

# 音视频常用基础了解

## 音频

首先我们简单了解一下声音的三要素

1. 音调：也可以理解为音频，主要与声波的频率有关，频率高则音调高。
2. 音量：主要与声波震动的幅度有关，幅度大则音量高。
3. 音色：主要与声波的波形相关。

## 采样位数

一个范围，也就是每次采样，采集到的数据值的范围，比如常用的16位，也就是采样值取值范围就是-32768--32767

## 采样频率

类似于帧率，也就是音频在每秒内的打点数。

## 声道

声音录制时的音源数量（也就是麦克风的个数）。

## 码率

一秒内音频数据的大小

计算方式：

比特率 (bit) = 声道 \* 位数 \* 频率

码率 (byte) = 比特率 / 8

## 常用音频格式：

wav 格式：音质高 无损格式 体积较大。  
mp3 格式：有损压缩，文件较小。  
aac 格式：相对于 mp3，aac 格式的音质更佳，文件更小，有损压缩。  
mpeg 格式：有损压缩，但是它的最大优势是以极小的声音失真换来了较高的压缩比。  
amr 格式：压缩比较大。  
flac 格式：无损压缩。  
cd 格式：高保真，类似于原声。

## 音频帧

音频不像视频那样可以把每一帧画面抽离出来，也就是说音频所说的帧率，只不过是我們为了与视频帧相对应在特定时间段内的数据体现。在实时应视频领域比较明显。

# 视频

## 帧率

每秒内采集的视频图片数量。

## 码率

一秒内视频数据的大小

## 分辨率

每一帧视频的分辨率

## IPB帧

PB帧主要是在实时音视频用中的传输层做优化使用，减少传出层数据压力，但同样会增大客户端的处理工作量。

## I帧

关键帧

## P帧

向前预测帧，主要保存于关键帧之间的数据差异

## B针

双向预测针，主要保存本帧与前后帧的差别

## 编码类型

### 软编码

利用CPU进行编码的过程，通常使用开源库FFmpeg。

### 硬编码

使用GPU、芯片等进行编码，性能高，对硬件要求较高。像VideoToolBox, AudioToolbox这两大框架。

### 软编码和硬编码的区别

个人理解软编码方法调用等比较方便，缺点就是各种参数比较多，不熟悉比较麻烦，虽然他的处理速度相对于硬编码要快，但CPU损耗比较严重，部分老设备效果差距较大。个人建议对码率帧率要求比较高的不建议使用。硬编虽然使用GPU等来处理数据，但还是依赖于CPU的计算，这也就是相对于软编码来说处理速度慢的原因之一。

## 视频帧（图片存储）格式

常见个格式有YUV，RGB, NV12，NV21等

### YUV

通常分为平面格式，半平面格式以及打包格式，苹果推荐默认格式，其中Y代表“灰度值”，UV代表“色度值”，可以理解为两或三个平面叠加组成一个图片。相对一起他格式特点是小，便于传输。

### RGB

按照三原色来表示彩色图像的。相对于YUV占用带宽比较大，但处理相对于YUV比较方便。

### NV

NV21Android默认格式与NV21 相差不大。

## 视频编解码器标准

H.264 , H.265, H.265标准在H.264基础上, 进行优化, 相同时间长度视频大小仅有之前的一半左右。

## Video Audio Track

Video Track: 视频通道存储视频影像, 经常用于视频的转场滤镜

Audio Track: 音轨, 也可理解为声音通道, 多用于声音的处理, 变声, 倍速播放等  
在常用的处理中, track可以理解为一个队列使用, 他同样具有队列的基本属性 (主要是增删改属性)

## Header

媒体数据的头, 主要包含媒体的信息, 像码率分辨率等。分析媒体的主要信息来源

## 流媒体协议:

主要有: RTMP RTP RTSP HLS FLV http https等, 常见使用较多的RTMP.

点播: 使用http https 较多

直播: rtmp , 可以说是为直播专门打造。