linesize 一行数据的宽度，例如yuv480p其中y是一行，uv是一行，分辨率处于480p的时候，那么y的linesize就是480，uv加起来是480。

传统的integer是以32位来表示的，存储需要4个字节，当如果整数大小在256以内，那么只需要用一个字节就可以存储这个整数，这样就可以节省3个字节的存储空间，Google varint就是根据这种思想来序列化整数的。Varint 中的每个 byte 的最高位 bit 有特殊的含义，如果该位为 1，表示后续的 byte 也是该数字的一部分，如果该位为 0，则结束。其他的 7 个 bit 都用来表示数字。因此小于 128 的数字都可以用一个 byte 表示。大于 128 的数字，会用两个字节。

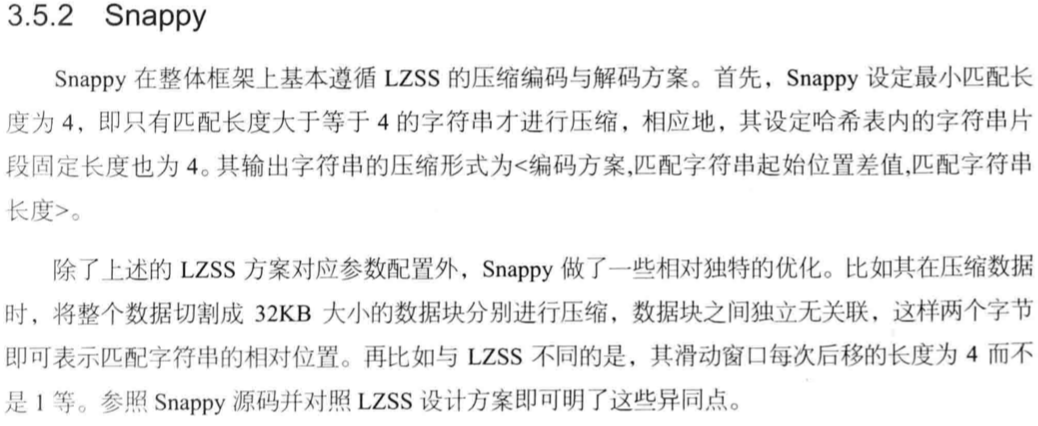
动态Huffuman压缩比静态好在一开始的树层数小，而且不需要将树作为数据加入压缩文件。

对于中文而言，Huffuman按照字节来压缩而不是按照字符为单位压缩比较好，因为字节的重复性更大。可以推广：英文字符按照概率从大到小排，观察1-8位之间什么情况，以多少位为单位能够得到最好的压缩？当然还要考虑位的压缩比，比如你是讲4位压缩到1位，而以字节为单位能够将8位压缩到1位。

多媒体老师的ppt中记录不少压缩算法



LZSS这种限制能够提高压缩与解压效率



Snappy是google开源的在合理压缩率基础上追求压缩与解压速度的压缩算法

base64编码：

由于不同的设备对字符的处理方式有一些不同，这样那些不可见字符就有可能被处理错误，这是不利于传输的。所以就先把数据先做一个Base64编码，统统变成可见字符，这样出错的可能性就大降低了。也就是说主要是为了处理不可见字符，但是这样也就是意味着迅雷和旋风下载链接使用base64编码有着其他的用途，可能就是想让一般人不懂或者使得自己的链接是本软件专用的。

迅雷的“专用地址”也是用Base64加密的。但是不是直接进行的，http://地址的前后分别添加AA和ZZ，之后再进行base64加密。

QQ旋风地址的编码更简单，不加任何字符，直接就对地址进行Base64编码了

固实压缩文件是 RAR 的一种特殊压缩方式存储的压缩文件，它把压缩文件中的全部文件都当成一个连续数据流来看待。固实压缩可增加压缩性能，特别是在添加大量的小文件的时候。

Gap Encoding类似压缩原理：

Given a sorted list of integers x, y, z, ⋯, represent by x, y-x, z-y, ⋯

---------------------------------------------------------------------------------------------------------------

图像的png格式也会有几种不同的压缩办法，jpg也有不同年代版本的压缩算法，png里面比较推荐的是tinypng（https://www.v2ex.com/t/151976）

颜色体系有两种，RGB和CMY，电脑等就是用RGB，印刷机等使用CMY。CMY从颜色混合来看就是RGB的缩减版。

---------------------------------------------------------------------------------------------------------------

音频编码的几种格式：raw（未压缩的pcm或wav格式）、speex（speex格式）、speex-wb（宽频speex格式）

web api比如腾讯云都会直接在下方简单介绍使用了什么技术，类似平台还有 speakin、科大飞讯、百度语音（百度云）、阿里云、Google、微软。腾讯云说他们的声纹识别可以拒绝仿冒录音回放的欺诈，基于声纹特征和说话内容同时验证。

识别音色，声纹识别(Voiceprint Recognition, VPR)，也称为说话人识别(Speaker Recognition)。通过机器学习识别声音没有机器学习识别图像容易，因为数据集上难得。有两类发展方向，即说话人辨认(Speaker Identification)和说话人确认(Speaker Verification)。前者用以判断某段语音是若干人中的哪一个所说的，是“多选一”问题；而后者用以确认某段语音是否是指定的某个人所说的，是“一对一判别”问题。音色本身是一个很笼统的概念，是指不同的声音的频率表现在波形方面总是有与众不同的特性，音色上的差异在信号处理的层面可以表示为在频域不同频段能量分布情况的差异。采用目前最先进的声纹技术Ivector技术，配合PLDA技术来降低信道差异的影响，提升识别性能，文本相关声纹认证实验性能达到等错率0.7%。声纹识别的结果是不可能直接弄成一串Hash之类的数字，因为特征提取本身是不完全的，所以对于每一段语音，最终得到的“Hash”都会不同，所以需要建立一个模型来帮助识别。

通过观察我们的频谱峰值和合并的峰值频率以及它们之间的时间差，我们可以得到一个可以当作歌曲的唯一数字指纹的的哈希。

VAD（Voice Activity Detection），通常作为对话切割的依据。

---------------------------------------------------------------------------------------------------------------

一个视频播放器包含：文件格式解析（Demuxer）、音视频解码（Decoder）、音视频渲染输出（Renderer），FFmpeg中都有对应的实现——avformat、avcodec、avdevice。

视频容器： MP4只是一个封装格式，它里面可能是MPEG4编码，也可能是H.264，或者H.265。MP4、rmvb、rm、flv、AVI、mov、WMV、mkv

视频部分与音频部分是通过数据块的时间戳来实现同步。如果音频慢了就会被丢弃一部分。

AVC是高级视频编码的简称（Advanced Video Coding），视频编解码技术有两套标准，H.261、国际电联（ITU-T）的标准H.263、H.263+等；还有ISO 的MPEG标准MPEG-1、MPEG-2、MPEG-4等等。HEVC是High Efficiency Video Coding的缩写，是一种新的视频压缩标准，用来以替代H.264/AVC编码标准，2013年1月26号，HEVC正式成为国际标准。

MPEG标准，现在到了MPEG-4

h264/AVC H.264 is a codec used to create MP4 files

h265/HEVC （265也是在264上修改优化的，有专利限制）265的编解码器现在都已经有了，但是编码器效率太低。

上次参加了个avs报告，自主的AVS2比h.265还好

有一次看阿里云直播，中间出了问题崩了，结果浏览器即使别人好了我也要过挺久的一段时间才好。不过切手机倒是能“马上”好，估计是川大的网络情况原因。