

# MP3 文件格式解析

汪 勇 熊前兴  
(武汉理工大学计算机学院 武汉 430063)

**摘 要** 随着 Internet 的迅猛发展,MP3 高密度和需要较小的传输带宽的特点使其占据乐坛霸主的地位,尽管网络不是音乐发布的唯一渠道,但 MP3 已成为一种公认的音乐数据格式。本文介绍了关于 MP3 文件的编码与解码原理,进而对 MP3 文件格式作了详细的分析。

**关键词** MP3 文件格式 帧 声音文件 音频压缩

## FILE FORMAT ANALYSIS OF MP3

Wang Yong Xiong Qianxing  
(School of Computer, Wuhan University of Technology, Wuhan 430063)

**Abstract** With the rapid development of Internet, the features of high density and needing smaller transmitting band width make MP3 occupy the header position of music area. Though network is not only channel to publish music, MP3 has become a common audio data format. At first, this paper introduces the encoding and decoding principle of MP3, then it analyses deliberately about MP3 file format.

**Keywords** MP3 File format Frame Audio file Audio compression

## 1 引 言

文件压缩技术的日新月异使得 MP3 成为时下最烫手的音乐格式,优质的音乐随着 0 与 1 的排列迅速散布到世界各地,撼动人心。何谓 MP3? MP3 的全称是 MPEG Audio Layer3,它是一种高效的计算机音频编码方案,它以较大的压缩比将音频文件转换成较小的扩展名为 .MP3 的文件,基本保持原文件的音质。MP3 是 ISO/MPEG 标准的一部分,ISO/MPEG 标准描述了使用高性能感知编码方案的音频压缩,此标准一直在不断更新以满足“质高量小”的追求,现已形成 MPEG Layer1、Layer2、Layer3 三个音频编解码方案。MPEG Layer3 压缩率可达 1:10 至 1:12,1M 的 MP3 文件可播放 1 分钟,而 1 分钟 CD 音质的 WAV 文件(44100Hz, 16bit, 双声道, 60 秒)要占用 10M 空间,这样算来,一张 650M 的 MP3 光盘播放时间应在 10 小时以上,而同样容量的一张 CD 盘播放时间在 70 分钟左右。MP3 的优势是 CD 难以比拟的。

## 2 MP3 原理浅析

### 2.1 MPEG Audio 标准

MPEG(Moving Picture Experts Group)是 ISO 下的一个动态图像专家组,它制定的 MPEG 标准广泛应用于各种多媒体中。MPEG 标准包括视频和音频标准,其中音频标准已制定出 MPEG-1、MPEG-2、MPEG-2 AAC 和 MPEG-4。

MPEG-1 和 MPEG-2 标准使用同一个音频编解码族—Layer1、2、3。MPEG-2 一个新特点是采用低采样率扩展降低数据流量,另一特点是多通道扩展,将主声道增加为 5 个。MPEG

-2 AAC(MPEG-2 Advanced Audio Coding)标准是 Fraunhofer IIS 同 AT&T 公司于 1997 年推出的,旨在显著减少数据流量,MPEG-2 AAC 采用的修正的离散余弦变换(MDCT, Modified Discrete Cosine Transform)算法,采样率可在 8KHz 到 96KHz 之间,声道数可在 1—48 之间。

MPEG Audio Layer1、2、3 三个层使用相同的滤波器组、位流结构和头信息,采样频率为 32KHz、44.1KHz 或 48KHz。Layer 1 是为数字压缩磁带 DCC(Digital Compact Cassette)设计的,数据流量为 384kbps,Layer 2 在复杂性和性能间作了权衡,数据流量下降到 256kbps—192kbps。Layer 3 一开始就为低数据流量而设计,数据流量在 128kbps—112kbps。Layer 3 增加了 MDCT 变换,使其频率分辨能力是 Layer 2 的 18 倍,Layer 3 还使用了与 MPEG Video 类似的平均信息量编码(Entropy Coding),减少了冗余信息。MP3 绝大部分使用的是 MPEG-1 标准。

### 2.2 音频压缩

MP3 格式始于 80 年代中期,德国 Erlangen 的 Fraunhofer 研究所致力于高质量、低数据率的声音编码。

MP3 音频压缩包含编码和解码两个部分。编码是将 WAV 文件中的数据转换成高压压缩率的位流形式,解码是接受位流并将其重建到 WAV 文件中。

MP3 采用了感知音频编码(Perceptual Audio Coding)这一失真算法。人耳感受声音的频率范围是 20Hz~20kHz,MP3 截掉了大量的冗余信号和无关的信号,编码器通过混合滤波器组将原始声音变换到频率域,利用心理学模型,估算刚好能被察觉到的噪声水平,再经过量化,转换成 Huffman 编码,形成 MP3 位流。

