

WAV 文件格式分析与应用

一. RIFF 概念

在 Windows 环境下, 大部分的多媒体文件都依循着一种结构来存放信息, 这种结构称为"资源互换文件格式"(Resources Interchange File Format), 简称 RIFF。例如声音的 WAV 文件、视频的 AV1 文件等等均是由此结构衍生出来的。RIFF 可以看做是一种树状结构, 其基本构成单位为 chunk, 犹如树状结构中的节点, 每个 chunk 由"辨别码"、"数据大小"及"数据"所组成。

块的标志符 (4BYTES)
数据大小 (4BYTES)
数据

图一、块的结构示意图

辨别码由 4 个 ASCII 码所构成, 数据大小则标示出紧跟其后数据的长度(单位为 Byte), 而数据大小本身也用掉 4 个 Byte, 所以事实上一个 chunk 的长度为数据大小加 8。一般而言, chunk 本身并不允许内部再包含 chunk, 但有两种例外, 分别为以"RIFF"及"L1ST"为辨别码的 chunk。而针对此两种 chunk, RIFF 又从原先的"数据"中切出 4 个 Byte。此 4 个 Byte 称为"格式辨别码", 然而 RIFF 又规定文件中仅能有一个以"RIFF"为辨别码的 chunk。

RIFF/LIST 标志符	
数据 1 大小	
数据 1	格式/列表类型
	数据

图二、RIFF/LIST 块结构

只要依循此一结构的文件, 我们均称之为 RIFF 档。此种结构提供了一种系统化的分类。如果和 MS-DOS 文件系统作比较, "RIFF"chunk 就好比是一台硬盘的根目录, 其格式辨别码便是此硬盘的逻辑代码(C: 或 D:), 而"L1ST"chunk 即为其下的子目录, 其他的 chunk 则为一般的文件。至于在 RIFF 文件的处理方面, 微软提供了相关的函数。视窗下的各种多媒体文件格式就如同在磁盘机下规定仅能放怎样的目录, 而在该目录下仅能放何种数据。

二. WAV 文件格式

WAVE 文件是非常简单的一种 RIFF 文件, 它的格式类型为"WAVE"。RIFF 块包含两个子块, 这两个子块的 ID 分别是"fmt"和"data", 其中"fmt"子块由结构 PCMWAVEFORMAT 所组成, 其子块的大小就是 sizeof(PCMWAVEFORMAT), 数据组成就是 PCMWAVEFORMAT 结构中的数据。

标志符 (RIFF)
数据大小
格式类型 ("WAVE")
"fmt"
Sizeof(PCMWAVEFORMAT)
PCMWAVEFORMAT
"data"
声音数据大小
声音数据

图三、WAVE 文件结构

PCMWAVEFORMAT 结构定义如下:

```

Typedef struct
{
    WAVEFORMAT wf;          //波形格式;
    WORD wBitsPerSample;    //WAVE 文件的采样大小;
} PCMWAVEFORMAT;

//WAVEFORMAT 结构定义如下:
typedef struct
{
    WORD wFormatTag;        //编码格式, 包括 WAVE_FORMAT_PCM, WAVEFORMAT_ADPCM 等
    WORD nChannels;        //声道数, 单声道为 1, 双声道为 2;
    DWORD nSamplesPerSec;   //采样频率;
    DWORD nAvgBytesperSec;  //每秒的数据量;
    WORD nBlockAlign;       //块对齐;
} WAVEFORMAT;
    
```

“data”子块包含 WAVE 文件的数字化波形声音数据, 其存放格式依赖于“fmt”子块中 wFormatTag 成员指定的格式种类, 在多声道 WAVE 文件中, 样本是交替出现的。如 16bit 的单声道 WAVE 文件和双声道 WAVE 文件的数据采样格式分别如图四所示:

16 位单声道:

采样一		采样二	
低字节	高字节	低字节	高字节

16 位双声道:

采样一			
左声道		右声道	
低字节	高字节	低字节	高字节

图四、WAVE 文件数据采样格式

WAV 文件格式实例分析:

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
00000000H	52	49	46	46	0A	06	01	00	57	41	56	45	66	6D	74	20
00000010H	12	00	00	00	01	00	02	00	44	AC	00	00	10	B1	02	00
00000020H	04	00	10	00	64	61	74	61	E6	05	01	00	00	00	00	00

	偏移地址	字节数	数据类型	内容
文件头	00H	4	char	“RIFF”； RIFF 标志
	04H	4	long int	0x00 01 06 0A (注意数据存储顺序)； 文件长度
	08H	4	char	“WAVE”； WAVE 标志
	0CH	4	char	“fmt ”； fmt 标志, 最后一位为空
	10H	4	long int	0x12; sizeof(PCMWAVEFORMAT)
	14H	2	int	1 (WAVE_FORMAT_PCM)； 格式类别, 1 表示为 PCM 形式的声音数据
	16H	2	int	2; 通道数, 单声道为 1, 双声道为 2
	18H	2	int	44100; 采样频率 (每秒样本数)
	1CH	4	long int	0x10B10000; 每秒数据量; 其值为通道数×每秒数据位数×每样本的数据位数 / 8。播放软件利用此值可以估计缓冲区的大小。
	20H	2	int	数据块的调整数 (按字节算的), 其值为通道数×每样本的数据位值 / 8。播放软件需要一次处理多个该值大小的字节数据, 以便将其值用于缓冲区的调整。
	22H	2		每样本的数据位数, 表示每个声道中各个样本的数据位数。如果有多个声道, 对每个声道而言, 样本大小都一样。
	24H	4	char	“data”； 数据标记符
28H	4	long int	0x00 01 05 E6; 语音数据大小	