

JPEG 文件格式实例详解

石志鑫 2009 年 8 月

Email:szhixin@gmail.com

一、JPEG 文件数据流

实例图片为 windows 自带的实例图片下的 Blue hills. jpg, 见附件中的实例用图。在 FlexHEX 中查看的数据流如下:

FF D8 FF E0 00 10 4A 46 49 46 00 01 02 01 01 2C 01 2C 00 00 FF ED 09 4C 50 68 6F 74 6F 73 68
6F 70 20 33 2E 30 00 38 42 49 4D 03 ED 0A 52 65 73 6F 6C 75 74 69 6F 6E 00 00 00 00 10 01 2C 00 00 00 01
00 01 01 2C 00 00 00 01 00 01 38 42 49 4D 04 0D 18 46 58 20 47 6C 6F 62 61 6C 20 4C 69 67 68 74 69 6E 67 20
41 6E 67 6C 65 00 00 00 00 04 00 00 00 78 38 42 49 4D 04 19 12 46 58 20 47 6C 6F 62 61 6C 20 41 6C 74 69 74
75 64 65 00 00 00 00 04 00 00 00 1E 38 42 49 4D 03 F3 0B 50 72 69 6E 74 20 46 6C 61 67 73 00 00 00 09 00 00
00 00 00 00 00 00 01 00 38 42 49 4D 04 0A 0E 43 6F 70 79 72 69 67 68 74 20 46 6C 61 67 00 00 00 00 01 00 00
38 42 49 4D 27 10 14 4A 61 70 61 6E 65 73 65 20 50 72 69 6E 74 20 46 6C 61 67 73 00 00 00 00 0A 00 01 00 00
00 00 00 00 00 02 38 42 49 4D 03 F5 17 43 6F 6C 6F 72 20 48 61 6C 66 74 6F 6E 65 20 53 65 74 74 69 6E 67 73
00 00 00 48 00 2F 66 66 00 01 00 6C 66 66 00 06 00 00 00 00 00 01 00 2F 66 66 00 01 00 A1 99 9A 00 06 00 00
00 00 00 01 00 32 00 00 00 01 00 5A 00 00 00 06 00 00 00 00 00 01 00 35 00 00 00 01 00 2D 00 00 00 06 00 00
00 00 00 01 38 42 49 4D 03 F8 17 43 6F 6C 6F 72 20 54 72 61 6E 73 66 65 72 20 53 65 74 74 69 6E 67 73 00 00
00 70 00 00 FF
FF
FF
FF
FF
00 10 00 00 00 01 00 00 02 40 00 00 02 40 00 00 00 00 38 42 49 4D 04 1E 0D 55 52 4C 20 6F 76 65 72 72 69 64
65 73 00 00 00 04 00 00 00 00 38 42 49 4D 04 1A 06 53 6C 69 63 65 73 00 00 00 00 75 00 00 00 06 00 00 00 00
00 00 00 00 00 00 02 58 00 00 03 20 00 00 00 0A 00 55 00 6E 00 74 00 69 00 74 00 6C 00 65 00 64 00 2D 00 31
00 00 00 01 00
00 02 58 00
38 42 49 4D 04 11 11 49 43 43 20 55 6E 74 61 67 67 65 64 20 46 6C 61 67 00 00 00 01 01 00 38 42 49 4D 04 14
17 4C 61 79 65 72 20 49 44 20 47 65 6E 65 72 61 74 6F 72 20 42 61 73 65 00 00 00 04 00 00 00 02 38 42 49 4D
04 0C 15 4E 65 77 20 57 69 6E 64 6F 77 73 20 54 68 75 6D 62 6E 61 69 6C 00 00 05 A4 00 00 00 01 00 00 00 70
00 00 00 54 00 00 01 50 00 00 6E 40 00 00 05 88 00 18 00 01 FF D8 FFE0 00 10 4A 46 49 46 00 01 02 01 00 48
00 48 00 00 FF EE 00 0E 41 64 6F 62 65 00 64 80 00 00 00 01 FF DB 00 84 00 0C 08 08 08 09 08 0C 09 09 0C
11 0B 0A 0B 11 15 0F 0C 0C 0F 15 18 13 13 15 13 13 18 11 0C 0C 0C 0C 0C 0C 0C 0C 0C 0C 0C 0C 0C 0C 0C 0C
0C
0E 0E 14 14 0E 0E 0E 0E 14 11 0C 0C 0C 0C 0C 11 11 0C 0C 0C 0C 0C 0C 11 0C 0C 0C 0C 0C 0C 0C 0C 0C 0C
0C
03 11 01 FF DD 00 04 00 07 FF C4 01 3F 00 00 01 05 01 01 01 01 01 01 00 00 00 00 00 00 00 03 00 01 02 04 05
06 07 08 09 0A 0B 01 00 01 05 01 01 01 01 01 01 00 00 00 00 00 00 00 01 00 02 03 04 05 06 07 08 09 0A 0B 10
00 01 04 01 03 02 04 02 05 07 06 08 05 03 0C 33 01 00 02 11 03 04 21 12 31 05 41 51 61 13 22 71 81 32 06 14 91
A1 B1 42 23 24 15 52 C1 62 33 34 72 82 D1 43 07 25 92 53 F0 E1 F1 63 73 35 16 A2 B2 83 26 44 93 54 64 45 C2
A3 74 36 17 D2 55 E2 65 F2 B3 84 C3 D3 75 E3 F3 46 27 94 A4 85 B4 95 C4 D4 E4 F4 A5 B5 C5 D5 E5 F5 56
66 76 86 96 A6 B6 C6 D6 E6 F6 37 47 57 67 77 87 97 A7 B7 C7 D7 E7 F7 11 00 02 02 01 02 04 04 03 04 05 06
07 07 06 05 35 01 00 02 11 03 21 31 12 04 41 51 61 71 22 13 05 32 81 91 14 A1 B1 42 23 C1 52 D1 F0 33 24 62

