

数字音频技术

声音的基本概念

声音是通过一定介质传播的连续的波



- 重要指标：
 - 振幅：音量的大小
 - 周期：重复出现的时间间隔
 - 频率：指信号每秒钟变化的次数 ###声音按频率分类： 次声波： < 20HZ 人耳可听： 20Hz~20kHz 超声波： > 20kHz

声音的类型

- 波形声音：包含了所有声音形式。
- 语音：不仅是波形声音，而且还有丰富的语言内涵(抽象→提取特征→意义理解)。
- 音乐：与语音相比，形式更规范。音乐是符号化的声音。

声音信号

类型

- 模拟声音
- 数字声音

转化



数字化三要素

- 采样频率
 - 采样频率越高,声音质量越好,数据量也越大
- 量化位数
 - 量化位数越多,音质越好,数据量也越大
- 声道数
 - 立体声比单声道的表现力丰富

#DSD与PCM

DSD包括DSF、DFF格式，包括DSD64、DSD128、DSD256、DSD512等。 PCM： 脉冲编码调制 DSD： 直接比特流数字编码调制

音频的文件格式

- MP1、MP2、MP3、WMA、AAX、M4A、OGG、FLAC、APE、AIF、DSF、DFF...

MP3结构（需授权）：

ID3v2(包含作者、作曲、专辑等，长度不固定，扩展了ID3v1的信息量)
Frame(帧，个数由文件大小和帧长决定)
...（每个帧的长度由位率bitrate决定）
Frame(每帧分帧头和数据，帧头记录了mp3的位率，采样率和版本等信息，每帧相互独立)
ID3v1(包含作者、作曲、专辑等信息，128BYTE)

APE结构（免费开源）：

header(文件的属性、sound的参数（如声道数、采样率等）、内部结构（如帧数、Seek Table，甚至可能包括WAV的header）

tag = 'MAC'

compression_level(压缩等级): fast, normal, high, extra-high, insane

APE Tag(存放metadata数据，比如歌名、演唱者、专辑名等)，有两个版本，APEv1和APEv2,v1一般放在文件末尾，

APEv2具有和ID3v2一样的灵活扩展性，v2的格式不像ID3v2那么繁琐，格式简单，实现起来也很方便，

并且存放位置可选，v1和v2版本的区别是：

- (a) APEv1使用的是ASCII编码，APEv2使用的是UTF-8编码，使得可以实现unicode支持。
- (b) APEv2标准里增加了一个APE Tags Header，APEv1里面没有。

OGG结构（开源免费）：

OGG文件由一个个的page组成，解析时通过page的header标记'Oggs'将page解析出来。
与文件、stream有关的信息，也是保存在Pages中的，这些Pages一般放在文件的开始位置

FLAC结构（开源免费）：

(FLAC编码设置只对编码时间有较大影响而对解码影响不大；
因为等级越高，编码器就会花越多的时间去寻找最佳的压缩算法，而解码器则根据给定的压缩算法直接解压)

基本结构：

1. 4byte 字符“fLaC”：flac标志，用于识别flac数据流
2. streaminfo文件信息描述块 (metadata block)：包含必须的信息（采样率、声道数...）
3. 可选的其他描述信息块：（解码时可以不识别）
4. 一个以上的音频帧 (frame)

AAC结构：

AAC“Advanced Audio Coding”的缩写，中文称为“高级音频编码”，被手机界称为“21世纪数据压缩方式”

AAC同时可以支持多达48个音轨，15个低频音轨，更多种取样率和比特率与及有多种言语的兼容能力，更高的译码效率

音效：

蝰蛇音效：主要功能是用于对MP3等有损音乐进行细节和高音补偿，指由声音所制造的效果，是指为增进一步之真实感、气氛或戏剧讯息，而加于声带上的杂音或声音
EQ音效(数字音效)：包括不限于：classic,pop,jazz,rock,normal,auto...

高保真：

音乐上的“高保真”，指重放时清晰度的保证、频响保证、失真度与信噪比的保证。