

JPEG 图象文件格式分析

许刚 (中国科学院软件研究所 100080)
 廖斌 (武汉水利电力大学计算机科学系 430072)
 李承毅 (湖南省公安厅 410001)

摘要:基于 DCT 变换的 JPEG 静态图象编码方案已成为国际标准。本文对 JPEG 方案的实现原理进行描述,同时给出 JPEG 文件的交换格式进行分析。

关键词:JPEG 方案 JPEG 交换文件格式

1. 简介

JPEG(Joint Photographic Experts Group)是由国际标准化组织(ISO: International Standardization Organization)和国际电话电报咨询委员会(CCITT: Consultation Committee of the International Telephone and Telegraph)为静态图象压缩制定的第一个国际数字图象压缩标准。JPEG 的压缩模式包括以下几种:

顺序式编码(sequential encoding)

一次将图象由左到右,由上到下顺序处理。

递增式编码(progressive encoding)

当图象传输的时间较长时,可将图象分数次处理,以较模糊至清晰的方式来传送图象。

无失真编码(lossless encoding)

压缩后经过还原的图象与原始图象一模一样。

阶梯式编码(hierarchical encoding)

图象以多种分辨率来压缩,目的是为了具有高分辨率的图象能在较低分辨率的设备上显示。

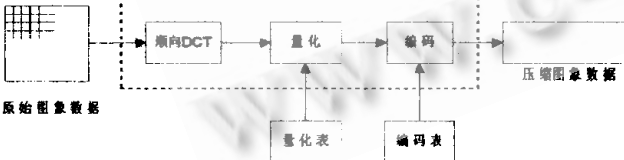


图1 JPEG 的编码过程

在图象编码和解码的处理过程中,包括基于 DCT 变换的有失真压缩方法和使用 DPCM 预测器的无失真压缩方法。本文介绍的是基于 DCT 变换的有失真图象压缩方法,压缩和解压是以 8x8 的象素块作为基本处理单

元。图象编码和解码的流程图如图 1、2:

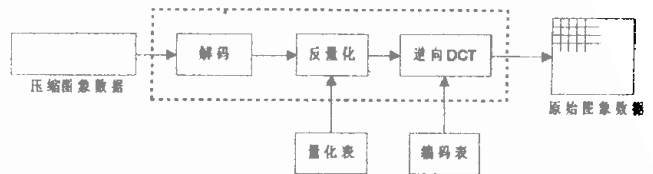


图2 JPEG 的解码过程

2. 离散余弦变换

离散余弦变换 (Discrete Cosine Transform, 简称 DCT)是指将一组光强数据转换成频率数据,以得知强度变化的情形。由于人的视觉系统对图象的低频数据比对高频数据具有更高的敏感度,因此若对高频数据进行量化处理,其还原所得数据虽与原始数据有些差异,但人的眼睛是不容易分辨的。由此可知,低频数据比高频数据重要得多。

$$FDCT: \hat{f}[u, v] = \frac{1}{4} C_u C_v \sum_{x=0}^7 \sum_{y=0}^7 f[x, y] \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16} \quad (1)$$

$$IDCT: f[x, y] = \frac{1}{4} \sum_{u=0}^7 \sum_{v=0}^7 C_u C_v \hat{f}[u, v] \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16} \quad (2)$$

$$\text{其中 } \begin{cases} C_u C_v = 1/\sqrt{2} & u, v = 0 \\ C_u C_v = 1 & \text{其它} \end{cases}$$

在压缩时,将原始图象数据分成许多数据单元,然后经过零均值化将每个字节的值从 0~255 转为 -128~+127,在作 DCT 变换(Forward DCT, 简称 FDCT)。FDCT 将每个数据单元的值转换为 64 个 DCT 系数,第一个系

数称作直流系数(DC Coefficient),而其他 63 个系数则称为交流系数(AC Coefficient)。在解码时,经过 DCT 逆变换(Inverse DCT,简称 IDCT)将 64 个 DCT 系数转回为 8x8 个像素块,然后形成完整的图象数据,上面就是 JPEG 所使用的二维 DCT 公式:

3. 量化

将 DCT 系数按比例缩小,并取其最接近的整数值的处理过程称为量化(Quantization);而将量化过后的数据再按原比例恢复为最接近原始 DCT 系数值的处理过程称为逆量化(Dequantization)。下面给出 JPEG 所提供的量化表和量化公式:

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

表 1 亮度(luminance)量化表

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

表 2 色调(Chrominance)量化表

$$P_q[u, v] = \text{Round} \left\{ \frac{P[u, v]}{Q[u, v]} \right\} \quad (3)$$

$$R[u, v] = P_q[u, v] * Q[u, v] \quad (4)$$

4. 色彩系统

JPEG 采用的是 YCbCr 色彩系统,它是从 YUV 色彩系统衍生出来的。RGB 和 YCbCr 的关系如下:

$$\begin{aligned} Y &= 0.2990R + 0.5870G + 0.1140B \\ Cb &= -0.1687R - 0.3313G + 0.5000B + 128 \\ Cr &= 0.5000R - 0.4187G - 0.0813B + 128 \end{aligned} \quad (5)$$

$$\begin{aligned} R &= Y + 1.40200(Cr - 128) \\ G &= Y - 0.34414(Cb - 128) - 0.71414(Cr - 128) \end{aligned}$$

