
RT-THREAD PLAYER

音频格式

RT-THREAD 文档中心

上海睿赛德电子科技有限公司版权 ©2018



WWW.RT-THREAD.ORG

Monday 15th April, 2019

目录

目录	i
1 Player 组件音频格式	1
1.1 版本和修订	1
1.2 Player 组件支持格式说明	1
1.3 音频格式简介	1
1.3.1 WAV 格式	2
1.3.2 MP3 格式	2
1.3.3 M4A 格式	2
1.3.4 AMR 格式	3
1.3.5 AAC 格式	3

第 1 章

Player 组件音频格式

1.1 版本和修订

Date	Version	Author	Note
2019-01-29	v1.0	liu2guang	初始版本

1.2 Player 组件支持格式说明

RT-Thread 提供的 Player 组件支持多种音频格式的播放, 目前支持的有以下的音频格式:

音频格式	后缀名	Player 组件支持详情
WAV 格式	.wav	支持单声道 (Mono) 与双声道 (Stereo), 支持 11.025KHz ~ 44.1KHz 采样率, 8bit ~ 16bit 采样精度 ¹
MP3 格式	.mp3	支持 CBR/VBR/ABR, 支持 96Kbps@8KHz ~ 320Kbps@48KHz, 支持带 ID3v2 的 mp3 音频
M4A 格式	.m4a	支持 AAC 编码的 M4A
AMR 格式	.amr	支持 AMR NB 编码 (语音带宽范围: 300 ~ 3400Hz, 8KHz 采用率)

1.3 音频格式简介

¹WAV 的支持情况与音频芯片和控制器芯片以及音频驱动实现有关.

1.3.1 WAV 格式

WAV 为微软公司 (Microsoft) 开发的一种声音文件格式, 它符合 RIFF 文件规范, 用于保存 Windows 平台的音频信息资源, 被 Windows 平台及其应用程序所广泛支持, 该格式也支持 MSADPCM, CCITTALAW 等多种压缩运算法, 支持多种音频数字, 取样频率和声道, 标准格式化的 WAV 文件和 CD 格式一样, 也是 44.1K 的取样频率, 16 位量化数字, 因此在声音文件质量和 CD 相差无几。WAV 打开工具是 WINDOWS 的媒体播放器。

通常使用三个参数来表示声音: 量化位数、取样频率和采样点振幅。量化位数分为 8 位、16 位、24 位三种, 声道有单声道和立体声之分, 取样频率一般有 11025Hz(11kHz)、22050Hz(22kHz) 和 44100Hz(44kHz) 三种, 不过尽管音质出色, 但在压缩后的文件体积过大! 相对其他音频格式而言是一个缺点, 其文件大小的计算方式为: WAV 格式文件所占容量 (B) = (取样频率量化位数声道) * 时间 / 8 (1 字节 = 8bit) 每一分钟 WAV 格式的音频文件的大小为 10MB, 其大小不随音量大小及清晰度的变化而变化。

WAV 是最接近无损的音乐格式, 所以文件大小相对也比较大。

1.3.2 MP3 格式

MPEG-1 or MPEG-2 Audio Layer III 是一种音频压缩技术, 其全称是动态影像专家压缩标准音频层面 3 (Moving Picture Experts Group Audio Layer III), 简称为 MP3, 是目前最流行的音频编码格式。

MP3 文件是由帧 (frame) 构成的, 帧是 MP3 文件最小的组成单位。MPEG 音频文件是 MPEG1 标准中的声音部分, 也叫 MPEG 音频层, 它根据压缩质量和编码复杂程度划分为三层, 即 Layer-1、Layer2、Layer3, 且分别对应 MP1、MP2、MP3 这三种声音文件, 并根据不同的用途, 使用不同层次的编码。MPEG 音频编码的层次越高, 编码器越复杂, 压缩率也越高, MP1 和 MP2 的压缩率分别为 4:1 和 6:1-8:1, 而 MP3 的压缩率则高达 10:1-12:1, 也就是说, 一分钟 CD 音质的音乐, 未经压缩需要 10MB 的存储空间, 而经过 MP3 压缩编码后只有 1MB 左右。不过 MP3 对音频信号采用的是有损压缩方式, 为了降低声音失真度, MP3 采取了“感官编码技术”, 即编码时先对音频文件进行频谱分析, 然后用过滤器滤掉噪音电平, 接着通过量化的方式将剩下的每一位打散排列, 最后形成具有较高压缩比的 MP3 文件, 并使压缩后的文件在回放时能够达到比较接近原音源的声音效果。

MP3 的比特率种类分别是 VBR、ABR 和 CBR 三种:

