

Report for Image Compression

计 76 张翔 2017011568

2020 年 4 月 23 日

1 DCT with Different Settings

1.1 1D vs 2D-DCT

容易发现，对行与列分别做 1D-DCT 与直接做 2D-DCT 得到的结果是相同的，在取 $1/n^2$ 个系数的情况下该结论也成立。

如果逐行做 DCT，可以得到

$$c_u(v) = \alpha(v) \sum_{y=0}^{N-1} f(x, y) \cos\left(\frac{(2y+1)v\pi}{2N}\right), v = 0, 1, \dots, N-1$$

取每行系数的前 $1/n$ ，只需要令

$$c'_u(v) = \begin{cases} c_u(v), v = 0, 1, \dots, \lfloor N/n \rfloor - 1 \\ 0, otherwise \end{cases}$$

逐列做 DCT，得到

$$c(u, v) = \alpha(u) \sum_{x=0}^{N-1} c'_u(v) \cos\left(\frac{(2x+1)v\pi}{2N}\right), u = 0, 1, \dots, N-1$$

取每列系数的前 $1/n$ ，得到

$$c'(u, v) = \begin{cases} c(u, v), u = 0, 1, \dots, \lfloor N/n \rfloor - 1 \\ 0, otherwise \end{cases}$$

即

$$c'(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos\left(\frac{(2y+1)v\pi}{2N}\right) \cos\left(\frac{(2x+1)v\pi}{2N}\right), u, v = 0, 1, \dots, \lfloor N/n \rfloor - 1$$

其中 $\alpha(x) = \sqrt{(2 - \delta(x, 0)) / N}$ 。上述结果正好就是 2D-DCT 后取矩阵的左上角一块 $\lfloor N/n \rfloor \times \lfloor N/n \rfloor$ 的子阵。通过程序也容易验证上述结论。

1.2 Number of Coefficients Retained

直观感受应是保留的 DCT 系数越少，结果图的信息量会越少，而由于保留的是低频区域的系数，相应地，图片会损失高频细节。如果保留所有系数，此时的 DCT 是可逆变换，可以完整恢复原图。例如，在将原图分割为 8×8 块的情况下，使用 $1/4, 1/16, 1/64$ 系数后得到的结果如下



(a) 原图



(b) 1/4



(c) 1/16



(d) 1/64

图 1: 保留不同比例 DCT 系数后得到的结果

保留 1/4 系数时，观感与原图接近，但细节部分略有缺失；保留 1/64 系数时，相当于只保留 8×8 块中的 DC 分量，使得每个小块中为纯色，最终呈现 mosaic 的效果。

上述三者的 PSNR 分别为 34.71, 28.19, 23.66，可看出随着保留系数的减少，PSNR 有明显下降。

1.3 Block Size

保留相同比例 DCT 系数时，结果图所保留的信息量应该相差不大，但观感可能有一定的差异，因为不同块之间可能出现不平滑的现象，尤其是保留系数较少的情况下（如上图 1/64 的情形）。下面保留 1/16 系数，块大小为 $8 - 512$ 之间变化



(a) 8×8



(b) 32×32



(c) 128×128



(d) 512×512

图 2: 不同块大小得到的结果

表 1: 不同块大小的 PSNR

Block Size	PSNR
8	28.19
16	28.99
32	29.43
64	29.66
128	29.78
256	29.85
512	29.88

可以看出，块大小增大时，PSNR 有一定的增加，但增加量较为有限。较小的块大小在保留的 DCT 系数较少时出现了明显的 blocky effect(上述 8×8 的样例)，而较大的块大小得到的图片更为平滑，但出现了类似衍射条纹的 artifacts。

2 Quantization

记 JPEG 的量化矩阵 \mathbf{Q}_J ，令 $a \in [0.1, 2]$ ，使用 $a\mathbf{Q}_J$ 量化上述 8×8 分块 DCT 后的图片，可得到如下结果

