目录

1. 媒体文件的编码方式 3
1.1. 音频编码 3
1.1.1. WAV 3
1.1.2. FLAC 3
1.1.3. MP3 3
1.1.4. AAC 4
1.1.5. Opus 5
1.1.6. Vorbis 5
1.2. 视频编码 6
1.2.1. H.264/MPEG-4 AVC 6
1.2.2. H.265/HEVC 7
1.2.3. VP8 7
1.2.4. VP9 7
1.2.5. MPEG-4/Xvid/DivX 8
1.3. 媒体文件的格式 8
1.4. 光盘的格式 8
1.4.1. CD 8
1.4.2. VCD 8
1.4.3. DVD 8

1. 媒体文件的编码方式

本文介绍常见音频、视频文件的编码与格式。内容冗长，如仅需结论可只看红字部分。实际操作请参考使用ffmpeg编码媒体文件。

1.1. 音频编码

音频编码很多，其中一部分编码主要在电话中被应用，这些编码在本文不作介绍。常用于音频、视频文件的编码方式有如下几种。其中以MP3和AAC最为常用。

1.1.1. WAV

WAV是一种未经压缩的原始音频编码，通常用于录音。理论上，它是**无损**[1]的。

WAV被各系统广泛支持。

1.1.2. FLAC

FLAC是一种**无损**压缩的音频编码，压缩率大约在一半左右。FLAC目前基本被各系统支持，但部分便携式音乐播放器可能尚未支持。

该项目的主页：FLAC

1.1.3. MP3

MP3是支持最为广泛的**有损**压缩的音频编码。

MP3有一些限制：

- 最高码率320kbps。一般认为该比特率接近无损。
- 最高采样频率48kHz。

目前使用最多的开源MP3编码器是LAME。LAME编码的128kbps音频人耳测试已难以与无损音频区别。LAME提供固定码率CBR、平均码率ABR、可变码率VBR三种方式。

1.1.3.1. CBR

固定码率，有如下选择（单位：kbps）：

8, 16, 24, 32, 40, 48, 64, 80, 96, 112, 128, 160, 192, 224, 256, 320[2]

固定码率文件较大。如需进一步节省空间，建议使用VBR。

MP3 CBR建议双声道码率128kbps以上，推荐192kbps。

1.1.3.2. ABR

平均码率，介于CBR和VBR之间。

1.1.3.3. VBR

可变码率，在设定等级（lame为参数-V，ffmpeg中为参数-q:a，两者等价）后编码器会选择一定的码率范围。见下表（单位：kbps，等级0码率最高，音质最好；等级9码率最低，音质最差。

VBR等级	平均码率	码率范围
0	245	220-260
1	225	190-250
2	190	170-210
3	175	150-195
4	165	140-185
5	130	120-150
6	115	100-130
7	100	80-120
8	85	70-105
9	65	45-85

如何选择等级：

- 等级0～3听觉无损。
- 等级4为默认。
- 等级6为可以接受。

MP3 VBR建议选择2级。 LAME在高码率时表现较佳。如码率较低，建议选择其他编码，如AAC。

1.1.4. AAC

AAC是一种**有损**音频压缩，最常见于MP4文件中的音频编码，在MKV文件中也常用。AAC有多种编码器，其中推荐使用Fraunhofer FDK AAC。

AAC的压缩比很高，可达到约18:1～20:1。支持多声道。最高采样频率为96kHz。AAC有多种规格，最主要的是AAC-LC、AAC-HE和AAC-HEv2。（以下码率如无说明，均指双声道。非双声道时根据声道数计算。比如5.1环绕声为6声道，将每声道的码率乘以声道数即可。）

- 低复杂度LC（为最常用的规格，码率在96～192kbps范围时，音质、压缩比均比较平衡。最广泛。
- 高性能HE（v1）：除LC的功能外，添加了谱带复制（SBR）技术，适合48～64kbps以下的低码率场合。常用于数字广播。
- HEv2：在HEv1的基础上增加了参数化双声道（PS）技术，适合码率低至32kbps的场合。**仅支持双声道**。常用于数字广播。

