

# JPEG 图像编解码与压缩质量分析

分工明细：

唐晓阳：代码实现与优化、答辩、ppt 修改（占比 70%）

朱紫阳：答辩 PPT 制作与大作业报告（占比 30%）

## 一、任务问题：实现 JPEG 图像编解码

### 1.基础：

- （1）将输入图像从 RGB 空间转换为 YCbCr 空间，对色度分量 Cb 和 Cr 进行下采样
- （2）对每个 8×8 的像素块执行离散余弦变换 (DCT)，将图像从空间域变换到频率域，最终对 DCT 系数进行量化和熵编码。同时实现 JPEG 解码

2.进阶：自定义量化表及实验（如加强高频压缩、保留边缘细节），给出不同量化方法对图像压缩质量的影响，对图像进行分析（如 PSNR/SSIM 等）

## 二、原理分析

数字图像爆炸式增长使传统存储扩容方案难以为继，网络带宽限制更制约高清内容传输。1980 年代全球影像通信需求激增，推动国际组织联合成立 JPEG 专家组，通过消除像素冗余与视觉冗余，制定首个通用静态图像压缩标准，解决存储成本与传输效率的核心矛盾，为多媒体数字化浪潮铺平道路。

JPEG 革命性采用 DCT 与人眼视觉模型，将图像转换至频域后量化压缩高频信息，辅以熵编码优化数据流。其“感知优先”策略在 10-20 倍压缩率下仍保持视觉可接受质量，直接推动数码影像普及，降低早期互联网图像传输门槛，奠定现代有损压缩技术框架，影响延续至视频编码等衍生领域。

以下是相关的关键技术原理分析：

### 1. YCbCr 色彩空间转换与下采样

#### （1）原理：

人眼感知特性：在有损压缩中，首先要做的事情就是“把重要的信息和不重要的信息分开”。由于人眼的构造,视网膜对亮度（Y 分量）敏感，是重要信息，对色度（Cb/Cr）分辨率需求低，YCbCr 能很好的将亮度与色差分开。

RGB→YCbCr 转换（数学公式）：

$$Y=0.299R+0.587G+0.114B$$

$$Cb=128-0.1687R-0.3313G+0.5B$$

$$Cr=128+0.5R-0.4187G-0.0813B$$

实现效果：



图一

下采样策略（4:2:0）：亮度 Y 保留全分辨率，色度 Cb/Cr 在水平和垂直方向均下采样至 1/2

（2）作用：直接减少 50%的数据量，且不影响主观视觉质量

## 2.离散余弦变换（DCT）

（1）能量集中性：将  $8 \times 8$  像素块从空域转换到频域，能量集中于左上角低频区域（直流和低频系数）

（2）高频特性：右下角高频分量对应图像细节（如边缘/纹理），对视觉贡献较小

（3）数学实现：

$$F_{mn}(u, v) = \frac{2}{N} C(u) C(v) \sum_{x=0}^7 \sum_{y=0}^7 B_{mn}(x, y) \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16}$$

其中  $C(0) = 1/\sqrt{2}$ ,  $C(k \neq 0) = 1$

## 3. 量化与心理视觉模型

（1）量化表设计：低频区域（左上）使用小步长→ 保留主要能量，高频区域（右下）使用大步长→ 主动丢弃细节

```
1. LTable = np.array([[16, 11, 10, 16, 24, 40, 51, 61],
2.                    [12, 12, 14, 19, 26, 58, 60, 55],
3.                    [14, 13, 16, 24, 40, 57, 69, 56],
4.                    [14, 17, 22, 29, 51, 87, 80, 62],
5.                    [18, 22, 37, 56, 68, 109, 103, 77],
6.                    [24, 35, 55, 64, 81, 104, 113, 92],
7.                    [49, 64, 78, 87, 103, 121, 120, 101],
8.                    [72, 92, 95, 98, 112, 100, 103, 99]], dtype=np.uint8)
9.
```

10.	CTable = np.array([[17, 18, 24, 47, 99, 99, 99, 99],
11.	[18, 21, 26, 66, 99, 99, 99, 99],
12.	[24, 26, 56, 99, 99, 99, 99, 99],
13.	[47, 66, 99, 99, 99, 99, 99, 99],
14.	[99, 99, 99, 99, 99, 99, 99, 99],
15.	[99, 99, 99, 99, 99, 99, 99, 99],
16.	[99, 99, 99, 99, 99, 99, 99, 99],
17.	[99, 99, 99, 99, 99, 99, 99, 99]], dtype=np.uint8)