[illegible]

F5 56 66 76 86 96 A6 B6 C6 D6 E6 F6 37 47 57 67 77 87 97 A7 B7 C7 D7 E7 F7 11 00 02 02 01 02
 04 04 03 04 05 06 07 07 06 05 35 01 00 02 11 03 21 31 12 04 41 51 61 71 22 13 05 32 81 91 14 A1
 B1 42 23 C1 52 D1 F0 33 24 62 E1 72 82 92 43 53 15 63 73 34 F1 25 06 16 A2 B2 83 07 26 35 C2
 D2 44 93 54 A3 17 64 45 55 36 74 65 E2 F2 B3 84 C3 D3 75 E3 F3 46 94 A4 85 B4 95 C4 D4 E4
 F4 A5 B5 C5 D5 E5 F5 56 66 76 86 96 A6 B6 C6 D6 E6 F6 27 37 47 57 67 77 87 97 A7 B7 C7 **FF**
DA 00 0C 03 01 00 02 11 03 11 00 3F 00 F4 D9 4A 54 77 04 DB BE F5 2D 35 38 99 CA 52 A1 21
 3C A5 4A E2 65 29 4A 8C A5 29 52 B8 99 4A 52 A3 29 4A 14 AE 26 5A A5 2A 32 94 A3 4A E2 65
 29 4A 8C A5 28 52 AD 94 A5 2A 32 94 A5 4A E2 65 29 4A 69 4A 52 4D B2 94 A5 46 52 94 A9 1C
 4C A5 25 19 4A 52 A5 5B 29 4A 54 65 29 4A 93 6C A5 29 51 94 A5 2A 55 B2 94 A5 34 A4 92 6D
 74 A5 34 A5 29 52 B8 99 4A 69 4D 29 4A 54 AE 25 E5 24 D2 94 A4 AB 5D 29 4C 9A 52 55 B2 94
 A5 46 52 94 A9 5C 4C 92 94 D2 94 A5 4A B5 E5 29 51 94 F2 12 A5 71 32 94 94 65 29 4A 95 6C
 A5 29 51 94 A5 04 DB 24 A5 46 52 94 95 6C 92 95 19 4A 52 55 B2 94 A5 46 53 A4 AB 5E 52 95
 19 4A 52 55 B2 94 A5 46 52 94 95 6C A5 32 69 49 15 5B 29 4A 54 52 49 56 CA 52 95 19 4A 52 55
 B2 94 A5 34 A6 94 15 6C A5 29 4C 94 A4 AB 5D 29 51 4A 52 55 B2 94 A5 46 52 94 95 6C 92 51
 94 F2 92 AD 79 4A 53 4A 52 92 AD 74 A5 46 53 CA 4A B5 D2 95 19 4F 29 26 D7 4A 53 4A 52 92
 AD 74 A5 34 A5 29 2A D7 94 A5 34 A6 94 A9 16 BC A5 29 A5 31 77 3D D1 A4 71 32 94 A5 09
 D6 3C 0D 1B A9 E0 20 EC BD C6 4B A3 C9 38 47 B9 A5 92 CB 5B 02 5B 72 94 AA 82 9B 41 FA
 7F 34 4A C3 87 2E 27 CD 23 11 DE D0 32 93 BC 48 4F 29 4A 80 91 DD