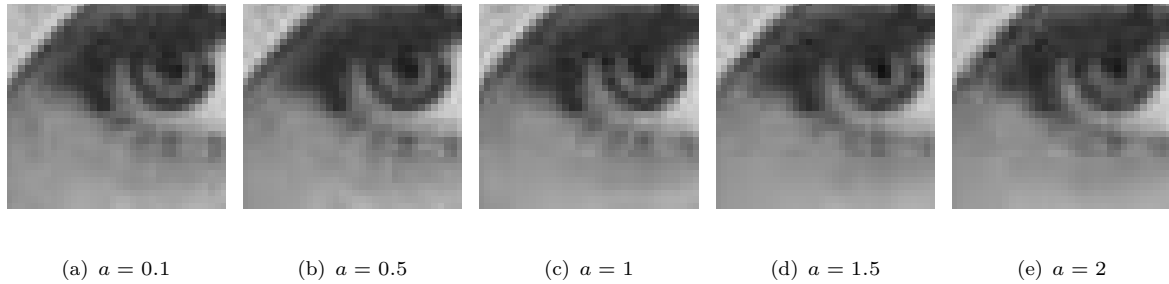


图 3: 不同 a 取值下得到的结果

可以看出，随着 a 的增大，量化矩阵的系数相应增大，对应 DCT 系数量化后的值会更加 coarse，使得结果图出现更加明显的 artifacts。

当 a 变化时，PSNR 的变化趋势如下 (a 取对数坐标)

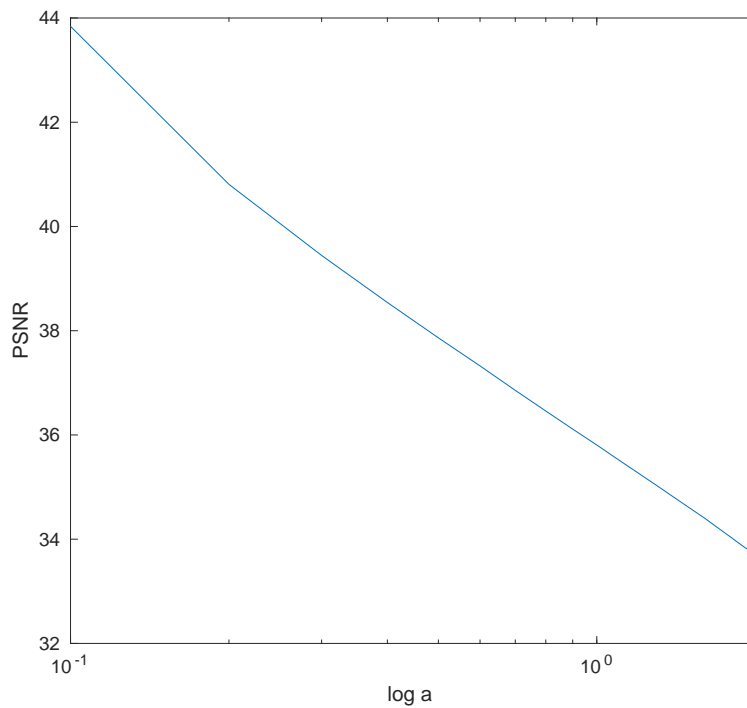


图 4: PSNR~ a 曲线

可以看出, PSNR 与 $\log a$ 呈近似线性关系。如果使用最小二乘拟合, 可以得到 $PSNR = -3.21 \log a + 35.83$, $R^2 = 0.9949$ 。考虑 $PSNR = 10 \log \left(\frac{255^2}{MSE} \right)$, 可以得到 $MSE = \left(\frac{255^2}{e^{3.523}} \right) a^{0.321}$, 说明在 a 较小时, MSE 与 PSNR 对 a 较为敏感。如果要设置一个用于表示 JPEG 图像品质的因子 γ ($\gamma \in (0, 1]$), 用于让用户指定图片压缩后的品质, 可以考虑 $\gamma < 0.5$ 时令 $a \propto \frac{1}{\gamma}$, 而 $\gamma > 0.5$ 时令 $a = C_1 - C_2 \gamma$, 在此基础上生成对应品质所需要的量化矩阵。

如果使用 Nikon 和 Canon 的量化矩阵, 结果如下



(a) Nikon



(b) Canon

图 5: Nikon vs Canon Quantization

Nikon 量化矩阵得到 $PSNR=42.93$, 量化后 DCT 系数中非零元占比 0.3979; Canon 量化矩阵 $PSNR=42.00$, 量化后 DCT 系数为 0.3568。这个结果比上面 JPEG 默认的量化矩阵得到的 $PSNR=35.83$ 要更好一些。