注意：支持AAC的播放器可能仅支持AAC-LC。

FDK AAC对以上三种规格均支持，FDK AAC有CBR、VBR两种方式。

1.1.4.1. CBR

每个声道建议64kbps，即双声道128kbps，5.1环绕声384kbps。

- AAC-LC，建议双声道码率128kbps以上。
- AAC-HE，建议双声道码率48～80kbps。
- AAC-HEv2，建议双声道码率48kbps以下。

1.1.4.2. VBR

在设定等级后编码器会选择相应的码率，见下表（单位：kbps，）等级5为最高，1为最低。

VBR等级	每声道码率	可指定规格
1	20-32	LC,HE,**HEv2**
2	32-40	LC,HE,**HEv2**
3	48-56	LC,**HE**,HEv2
4	64-72	**LC**
5	96-112	**LC**

LC为默认规格。当码率较低时，指定profile为aac_he或aac_he_v2会获得更高压缩比。上表加粗部分为该码率范围内的推荐规格。

1.1.5. Opus

Opus编码也是**有损**压缩，其尚未在媒体文件中被广泛使用，但在任何码率中都全面领先现有的其他编码，可能在未来的媒体文件中被广泛应用。目前，Opus在语音通话中已成为首选编码。

为获得较好音质，双声道码率至少32kbps，建议64kbps以上。

1.1.6. Vorbis

Vorbis是一个完全无专利的开源音频编码方式，每声道的码率范围16～128kbps，Vorbis支持固定码率和可变码率，这里仅介绍可变码率，VBR的指定方法类似MP3，但数字越大码率越高。见下表（单位：kbps）

VBR等级	码率范围
0	64-80
1	80-96
2	96-112
3	112-128
4	128-160
5	160-192
6	192-224
7	224-256

VBR等级	码率范围
8	256-320
9	320-500
10	500-1000

如何选择等级：

- 等级5以上认为与无损格式几乎无区别。
- 等级3为默认。
- 等级3～6的质量都比较好。

总结：如何选择音频编码

- 无损：FLAC
- 最兼容：MP3
- 兼容且音质好：AAC-LC
- 体积小、省流量且音质可：AAC-HE、AAC-HEv2
- 不考虑兼容性，有损编码中的最高音质：Opus

还有一些编码未介绍。这些编码的音质排序大致如下（名称按ffmpeg中的命名）：

libopus > libvorbis ≥ libfdk_aac > libmp3lame ≥ eac3/ac3 > libtwolame

未在此列的编码不建议使用。

1.2. 视频编码

视频编码较多，且多数为有损压缩算法，故视频编辑时应以原视频而不是导出（渲染）的视频为基础进行改动。视频编码对机器性能要求高，目前主流编码方式一般可经GPU加速运算，推荐使用合适的显卡，并正确安装含硬件解码功能的驱动程序。以下仅介绍部分常见的视频编码。

1.2.1. H.264/MPEG-4 AVC

H.264是目前**最常见**的视频编码，在卫星电视、视频网站（包括直播）等方面均被广泛使用。推荐使用的编码器为x264。如果有硬件支持，则推荐使用h264_vaapi（Intel）或h264_nvenc（NVIDIA）等。不同编码器的参数不同。以下分别详述libx264、h264_vaapi、h264_nvenc三个编码器应了解的参数设置。

1.2.1.1. libx264

libx264有CRF和Two-Pass ABR两种码率控制。前者指定所需要的质量等级，无法知道文件大小；后者则指定文件大小。故质量优先时选前者，文件大小有限制时选后者。这里不介绍后一种方式。