.....

BC 35 D1 31 04 48 58 49 22 87 77 59 94 BB 4A C2 49 25 3B DA FC 14 75 DD 1F 82 C3 49 25 07
 A0 89 80 99 C2 1C 40 FC 16 02 49 21 DE 88 0A 23 8D 38 58 69 24 90 EF 0F A3 F1 29 9A 60 98
 58 49 24 AE EE F3 FC 52 60 F6 C4 6A B0 52 49 5D 1D E2 3D 93 09 37 51 E6 B0 52 49 5D 3E AE
 FB 75 10 74 95 09 8F B9 61 A4 92 87 57 7D 9D 87 DE A2 DD 0C 7C 96 12 49 2B BB BD C3 88
 F0 EE 9F 92 0A C0 49 24 3B A4 43 B4 EE 54 CE 8B 9F 49 25 1E 8E EB 8C 47 9A 93 4A C0 49 24
 F4 7F **FF D9**

二、 JPEG 格式各字段详解

-----SOI标记-----

FF D8

标记码由两个字节构成，其前一个字节是固定值 **0xFF**，后一个字节则根据不同意义有不同数值。**SOI**的标记代码为 **0xFFD8**，即在JPEG文件中的如果出现数据 **0xFFD8**，则表示此处为一个**SOI**标记。

-----End SOI标记-----

----- APP0 标记-----

FF E0

APP0, Application, 应用程序保留标记

标记代码 2 字节 固定值 0xFFE0

包含 9 个具体字段：

00 10

① 数据长度 2 字节 ①~⑨9 个字段的总长度

5 字节 + 2 字节 + 1 字节 + 2 字节 + 2 字节 + 1 字节 + 1 字节 = 16 字节

4A 46 49 46 00

② 标识符 5 字节 固定值 **0x4A46494600**，即字符串“JFIF0”

01 02

③ 版本号 2 字节 一般是 **0x0102**，表示JFIF的版本号 **1.2**，可能会有其他数值代表其他版本

01

④ X和Y的密度单位 1 字节 只有三个值可选 **0**：无单位；**1**：点数/英寸；**2**：点数/厘米

01 2C

⑤ X方向像素密度 2 字节 取值范围未知

01 2C

⑤ Y方向像素密度 2 字节 取值范围未知

00

⑦ 缩略图水平像素数目 1 字节 取值范围未知

00

⑧ 缩略图垂直像素数目 1 字节 取值范围未知

本标记段可以包含图像的一个微缩版本，存为 **24** 位的**RGB**像素。如果没有微缩图像（这种情况更常见），则字段⑦“缩略图水平像素数目”和字段⑧“缩略图垂直像素数目”的值均为 **0**。

-----End APP0 标记-----

----- APPn标记-----

FF ED //APPD

APPn, Application，应用程序保留标记n，其中n=1~15(任选)