收稿日期:2003-09-09 汪勇,副教授,主研领域:数据库与信息系统,电子商务。

解码器要简单得多, 它的任务是从编码后的谱线成分中, 经过反量化和逆变换, 提取出声音信号。

在压缩音频数据时, 先将原始声音数据分成固定的分块, 然后作顺向 MDCT 变换, MDCT 本身并不进行数据压缩, 只是将一组时域数据转换成频域数据, 以得知时域变化情况, 顺向 MDCT 将每块的值转换为 512 个 MDCT 系数。量化使数据得到压缩, 在对量化后的变换样值进行比特分配时要考虑使整个量化块最小, 这就成为有损压缩了。解压时, 经反向 MDCT 将 512 个系数还原成原始声音数据, 前后的原始声音数据是不一致的, 因为在压缩过程中, 去掉了冗余和不相关数据。

3 MP3 文件格式分析

经过压缩后的 MP3 文件数据由多个帧组成, 帧是 MP3 文件最小组成单位。每个帧又由帧头、附加信息和声音数据组成。每个帧播放时间是 0.026 秒, 其长度随位率的不同而不同, 有些 MP3 文件末尾有些额外字节存放非声音数据的说明信息。MP3 文件结构如图 1 所示。

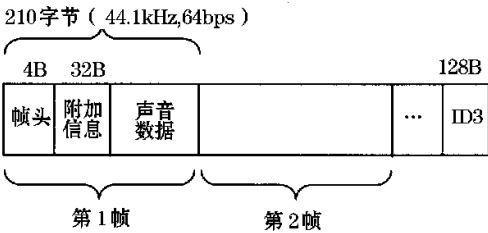


图 1 MP3 文件结构

3.1 帧头格式

帧头长 4 字节, 对于固定位率的 MP3 文件, 所有帧的帧头格式一样其数据结构如下:

```
typedef FrameHeader{
    unsigned int syn; 11; //同步信息
    unsigned int version; 2; //版本
    unsigned int layer; 2; //层
    unsigned int protection; 1; //CRC 校验
    unsigned int bitrate; 4; //位率
    unsigned int frequency; 2; //频率
    unsigned int padding; 1; //帧长调节
    unsigned int private; 1; //保留字
    unsigned int mode; 2; //声道模式
    unsigned int mode extension; 2; //扩充模式
    unsigned int copyright; 1; //版权
    unsigned int original; 1; //原版标志
    unsigned int emphasis; 2; //强调模式
} HEADER, *LPHEADER;
```

帧头 4 字节使用说明见表 1。

表 1 MP3 帧头字节使用说明

名称	长度(位)	说明
同步信息	11	所有位均为 1, 第 1 字节恒为 FF.
版本	2	00-MPEG 2 5 01-未定义 10-MPEG 2 11-MPEG 1
层	2	00-未定义 01-Layer 3 10-Layer 2 11-Layer 1
CRC 校验	1	0-校验 1-不校验

名称	长度(位)	说明
位率	4	取率率, 单位是 kbps, 例如采用 MPEG-1 Layer 3, 64kbps 是, 值为 0101.
频率	2	采样频率, 对于 MPEG-1: 00-44.1kHz 01-48kHz 10-32kHz 11-未定义
帧长调节	1	用来调整文件头长度, 0-无需调整, 1-调整, 具体调整计算方法见下文。
保留字	1	没有使用。
声道模式	2	表示声道: 00-立体声 01-Joint Stereo 10-双声道 11-单声道
扩充模式	2	当声道模式为 01 是才使用。
版权	1	文件合法性, 0-不合法 1-合法
原版标志	1	是否原版, 0-非原版, 1-原版
强调方式	2	用于声音经降噪压缩后再补偿的分类, 很少用到, 今后也可能不会用。 00-未定义 01-50/15ms 10-保留 11-CCITT J 17

MP3 帧长取决于位率和频率, 计算公式为:

帧长=⌊ 144× bitrate/frequency+ padding⌋

例如: 位率为 64kbps, 频率为 44.1kHz, padding 为 1 时, 帧长为 210 字节。帧头后面是可变长度的附加信息, 对于标准的 MP3 文件来说, 其长度是 32 字节, 紧接其后的是压缩的声音数据, 当解码器读到此处时就进行解码了。