(6)

$$B = Y + 1.77200(Cb - 128)$$

5. 数据抽样

从视觉上来看,人类的视觉系统对低频的数据对高频的数据具有更高的敏感度。事实上,人类的眼睛对明视度的改变比对色彩的改变要敏感得多。在 JPEG 的应用上,一般来说都是只处理灰度和真彩图象。真彩图象由 Y、Cb 和 Cr 三个颜色成分(Component)组合而成,而灰度图象只有 Y 一个成分。Y 代表的是图象的明视度,而 Cb 和 Cr 代表图象的色调。因此, Y 成分的数据相对重要, Cb 成分和 Cr 成分的数据相对不重要,那么就可以只取 Cb 和 Cr 的部分数据来处理,以增加压缩的比例。这样的过程称作部分数据取样(Subsampling)。

这里介绍两种常用的部分取样方式: YUV411 和 YUV422,它们所代表的意义是 Y、Cb 和 Cr 三个成分的数据取样比例。举例来说,如果 Y 取四个数据单元,即水平取样因子(Horizontal Sampling Factor)H_y乘以垂直取样因子(Vertical Sampling Factor)V_y的值等于 4(H_y * V_y = 4),而 Cb 和 Cr 各取一个数据单元,即 H_{cb} * V_{cb} = 1,这样的部分取样称为 YUV411。由于部分取样的关系,被压缩的数据量就大大减少。但这并不意味着部分取样为 YUV911, YUV1611...时,压缩比例就可以更高。因为将会使图象的质量下降。因此 JPEG 定义了最小编码单位(Minimum Coded Unit,简称 MCU),其值不能大于 10:

$$\sum_{i=1}^{Ns} H_i * V_i \leq 10 \quad (7)$$

Ns: 图象成分的个数

H_i: 成分 i 的水平抽样因子

V_i: 成分 i 的垂直抽样因子

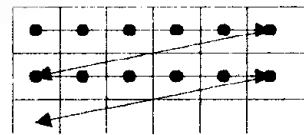


图 3 非交错式数据排列

在编码过程中,由数个数据单元组合而成的集合称为 MCU。在图象扫描中若仅有一个成分,则数据是以非交错的方式处理,MCU 相当于一个数据单元。若扫描中有数个成分,则数据是以交错的方式来处理,而 MCU 就是由数个不同成分中的数据单元组成,其排列方式与 H_i 和 V_i 有密切关系。

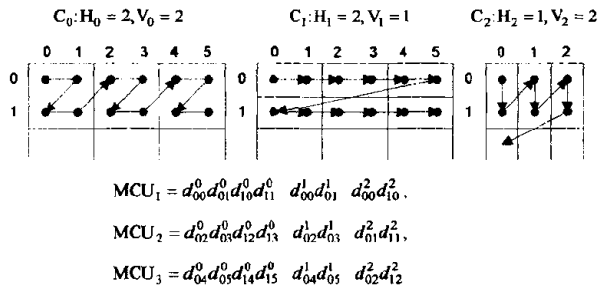


图4 交错式数据排列

4	-15..-8,8..15
5	-31..-16,16..31
6	-63..-32,32..63
7	-127..-64,64..127
8	-255..-128,128..255
9	-511..-256,256..511
10	-1023..-512,512..1023
11	-2047..-1024,1024..2047

图5 差值分类表

6. 熵编码

原始图象数据经过离散余弦变换和量化以后, 然后进行熵编码(Entropy Coding)。JPEG的熵编码是利用统计模型将描述符号(Descriptors)转换成为符号(Symbols), 并且指出这些符号的概率或是码字(Code words), 前者是使用算术编码(Arithmetic coding)来完成, 而后者则是通过霍夫曼编码(Huffman coding)完成。

首先将量化过的64个DCT系数转换成中间符号, 再分别以可变长度码(Variable-Length Code, 简称VLC)和可变长度整数(Variable-Length Integer, 简称VLI)来进行数据编码, 其中DC和AC是分开的:

(VLC)	(VLI)	(VLC)	(VLI)
symbol - 1	symbol - 2	symbol - 1	symbol - 2
(SSSS)	(VVVV)	(RRRR/SSSS)	(VVVV)

由DC差值转换而成的中间符号 由AC差值转换而成的中间符号

(1)DC系数。DC是采用一种称为差值脉冲编码调制(Differential Code Modulation, 简称DPCM)的差值编码方法, 也就是在同一个图象分量中取每个DC值与前一个DC值(PRED)的差值(DIFF)来编码:

$$DIFF = DC - PRED \quad (8)$$

这个差值被划分为12类, 用SSSS表示。然后用Huffman编码对这12类符号进行编码。分类方法如图5所示:

SSSS	差值
0	0
1	-1, 1
2	-2, -1, 1, 2
3	-7..-3, 3..7

对每一类而言, 除去SSSS=0, 需要额外的位数来表示实际的差值。SSSS表示了所需的额外位数。如果DIFF为正数, 则DIFF的低SSSS位被传输。如果DIFF为负数, 则(DIFF-1)的低SSSS位被传输。

(2)AC系数。AC系数首先要按照“之”字形排列顺序处理, 如图6所示:

0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63

图6 Zig-Zag顺序

每个非零AC系数由(RRRR/SSSS)和(VVVV)来表示, (RRRR/SSSS)是一个字节, RRRR指字节的高4位, 它表示当前非零系数相对前一个非零系数的位置。SSSS指字节的低4位, 它表示的含义与前相同。由于系数为零的游程可能超过15, 因此用值(RRRR/SSSS) = X 'F0' 表示。此外, 当块中的剩余系数全部为零时, (RRRR/SSSS) = '0000000' 用来表示块的结束(EOB)。

7. JPEG文件的交换格式

JPEG委员会在制定JPEG规格时, 定出了各种标记码(marker)来识别图象压缩数据的相关信息。JPEG的每个标记码都是由两个字节组成, 其中高字节一定是0xFF, 而在每个标记之前也可以加上个数不限的填充字节0xFF。

(来稿时间:1998年7月)