标记代码 2 字节 固定值 **0xFFE1~0xFFFF**

例如，**Adobe Photoshop**生成的JPEG图像中就用了**APP1** 和**APP13** 两个标记段分别存储了一幅图像的副本。

包含 **2** 个具体字段：

09 4C (2380)

① 数据长度 2 字节 ①~②2 个字段的总长度 即不包括标记代码，但包括本字段

```
50 68 6F 74 6F 73 68 6F 70 20 33 2E 30 00 38 42 49 4D 03 ED 0A 52 65 73 6F 6C 75 74 69 6F 6E
00 00 00 00 10 01 2C 00 00 00 01 00 01 01 2C 00 00 00 01 00 01 38 42 49 4D 04 0D 18 46 58 20
47 6C 6F 62 61 6C 20 4C 69 67 68 74 69 6E 67 20 41 6E 67 6C 65 00 00 00 00 04 00 00 00 78 38
42 49 4D 04 19 12 46 58 20 47 6C 6F 62 61 6C 20 41 6C 74 69 74 75 64 65 00 00 00 00 04 00 00
00 1E 38 42 49 4D 03 F3 0B 50 72 69 6E 74 20 46 6C 61 67 73 00 00 00 09 00 00 00 00 00 00 00
00 01 00 38 42 49 4D 04 0A 0E 43 6F 70 79 72 69 67 68 74 20 46 6C 61 67 00 00 00 00 01 00 00 38
```