1. CRF
    - 范围0～51[3]，0为无损，51最差；
    - 建议17～28；
    - 17～18视觉无损；
    - 默认23；

    - 每+6，文件大小加倍。
2. preset
    - 可选择（ultrafast, superfast, veryfast, faster, fast, medium, slow, slower, veryslow, placebo）
    - 默认medium）
    - 不要选择placebo（此时压缩率提升已经很有限，但用时更长，无意义；
    - 越慢则压缩率越高，根据需求和耐心选择。
3. tune
    - 用于特殊情况下的优化，可选项见下，可多选（以逗号隔开）：
        1. film（用于高质量电影，减少马赛克；
        2. animation（用于卡通，进一步减少马赛克，增加参考帧；
        3. grain（保留老电影的颗粒感；
        4. stillimage（适合类似幻灯片的静止画；
        5. fastdecode（通过禁止部分滤镜以加速解码；
        6. zerolatency（快速编码，适合低延迟直播；
4. profile
    - 该选项为了兼容性[4)]考虑，通常选择high（如不兼容，改为main（选择baseline可兼容所有播放器。

1.2.1.2. h264_vaapi

利用Intel集成显卡中的硬件编码以加速，基于vaapi库。另有一个闭源库libmfx（利用Intel Quick Sync技术，在Linux平台限制多且使用较麻烦，不建议使用。

需要额外指定的参数：

```
-vaapi_device /dev/dri/renderD128 -vf 'format=nv12,hwupload'
```

可调整的主要参数是-qp（范围0～52，默认20。其含义同libx264中的crf（其他参数按默认即可。

按上述方法，编码时就利用硬件加速了，但这里仅仅是编码时加速；如果源文件的编码被硬件支持，则在读入文件过程中也可利用硬件加速。其具体方法以及上述额外参数的含义，见使用ffmpeg编码媒体文件（

1.2.1.3. h264_nvenc

利用NVIDIA显卡中的硬件编码以加速。这里假定文件用于保存、视频网站投稿等非直播情况，简单粗暴地给出质量最佳且兼容较好的参数。

- -preset slow
- -profile yuv420p
- -cq （同libx264的参数crf（范围0～51）

1.2.2. H.265/HEVC

1.2.3. VP8

1.2.4. VP9

1.2.5. MPEG-4/Xvid/DivX

1.3. 媒体文件的格式

通常，扩展名即表示了媒体文件的格式。需要注意的是，尽管部分文件的扩展名与该文件的编码相同，但一些格式仅仅是一种“容器”，其支持的编码不止一种，因而在扩展名相同的时候，也可能因为系统缺少一些解码器而不能播放所有文件。下表列举了常见格式及其支持的编码，加粗部分为ffmpeg默认编码。

格式（扩展名）	视频编码	音频编码
MP4 (.mp4, .m4a[5)])	**H.264**, MPEG-4, MPEG-2, MPEG-1	**AAC**, MP3, others
AVI (.avi)	**MPEG-4**, any	**MP3**, any
Matroska (.mkv)	**H.264**, any	**Vorbis**, any
WebM (.webm)	**VP9**, VP8	**Opus**, Vorbis
Ogg Video (.ogg)	**Theora**, Dirac	**Vorbis**, FLAC
MPEG-1 (.mpg, .mpeg)	**MPEG-1**	MP3, **MP2**
MPEG-2 (.mpg, .mpeg)	MPEG-2	MP2, AAC
RealMedia (.rm)	RealVideo	AC-3

1.4. 光盘的格式

1.4.1. CD

音频：双声道，16位PCM（44.1kHz）

1.4.2. VCD

1.4.2.1. 视频

编码：MPEG-1 分辨率：NTSC 352×240，PAL 352×288 帧率：NTSC 29.97，PAL 25 码率：1150kbps

1.4.2.2. 音频

双声道 编码：MP2 采样频率：44.1kHz 码率：224kbps

1.4.3. DVD

1)
如果是其他已经有损格式转换成WAV格式，就不是无损了。
2)
320kbps是比较浪费的。用最高码率的有损压缩不如用无损压缩。
3)
x264为8位颜色，为多数情况；少数情况为10位，此时范围0～63。
4)
通常还需要指定参数-pix_fmt yuv420p以兼容多数设备，尽管这样做可能导致失真。

[5]

本质上和mp4相同，但只包含音频。改为mp4没有问题。

From:
https://irdya.top/ - **漂流記**

Permanent link:
https://irdya.top/zh/misc/codec

Last update: **2022/05/26 03:24**