

音视频学习从零到整(2)

音视频学习笔记出自:CC老师

- 音视频学习,越到实用环节,你会发现一些细节部分知识的缺失,可能会导致整个案例的理解出现很严重的错误.
- 希望大家学习过程,对技术保持一种匠人精神.沉淀下来,学习.

音视频学习从零到整(1)

音频编码

一.为什么要做音频编码?

之前的文章中,我带着大家来计算过CD音质的数据采集,每分钟需要存储空间约为10.1MB.从存储的角度或者网络实时传播的角度.这个数据量都是太大了.对于存储和传输都是非常具有挑战的.所以我们需要通过压缩编码

压缩编码的可能性

压缩编码的基本指标就是压缩比,压缩比通常小于1(如果等于或者大于1,是不是就失去的压缩的意义了,压缩目的就是为了减少数据体量).压缩算法分为2种,有损压缩和无损压缩.

- 无损压缩:解压后的数据可以完全复原.在常用的压缩格式中,用的较多的都是有损压缩.
- 有损压缩:解压后的数据不能完全复原,会丢失一部分信息.压缩比越小,丢失的信息就会越多,信号还原的失真就会越大.

需要根据不同的场景(考虑因素包括存储设备,传输网络环境,播放设备等),可以选用不同压缩编码算法.

压缩编码的原理实际上就是压缩冗余的信号.冗余信号就是指不能被人耳感知的信号.包括人耳听觉范围之外的音频信号以及被掩盖掉的音频信号.

- 拓展小课堂

人耳掩盖效应

- 主要表现在频域掩盖效应与时域掩盖效应.无论是在时域还是频域上,被掩盖掉的信息都认为是冗余信息,不进行编码处理
- 掩蔽效应指人的耳朵只对最明显的声音反应敏感,而对于不明显的声音,反应则较不为敏感。例如在声音的整个频率谱中,如果某一个频率段的声音比较强,则人就对其它频率段的声音不敏感了。应用此原理,人们发明了mp3等压缩的数字音乐格式,在这些格式的文件里,只突出记录了人耳朵较为敏感的中频段声音,而对于较高和较低的频率的声音则简略记录,从而大大压缩了所需的存储空间。在人们欣赏音乐时,如果设备对高频响应得比较好,则会使人感到低频响应不好,反之亦然。

常用压缩编码格式

WAV编码

WAV编码的一种实现方式(其实它有非常多实现方式,但都是不会进行压缩操作).就是在源PCM数据格式的前面加上44个字节.分别用来描述PCM的采样率,声道数,数据格式等信息.

- 特点:音质非常好,大量软件都支持其播放
- 适合场合:多媒体开发的中间文件,保存音乐和音效素材

MP3编码

MP3编码具有不错的压缩比,而且听感也接近于WAV文件,当然在不同的环境下,应该调整合适的参数来达到更好的效果.

* 特点:音质在128Kbit/s以上表现不错,压缩比较高.大量软件和硬件都支持.兼容性高.

* 适合场合:高比特率下对兼容性有要求的音乐欣赏.

AAC编码

AAC是目前比较热门的有损压缩编码技术,并且衍生了LC-AAC,HE-AAC,HE-AAC v2 三种主要编码格式.

- LC-AAC 是比较传统的AAC,主要应用于中高码率的场景编码($\geq 80\text{Kbit/s}$)
- HE-AAC 主要应用于低码率场景的编码($\leq 48\text{Kbit/s}$)
- 特点:在小于 128Kbit/s 的码率下表现优异,并且多用于视频中的音频编码
- 适合场景:于 128Kbit/s 以下的音频编码,多用于视频中的音频轨的编码

Ogg编码

Ogg编码是一种非常有潜力的编码,在各种码率下都有比较优秀的表现.尤其在低码率场景下.Ogg除了音质好之外,Ogg的编码算法也是非常出色.可以用更小的码率达到更好的音质. 128Kbit/s 的Ogg比 192Kbit/s 甚至更高码率的MP3更优质.但目前由软件还是硬件支持问题,都没法达到与MP3的使用广度.

- 特点:可以用比MP3更小的码率实现比MP3更好的音质,高中低码率下均有良好的表现,兼容不够好,流媒体特性不支持.
- 适合场景:语言聊天的音频消息场景

总结

简单介绍了为什么需要做音频编码?音频压缩编码的可能性来自哪里?音频的数据冗余信息以及压缩编码格式的使用场景.