32 E6 F2 43 BE 23 FB 91 F6 0F 82 86 57 53 D5 4B D5 59 CD CE 3F 9C D9 F3 06 3F 2A 28 CB
A4 89 DE 1B E4 ED 0A 69 C4 47 45 C3 25 F5 6E 7A A9 7A 81 50 76 7D 0D E1 C5 DF 01 FC 5D
0A 1F B4 99 AE D6 38 9F 88 84 46 19 1E 85 5E E8 EE E9 7A A1 3E F5 8C EE B3 5B 0C 3C D6
DF 22 FD 52 3F 58 7A 58 D0 D8 E9 FE A9 23 FC E4 7E EF 3F DD 27 CB 54 8C A3 BB B1 BD 2F
51 63 8E BB 86 EF E6 CE E3 D8 17 00 A2 EE AB 79 87 37 63 5B F7 FF 00 D2 71 44 72 D3 ED
5E 6A 39 83 FF D1 BB EA 29 7A A7 89 5C F3 FA BE 49 9D BB 1A 3E 04 FE 52 81 66 76 4D 9F
4E D7 11 E0 0E D1 FF 00 45 75 1E DB CC FB 65 E9 2C CC A2 AF E7 6D 63 0F 83 88 9F F3 55
6B 3A E6 23 34 66 EB 4F F2 44 0F F3 9F B5 73 DB BB F8 F2 52 DC 8F B5 15 DE DB D0 33 AF
63 B8 C3 D8 FA FC CC 38 7F D1 F7 23 B3 A9 E2 58 61 B7 36 4F 00 CB 7F EA D7 31 B9 3E F4
BD B8 AB 81 EB 3D 56 EA 49 1A 73 AF 08 16 F5 1C 6A F4 9D EE F0 6E BF F4 BE 8A E7 01 F0
3F 14 FB DD DC FD E9 0C 61 6F 03 AB 6F 56 BD DF CD B5 B5 8F 3F 71 FF 00 C8 AA B6 E5
DF 6E 96 58 E7 0F 09 D3 EE 0A A1 79 4C 5E 4A 70 88 1B 05 C2 29 BD 40 3C 12 F5 07 C1 57 94
F2 9C 9E 14 E5 E3 C4 14 B7 B7 E4 AB EE 4B 70 49 5C 2F FF D2 E7 09 4D 2B 3D 99 16 30 90
1C 63 CD 11 B9 6E 3C C2 EA 46 40 E2 9C 32 0D CD C9 C3 95 5F B5 3E 34 01 41 D7 58 EE 4F
3D 91 E3 08 F6 8F 80 6E 1B 58 39 70 9F C5 0D D9 5F BA 3E F5 53 71 F1 4A 50 E2 2B 86 21 D7
54 E6 FB 0F E7 1F 96 89 DB 93 60 E4 EE 1E 07 55 5F 71 48 94 2C A7 DB 1D 83 73 ED 60 8E 20
F9 F0 97 AE F3 A8 21 53 05 48 3B B8 47 88 A3 DA 88 D8 36 C5 EE EE 01 F8 29 8B 5A 7B C2 A6
1D 3C A9 4C 41 F1 44 49 69 C6 1B 73 F7 27 92 AA 6E 4E 1F E6 97 12 DE 07 FF D3 E1 24 FD C9
C1 8F 82 8C F8 6A 96 8B A2 B7 3E 99 EE 2A 42 C3 DF 54 2F F5 94 E2 7B A3 68 20 25 DE 0F 29
02 86 0F 74 FF 00 C5 1B 45 33 94 F3 A2 80 27 EF 4F BB EE 46 D1 4C BB A7 95 19 11 A2 52 8D
A2 99 C9 52 0F F3 43 F8 26 94 AD 06 29 A4 19 8D 0A 53 08 41 CA 41 DD B9 08 DA 29 FF D4
E0 B5 4B B2 C7 49 74 2D 07 67 B7 92 7D 56 2A 48 A1 DB F1 FC 12 FC BA 42 C4 49 14 3B 82
63 4E 13 19 EF E0 B1 12 49 5D 5D C1 DA 79 EE 9C 4C 79 2C 24 91 08 2E F7 7D 3C 53 EA B0
12 4E 43 BE 9C 4A E7 D2 49 0F **FF D9** 38 42 49 4D 04 21 1A 56 65 72 73 69 6F 6E 20 63 6F 6D
70 61 74 69 62 69 6C 69 74 79 20 69 6E 66 6F 00 00 00 00 55 00 00 00 01 01 00 00 00 0F 00 41 00
64 00 6F 00 62 00 65 00 20 00 50 00 68 00 6F 00 74 00 6F 00 73 00 68 00 6F 00 70 00 00 00 13 00
41 00 64 00 6F 00 62 00 65 00 20 00 50 00 68 00 6F 00 74 00 6F 00 73 00 68 00 6F 00 70 00 20 00
36 00 2E 00 30 00 00 00 01 00 38 42 49 4D 04 06 0C 4A 50 45 47 20 51 75 61 6C 69 74 79 00 00
00 00 07 00 01 00 00 00 01 01 00

FF EE //APPE

00 0E

41 64 6F 62 65 00 64 80 00 00 00 01

----- End APPn标记-----

-----DQT, Define Quantization Table, 定义量化表-----

FF DB

标记代码 2 字节 固定值 0xffdb

包含 2 个具体字段:

00 84

① 数据长度 2 字节 字段①和多个字段②的总长度,即不包括标记代码,但包括本
字段,84=16*8+4=132

00 0C 08 08 08 09 08 0C 09 09 0C 11 0B 0A 0B 11 15 0F 0C 0C 0F 15 18 13 13 15 13 13 18 11 0C
0C 0C 0C 0C 0C 11 0C
0C 0C 0C 0C 0C 0C 0C

01 0D 0B 0B 0D 0E 0D 10 0E 0E 10 14 0E 0E 0E 14 14 0E 0E 0E 14 11 0C 0C 0C 0C 0C 11
11 0C 0C 0C 0C 0C 0C 11 0C 0C 0C 0C 0C 0C 0C 0C 0C 0C 0C 0C 0C 0C 0C 0C 0C 0C 0C
0C 0C 0C 0C 0C 0C 0C

② 量化表 数据长度-2 字节, $132-2=130$

a) 精度及量化表 ID 1 字节

高 4 位: 精度, 只有两个可选值 0: 8 位; 1: 16 位;