## (2) 量化操作:

$$QF_{mn}(u, v) = \text{round} \left( \frac{F_{mn}(u, v)}{Q(u, v)} \right)$$

$Q(u, v)$  为标准量化表值, 满足  $Q(u, v) \propto \sqrt{u^2 + v^2}$  (高频区步长递增)

(3) 人眼适应性: 对高频噪声不敏感 → 允许更大的量化失真, 对低频亮度变化敏感 → 严格保护低频分量

## 4. 熵编码与统计冗余消除

(1) 之字形扫描 (Zigzag Scan): 将  $8 \times 8$  量化系数按频率从低到高排列为 一维序列, 产生长串零值系数 → 便于游程编码

(2) 霍夫曼编码: 对 (零游程长度, 非零系数值) 组合分配短码字, 高频区域的连续零值通常合并为 EOB 标记

(3) 压缩增益: 典型图像熵编码可再减少 20%-40% 的数据量

## 三、算法设计与分析

### 1. 算法流程概述

(1) 压缩: RGB 转 YCbCr → 色度下采样 → 分块 DCT → 量化 → ZigZag 扫描 → RLE + 哈夫曼编码 → 封装 JPEG 文件头。

(2) 解压缩: 解析 JPEG 头 → 哈夫曼解码 → 逆 ZigZag → 逆量化 → 逆 DCT → 色度上采样 → YCbCr 转 RGB。

### 2. 关键算法设计分析

#### (1) 颜色空间转换:

RGB 转 YCbCr: 使用标准转换矩阵分离亮度和色度分量, 利用人眼对亮度敏感的特性, 后续对色度 (Cb、Cr) 进行  $2 \times 2$  下采样, 减少数据量。

YCbCr 转 RGB：通过逆矩阵重建原始颜色空间。

(2) 离散余弦变换 (DCT)：对每个  $8 \times 8$  块应用 DCT，将空域信号转换到频域，能量集中在低频区域，高频部分可被量化舍去。

(3) 量化与 ZigZag 扫描：

量化表调整：根据压缩质量参数缩放标准量化表，质量越低，量化步长越大，压缩率越高。

ZigZag 扫描：将  $8 \times 8$  矩阵按 Z 字形排列成一维序列，使低频系数优先排列，便于后续编码。

(4) 哈夫曼编码：

DC 系数：使用差分编码（当前块 DC 值与前一块差值），减少冗余。

AC 系数：对连续零值进行游程编码 (Run-Length Encoding)，结合哈夫曼表生成变长码字。

码表生成：基于标准 JPEG 哈夫曼表，按码长分配二进制编码。

(5) JPEG 文件封装：添加文件头 (SOI、APP0、DQT、SOF0、DHT、SOS) 和压缩数据，处理特殊标记（如 FF 后补 00）。

(6) 解压缩与重建：逆过程解析哈夫曼码流，恢复量化后的 DCT 系数，逆 DCT 重建图像，色度上采样后转换回 RGB。

## 四、实验结果与分析

### 1. 压缩效果

在量化过程中，若对高频分量压缩过高会导致：(1) 出现色斑；(2) 在高频区域会出现较为严重的失真（如下图  $qs=99$  时）；(3)  $quality\_scale = 10$  时，能够做到基本无失真。





2.总体效果分析：

实验数据表明（由表一），量化表设计直接决定 JPEG 压缩的性能边界。标准量化表以中等细节损失（PSNR 28.19 dB，SSIM 0.9665）实现高压缩率，适用于通用场景；增强边缘保留策略通过精细量化高频分量（步长缩小 50%），显著提升细节保留能力（PSNR 32.02 dB，SSIM 0.9810），但压缩率显著降低，适合图像存档等高保真需求；增加高频压缩则以牺牲视觉质量为代价（PSNR 27.29 dB，块效应显著），将压缩率推至极高水平，仅适用于带宽严苛的实时传输。量化步长与高频分量的交互是核心矛盾：高频分量控制本质是细节保留与压缩效率的博弈。

策略	文本文 件大小	PNG 文 件大小	压缩 比	PSNR	SSIM
标准 量化	270638 字节	1517849 字节	0.18	28.19dB	0.9665
增强 边缘 保留	514970 字节	1517849 字节	0.34	32.02dB	0.9810
加强 高频 压缩	159278 字节	1517849 字节	0.10	27.29dB	0.9223

表一

## 五、大语言模型使用说明

以下内容使用了大模型辅助

1. 代码辅助解读（报告撰写人）
2. 4:2:0 下采样的具体做法
3. 对二进制码进行扫描封装的结构
4. 一些函数库的使用