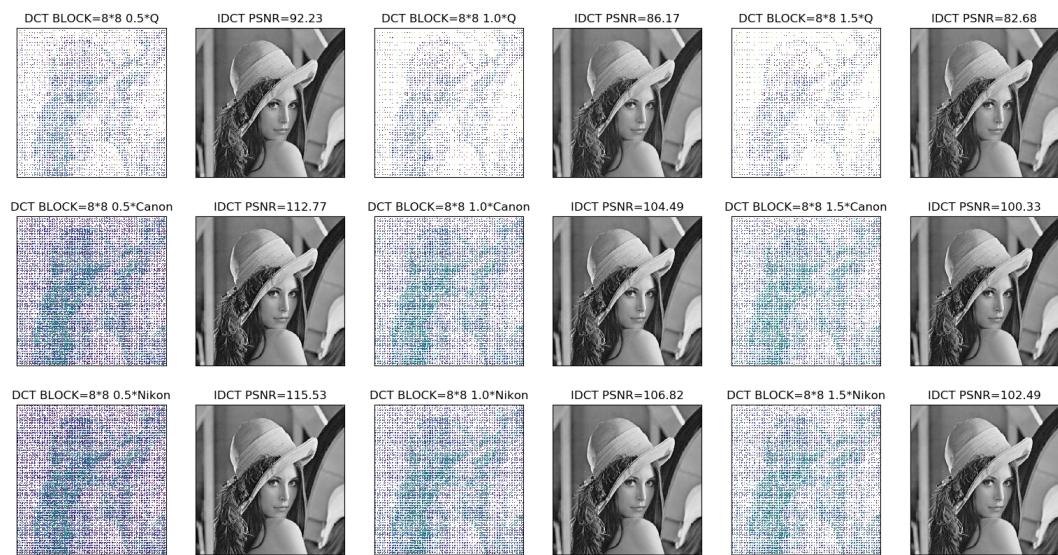


Figure 10: Some examples with different A and Q

3 Conclusion

总的来说这次小作业和前面三次小作业相比最大的区别在于将图像的处理从时域转到了频域，对我个人来说是很开阔视野的。