低 4 位: 量化表 ID, 取值范围为 0~3

b) 表项 $(64 \times (\text{精度} + 1))$ 字节 例如 8 位精度的量化表, 其表项长度为 $64 \times (0 + 1) = 64$ 字节

本标记段中, 字段②可以重复出现, 表示多个量化表, 但最多只能出现 4 次。

-----End DQT, Define Quantization Table, 定义量化表-----

-----SOF0, Start of Frame, 帧图像开始-----

FF C0

标记代码 2 字节 固定值 0xFFC0, 包含 9 个具体字段:

00 11

① 数据长度 2 字节 ①~⑥六个字段的总长度, 即不包括标记代码, 但包括本字段, $11 = 16 \times 1 + 1 = 17$

08

② 精度 1 字节 每个数据样本的位数, 通常是 8 位, 一般软件都不支持 12 位和 16 位

02 58 (600)

③ 图像高度 2 字节 图像高度 (单位: 像素), 如果不支持 DNL 就必须 > 0

03 20 (800)

④ 图像宽度 2 字节 图像宽度 (单位: 像素), 如果不支持 DNL 就必须 > 0

03

⑤ 颜色分量数 1 字节 只有 3 个数值可选

1: 灰度图; 3: YCrCb 或 YIQ; 4: CMYK, 而 JFIF 中使用 YCrCb, 故这里颜色分量数恒为 3

01 22 00 02 11 01 03 11 01

⑥ 颜色分量信息 颜色分量数 $\times 3$ 字节 (通常为 9 字节)

a) 颜色分量 ID 1 字节

b) 水平/垂直采样因子 1 字节

高 4 位：水平采样因子 低 4 位：垂直采样因子（曾经看到某资料把这两者调转了）

c) 量化表号 1 字节 当前分量使用的量化表的 ID

本标记段中，字段⑥应该重复出现，有多少个颜色分量（字段⑤），就出现多少次（一般为 3 次）。

-----End SOF0, Start of Frame, 帧图像开始-----

-----DRI, Define Restart Interval, 定义差分编码累计复位的间隔-----

FF DD

标记代码 2 字节 固定值 0xFFDD

包含 2 个具体字段：

00 04

①数据长度 2 字节 固定值 0x0004，①~②两个字段的总长度，即不包括标记代码，但包括本字段

00 32

②MCU块的单元中的重新开始间隔 2 字节 设其值为n，则表示每n个MCU块就有一个RSTn标记。第一个标记是RST0，第二个是RST1等，RST7后再从RST0重复。

-----End DRI, Define Restart Interval, 定义差分编码累计复位的间隔-----

-----DHT, Define Huffman Table, 定义哈夫曼表-----

FF C4

标记代码 2 字节 固定值 0xFFC4

包含 2 个具体字段：

01 3F

①数据长度 2 字节 字段①和多个字段②的总长度，即不包括标记代码，但包括本字段，01 3F=1*16*16+3*16+15=319

② 哈夫曼表 数据长度-2 字节 (319-2=(1+16)*4+12+12+114+111)

00 DC 直流第 0 个哈夫曼表

a) 表 ID 和表类型 1 字节

高 4 位：类型，只有两个值可选 0：DC 直流；1：AC 交流

低 4 位：哈夫曼表 ID，注意，DC 表和 AC 表分开编码

00 01 05 01 01 01 01 01 01 00 00 00 00 00 00 00 (12)

b) 不同位数的码字数量 16 字节

03 00 01 02 04 05 06 07 08 09 0A 0B

c) 编码内容 16 个不同位数的码字数量之和（字节）

本标记段中，字段②可以重复出现（一般 4 次），也可以只出现 1 次。例如，Adobe Photoshop 生成的 JPEG 图片文件中只有 1 个 DHT 标记段，里边包含了 4 个哈夫曼表；而 Macromedia Fireworks 生成的 JPEG 图片文件则有 4 个 DHT 标记段，每个 DHT 标记段只有一个哈夫曼表。

01 DC 直流第 1 个哈夫曼表

00 01 05 01 01 01 01 01 01 00 00 00 00 00 00 00 (12)