对于固定位率(CBR, Constant Bitrate)的 MP3 文件, 并不是所有的帧都是等长的, 有的帧可能多一个或几个字节。还有一种可变位率(VBR, Variable Bitrate)的 MP3 文件, 是为了使 MP3 文件长度最小同时又保证声音质量, 与 CBR 文件相比, 除了第一帧不同外, 其余的都一样。VBR 的第一帧不包含声音数据, 其长度是 156 个字节, 用来存放标准的声音帧头(4 字节)、VBR 文件标识、帧数、文件字节数等信息, 具体结构说明见表 2。

表 2 VBR 文件第一帧结构

字 节	说 明
1—4	与 CBR 相同的标准声音帧头
5—40	存放 VBR 文件标识“Xing”(58 69 6E 67), 此标识具体位置视采用的 MPEG 标准和声道模式而定。标识的前后字节没有使用。 36—39 MPEG-1 和非单声道(常见) 21—24 MPEG-1 和单声道 21—24 MPEG-2 和非单声道 13—16 MPEG-2 和单声道
41—44	标志, 说明是否存储了帧数、文件长度、目录表和 VBR 规模信息, 如果存储了, 则 01 02 04 08.
45—48	帧数(包括第一帧)
49—52	文件长度
53—152	目录表, 用来按时间进行字节定位。
153—156	VBR 规模, 用于高级应用。

3.2 ID3 标准

MP3 帧头中除了存储一些像 private、copyright、original 的简单音乐说明信息以外, 没有考虑存放歌名、作者、专辑名、年份等复杂信息, 而这些信息在 MP3 应用中非常必要。1996 年, FricKemp 在“Studio3”项目中提出了在 MP3 文件尾增加一块用于存放歌曲的说明信息, 形成了 ID3 标准, 至今已制定出 ID3 V1.0 V1.1, V2.0 V2.3 和 V2.4 标准。版本越高, 记录的相关信息就越丰富详尽。

ID3 V 1.0 标准并不周全, 存放的信息少, 无法存放歌词, 无法录入专辑封面、图片等。V2.0 是一个相当完备的标准, 但给编写软件带来困难, 虽然赞成此格式的人很多, 在软件中真正实现的却极少。绝大多数 MP3 仍使用 ID3 V1.0 标准。此标准是将 MP3 文件尾的最后 128 个字节用来存放 ID3 信息, 这 128 个字节使用说明见表 3。

表 3 ID3 V1.0 文件尾说明

字 节	长度(字节)	说 明
1~3	3	存放“TAG”字符, 表示 ID3 V1.0 标准, 紧接其后的是歌曲信息
4~33	30	歌名
34~63	30	作者
64~93	30	专辑名
94~97	4	年份
98~127	30	附注
128	1	MP3 音乐类别, 共 147 种

3.3 文件实例

在VC++ 中打开一个名为 test. mp3 文件, 其内容如下:

```
000000 FF FB 52 8C 00 00 01 49 09 C5 05 24 60 00 2A C1
000010 19 40 A6 00 00 05 96 41 34 18 20 80 08 26 48 29
000020 83 04 00 01 61 41 40 50 10 04 00 C1 21 41 50 64
.....
0000D0 FE FF FB 52 8C 11 80 01 EE 90 65 6E 08 20 02 30
0000E0 32 0C CD C0 04 00 46 16 41 89 B8 01 00 08 36 48
0000F0 33 B7 00 00 01 02 FF FF FF F4 E1 2F FF FF FF FF
.....
0001A0 DF FF FF FB 52 8C 12 00 01 FE 90 58 6E 09 A0 02
0001B0 33 B0 CA 85 E1 50 01 45 F6 19 61 BC 26 80 28 7C
0001C0 05 AC B4 20 28 94 FF FF FF FF FF FF FF FF FF
.....
001390 7F FF FF FF FD 4E 00 54 41 47 54 45 53 54 00 00
0013A0 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
.....
0013F0 00 00 00 00 04 19 14 03 00 00 00 00 00 00 00
001400 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
001410 00 00 00 00 00 00 4E
```

该文件长度 1416H(5. 142K), 帧头为: FF FB 52 8C, 转换成二进制为: 11111111 11111011 01010010 10001100。对照表 1 可知, test. mp3 帧头信息见表 4。