1. CBR CBR (Constant Bitrate) 常数比特率指文件从头到尾都是一种位速率。相对于 VBR 和 ABR 来讲, 它压缩出来的文件体积很大, 而且音质相对于 VBR 和 ABR 不会有明显的提高。
2. VBR VBR (Variable Bitrate) 动态比特率也就是没有固定的比特率, 压缩软件在压缩时根据音频数据即时确定使用什么比特率, 这是以质量为前提兼顾文件大小的方式, 推荐编码模式。
3. ABR ABR (Average Bitrate) 平均比特率是 VBR 的一种插值参数。LAME 针对 CBR 不佳的文件体积比和 VBR 生成文件大小不定的特点独创了这种编码模式。ABR 在指定的文件大小内, 以每 50 帧 (30 帧约 1 秒) 为一段, 低频和不敏感频率使用相对低的流量, 高频和大动态表现时使用高流量, 可以做为 VBR 和 CBR 的一种折衷选择。

1.3.3 M4A 格式

M4A 是 MPEG-4 音频标准的文件的扩展名。在 MPEG4 标准中提到, 普通的 MPEG4 文件扩展名是“.mp4”。自从 Apple 开始在它的 iTunes 以及 iPod 中使用“.m4a”以区别 MPEG4 的视频和音频文件以来,

“m4a”这个扩展名变得流行了。目前，几乎所有支持 MPEG4 音频的软件都支持“m4a”。最常用的.m4a 文件是使用 AAC 格式的（文件），不过其他的格式，比如 Apple Lossless 甚至 mp3 也可以被放在.m4a 容器里。可以安全地把只包含音频的.mp4 文件的扩展名改成.m4a，以便让它能在你喜欢的播放器里播放，反之亦然。

1.3.4 AMR 格式

AMR（Adaptive Multi-Rate）自适应多速率音频压缩音频编码格式，是一个使语音编码最优化的专利，专用于有效地压缩语音频率。

AMR 音频主要用于移动设备的音频压缩，压缩比非常高，但是音质比较差，主要用于语音类的音频压缩，不适合对音质要求较高的音乐类音频的压缩。AMR 被标准语音编码 3GPP 在 1998 年 10 月选用，现在广泛在 GSM 和 UMTS 中使用。它使用 1-8 个不同的位速编码。之前的手机里有很多 amr 的音频文件，可分成：1. AMR-NB（AMR-NarrowBand）：语音带宽范围：300-3700Hz，8KHz 采样频率，每 20ms 编码一帧，每个帧中包含 160 个语音样点；2. AMR-WB（Adaptive Multi-Rate - Wideband Speech Codec）：语音带宽范围 50-7000Hz，16KHz 采样频率。但考虑语音的短时相关性，每帧长度均为 20ms；3. AMR-WB+（Extended Adaptive Multi-Rate - Wideband Speech Codec）：amr-nb 和 amr-wb 都属于 speech codec，对 audio 的编码效果并不好，为了提高对 audio 的编码效果，出现了 amr-wb+。amr-wb+ 可以支持更高的采样率，对 speech 和 audio 采用不同的编码算法，对 speech 采用 ACELP 编码，对 audio 采用变换编码。amr-wb+ 在低比特率上对 audio 的编码效果与 he aac + 相当。amr-wb+ 包含 amr-wb，但复杂度更高。

1.3.5 AAC 格式

AAC 是高级音频编码（Advanced Audio Coding）的缩写，出现于 1997 年，最初是基于 MPEG-2 的音频编码技术。由 Fraunhofer IIS、Dolby Laboratories、AT&T、Sony 等公司共同开发，目的是取代 MP3 格式。2000 年，MPEG-4 标准出台，AAC 重新集成了其它技术（PS，SBR），为区别于传统的 MPEG-2 AAC，故含有 SBR 或 PS 特性的 AAC 又称为 MPEG-4 AAC。

AAC 是新一代的音频有损压缩技术，它通过一些附加的编码技术（比如 PS，SBR 等），衍生出了 LC-AAC，HE-AAC，HE-AACv2 三种主要的编码，LC-AAC 就是比较传统的 AAC，相对而言，主要用于中高码率（ $\geq 80\text{Kbps}$ ），HE-AAC（相当于 AAC+SBR）主要用于中低码率（ $\leq 80\text{Kbps}$ ），而新近推出的 HE-AACv2（相当于 AAC+SBR+PS）主要用于低码率（ $\leq 48\text{Kbps}$ ），事实上大部分编码器设成 $\leq 48\text{Kbps}$ 自动启用 PS 技术，而 $> 48\text{Kbps}$ 就不加 PS，就相当于普通的 HE-AAC。