01 00 02 03 04 05 06 07 08 09 0A 0B

10 AC 交流第 0 个哈夫曼表

00 01 04 01 03 02 04 02 05 07 06 08 05 03 0C 33 (114)

01 00 02 11 03 04 21 12 31 05 41 51 61 13 22 71 81 32 06 14 91 A1 B1 42 23 24 15 52 C1 62 33 34
72 82 D1 43 07 25 92 53 F0 E1 F1 63 73 35 16 A2 B2 83 26 44 93 54 64 45 C2 A3 74 36 17 D2 55
E2 65 F2 B3 84 C3 D3 75 E3 F3 46 27 94 A4 85 B4 95 C4 D4 E4 F4 A5 B5 C5 D5 E5 F5 56 66
76 86 96 A6 B6 C6 D6 E6 F6 37 47 57 67 77 87 97 A7 B7 C7 D7 E7 F7

11 AC 交流第 1 个哈夫曼表

00 02 02 01 02 04 04 03 04 05 06 07 07 06 05 35 (111)

01 00 02 11 03 21 31 12 04 41 51 61 71 22 13 05 32 81 91 14 A1 B1 42 23 C1 52 D1 F0 33 24 62
E1 72 82 92 43 53 15 63 73 34 F1 25 06 16 A2 B2 83 07 26 35 C2 D2 44 93 54 A3 17 64 45 55 36
74 65 E2 F2 B3 84 C3 D3 75 E3 F3 46 94 A4 85 B4 95 C4 D4 E4 F4 A5 B5 C5 D5 E5 F5 56 66
76 86 96 A6 B6 C6 D6 E6 F6 27 37 47 57 67 77 87 97 A7 B7 C7

-----End DHT, Define Huffman Table, 定义哈夫曼表-----

-----SOS, Start of Scan, 扫描开始 12 字节-----

FF DA

标记代码 2 字节 固定值 0xFFDA

包含 2 个具体字段:

00 0C

①数据长度 2 字节 ①~④两个字段的总长度, 即不包括标记代码, 但包括本字段

03

②颜色分量数 1 字节 应该和SOF中的字段⑤的值相同, 即:

1: 灰度图是; 3: YCrCb或YIQ; 4: CMYK。

01

③颜色分量信息

a) 颜色分量ID 1 字节

b) 直流/交流系数表号 1 字节 高 4 位: 直流分量使用的哈夫曼树编号
低 4 位: 交流分量使用的哈夫曼树编号

00 02 11 03 11 00 3F 00

④ 压缩图像数据

a)谱选择开始	1 字节	固定值 0x00
b)谱选择结束	1 字节	固定值 0x3F
c)谱选择	1 字节	在基本JPEG中总为 00

本标记段中，字段③应该重复出现，有多少个颜色分量（字段②），就出现多少次（一般为 3 次）。本段结束后，紧接着就是真正的图像信息了。图像信息直至遇到一个标记代码就自动结束，一般就是以EOI标记表示结束。