表 4 test. mp3 文件帧头信息

名称	位值	说 明
同步信息	11111111 1111	第 1 字节恒为 FF, 11 位均为 1
版本	11	MPEG 1
层	01	Layer 3
CRC 校验	1	不校验
位率	0101	64kbps
频率	00	44 1kHz
帧长调节	1	调整, 帧长是 210 字节
保留字	0	没有使用
声道模式	10	双声道
扩充模式	00	未使用
版权	1	合法
原版标志	1	原版
强调方式	00	未定义

第 1397H 开始的三个字节是 54 41 47, 存放的是字符“TAG”, 表示此文件有 ID3 V1.0 信息。  
139AH 开始的 30 个字节存放歌名, 前 4 个非 00 字节是 54 45 53 54, 表示“TEST”。  
13F4H 开始的 4 个字节是 04 19 14 03, 存放年份“04/25/2003”。  
最后的 1 个字节是 4E, 表示音乐类别, 代号为: 78, 亦即

“Rock&Roll”。  
其它字节均为 00, 未存储信息。

4 结束语

声音作为一类重要的多媒体数据, 人们总是在不断寻求更高效率的压缩方法和新的声音文件格式。通过对 MP3 文件格式的分析, 不难发现它的欠缺。一方面, MP3 文件每个帧都有 4 字节的相同帧头, 对于一个含有大量帧的 MP3 文件来说, 需要一定的空间开销, 如果将第一帧作为文件头, 用来存放文件信息, 其余帧专门用来存放数据, 将会有效地解决这个问题。另一方面, ID3 存放音乐说明信息, 帧头中的 private、copyright 等信息也是简单的说明信息, 音乐的说明信息缺乏统一的管理。若将 ID3 作为文件尾, 专门存放音乐说明信息, 避免同类信息的零乱存放。

无论如何, MP3 的发展势不可挡, MP3 已成为公认的声音数据格式, MP3 正与 JPEG 图像、PDF 文档成为多媒体信息处理领域的热点。

参 考 文 献

[ 1 ] Cervera Teresa, The effect of MPEG audio compression on multidimensional set of voice parameters, Logopedics Phoniatrics Vocology, 2001. 8, Vol. 26, p 124.  
[ 2 ] Bosse Lincoln, An Experimental High Fidelity Perceptual Audio Coder Project in MUS420 Win 97, Computer Based Learning Unit, University of Leeds, 1996.  
[ 3 ] Mathew A. Watson, Peter Buettner, Design and Implementation of AAC Decoders, IEEE Transactions on Consumer Electronics, 2000. 6.  
[ 4 ] 林福宗, 多媒体技术基础[M], 北京: 清华大学出版社, 2001.  
[ 5 ] <http://www.iis.fraunhofer.de/am/techinf/layer3/index.html#top>.  
[ 6 ] <http://www.kom.auc.dk/dsp/doc/1014>.  
[ 7 ] <http://www.id3.org>.  
[ 8 ] [http://mpgedit.org/mpgedit/mpeg\\_format/MP3Format.html](http://mpgedit.org/mpgedit/mpeg_format/MP3Format.html).  
[ 9 ] <http://www.psionics.demon.co.uk/mp3/mp3inf>.

(上接第 43 页)  
数据库数据的完整性; 达不到优化和整理数据库目的, 只能实现一对一的数据库同步; 优点是: 实现简单, 出错机会少, 实现时间短(70GB 数据库同步时间约 4 小时), 不会影响业务。

6 结束语

数据同步是复制服务器的重要内容, 在实际工作中, 如果只涉及某几张表较短时间内的同步问题, 用 BCP 方法实现比较合适; 如复制宕机时间长, 涉及复制表较多, 建议用 Dump/ Load 实现整库数据同步。

参 考 文 献

[ 1 ] 宋兴彬、徐力、张广涛, “基于 BCP 方法实现 Sybase 复制数据同步的研究”, 《山东电子》, 2003 03(待发表).  
[ 2 ] Sybase Inc, Fast Track to Replication Server Administration 12.0 Student Guide volume1, 2 Version 1.1 Sybase Inc.  
[ 3 ] 王珊、周胜, Sybase Replication Server 使用高级指南, 水利水电出版社, 2000. 3.  
[ 4 ] Sybase Inc Sybase Replication Server 12.0 管理指南, 2000. 1.  
[ 5 ] 宋兴彬 “基于 Sybase 复制技术的分布式数据库系统的建立”, 《山东科学》2000 年第 1 期.