-----End SOS, Start of Scan, 扫描开始 12 字节-----

-----图像信息-----

F4 D9 4A 54 77 04 DB BE F5 2D 35 38 99 CA 52 A1 21 3C A5 4A E2 65 29 4A 8C A5 29 52 B8 99
4A 52 A3 29 4A 14 AE 26 5A A5 2A 32 94 A3 4A E2 65 29 4A 8C A5 28 52 AD 94 A5 2A 32 94 A5
4A E2 65 29 4A 69 4A 52 4D B2 94 A5 46 52 94 A9 1C 4C A5 25 19 4A 52 A5 5B 29 4A 54 65 29
4A 93 6C A5 29 51 94 A5 2A 55 B2 94 A5 34 A4 92 6D 74 A5 34 A5 29 52 B8 99 4A 69 4D 29 4A
54 AE 25 E5 24 D2 94 A4 AB 5D 29 4C 9A 52 55 B2 94 A5 46 52 94 A9 5C 4C 92 94 D2 94 A5 4A
B5 E5 29 51 94 F2 12 A5 71 32 94 94 65 29 4A 95 6C A5 29 51 94 A5 04 DB 24 A5 46 52 94 95 6C
92 95 19 4A 52 55 B2 94 A5 46 53 A4 AB 5E 52 95 19 4A 52 55 B2 94 A5 46 52 94 95 6C A5 32 69
49 15 5B 29 4A 54 52 49 56 CA 52 95 19 4A 52 55 B2 94 A5 34 A6 94 15 6C A5 29 4C 94 A4 AB
5D 29 51 4A 52 55 B2 94 A5 46 52 94 95 6C 92 51 94 F2 92 AD 79 4A 53 4A 52 92 AD 74 A5 46
53 CA 4A B5 D2 95 19 4F 29 26 D7 4A 53 4A 52 92 AD 74 A5 34 A5 29 2A D7 94 A5 34 A6 94 A9
16 BC A5 29 A5 31 77 3D D1 A4 71 32 94 A5 09 D6 3C 0D 1B A9 E0 20 EC BD C6 4B A3 C9 38
47 B9 A5 92 CB 5B 02 5B 72 94 AA 82 9B 41 FA 7F 34 4A C3 87 2E 27 CD 23 11 DE D0 32 93
BC 48 4F 29 4A 80 91 DD

BC 35 D1 31 04 48 58 49 22 87 77 59 94 BB 4A C2 49 25 3B DA FC 14 75 DD 1F 82 C3 49 25 07
A0 89 80 99 C2 1C 40 FC 16 02 49 21 DE 88 0A 23 8D 38 58 69 24 90 EF 0F A3 F1 29 9A 60 98
58 49 24 AE EE F3 FC 52 60 F6 C4 6A B0 52 49 5D 1D E2 3D 93 09 37 51 E6 B0 52 49 5D 3E AE
FB 75 10 74 95 09 8F B9 61 A4 92 87 57 7D 9D 87 DE A2 DD 0C 7C 96 12 49 2B BB BD C3 88
F0 EE 9F 92 0A C0 49 24 3B A4 43 B4 EE 54 CE 8B 9F 49 25 1E 8E EB 8C 47 9A 93 4A C0 49 24
F4 7F

-----End 图像信息-----

-----EOI, End of Image-----

FF D9

标记代码 2 字节 固定值 0xFFD9

这里补充说明一下，由于在 JPEG 文件中 0xFF 具有标志性的意思，所以在压缩数据流(真正的图像信息)中出现 0xFF，就需要作特别处理。具体方法是，在数据 0xFF 后添加一个没有意义的 0x00。换句话说，如果在图像数据流中遇到 0xFF，应该检测其紧接着的字符，如果是

1) 0x00，则表示 0xFF 是图像流的组成部分，需要进行译码；

2) 0xD9，则与 0xFF 组成标记 EOI，则图像流结束，同时图像文件结束；

3) 0xD0~0xD7, 则组成 RSTn 标记, 则要忽视整个 RSTn 标记, 即不对当前 0xFF 和紧接的 0xDn 两个字节进行译码, 并按 RST 标记的规则调整译码变量;

3) 0xFF, 则忽视当前 0xFF, 对后一个 0xFF 再作判断;

4) 其他数值, 则忽视当前 0xFF, 并保留紧接的此数值用于译码。

三、 参考文献

[1] JPEG 文件编/解码详解, cat_ng 猫猫,

<http://blog.csdn.net/ShiZhixin/archive/2009/07/24/4378260.aspx>转载

[2] Visual C++实现 MPEG/JPEG 编解码技术, 张益贞、刘滔, 人民邮电出版社, 2002 年 11 月

注: 本文章为个人学习笔记, 参考了很多文献, 加上了个人的理解形成此文。在此, 对参考文献中的原作者以及由于种种原因未列出的原作者表示感谢! 如有侵犯版权, 请 Email 告知。再次声明, 本资料不是论文, 是个人学习笔记, 仅供学习交流之用。

四、 附件

实例用图:

