**H264格式**

一、概述

H264基本流（elementary stream,ES）分为两层：

1. VCL（video coding layer），视频编码层，对视频采用空间和时间等冗余算法进行编码压缩，输出编码后的裸数据bit流SODB。
2. NAL（network abstraction layer），网络提取层，对编码数据打包和传送，用于网络传输。NAL将SODB打包成RBSP（字节流）然后加上NAL头，组成一个NAL unit。

引入NAL并使之与VCL分离带来的好处包括两方面：1、使信号处理和网络传输分离，VCL和NAL可以在不同的处理平台上实现；2、VCL 和NAL 分离设计，使得在不同的网络环境内，网关不需要因为网络环境不同而对VCL比特流进行重构和重编码。

SODB与RBSP定义：

SODB（String Of Data Bits）：原始数据比特流, 长度不一定是8的倍数，故需要补齐

RBSP（Raw Byte Sequence Payload）：原始数据字节流，SODB+RBSP trailing bits=RBSP，添加trailing bits是为了使一个RBSP为整字节数，把SODB尾部字节对齐后得到RBSP。

H264最终由连续的NALU组成，每个NALU前面有一个起始码0x00000001（四字节表示一帧图片的第一个NAL起始码）或0x000001（三字节表示一帧图片的非第一个NAL的起始码）。为了防止数据中存在0x000001或0x00000001干扰起始码的判断，编码器会对原始数据中出现0x000001或0x00000001的地方在倒数第二个字节加上03，比如0x00000301，解码使去掉，专业说法：脱壳。

1. NALU头

NALU头长度一个字节：

F（forbidden\_zero\_bit）：1 位，初始为0。当网络识别此单元存在比特错误时，可将其设为 1，以便接收方丢掉该单元。校验位。

NRI（nal\_ref\_idc）：2 位，用来指示该NALU 的重要性等级。值越大，表示当前NALU越重要。如果为0表示该nal没有预测，丢失后不会导致误码扩散。

Type（nal\_unit\_type）：5 位，指出NALU 的类型：

0：未指定

1：标识是否为IDR的片，该值为0表示非IDR

2 3 4：数据分区 A B C，IDR图像不能分区

5：IDR图像的片

6：SEI

7：SPS 序列参数集

8：PPS 图像参数集

关于sps和pps：每个VCL NAL单元包含一个标识，指向有关的图像参数集，每个图像参数集包含一个标识，指向有关的序列参数集，因此，只用少数的指针信息，减少每个VCL NAL单元重复传送信息。

每一帧的解码都需传给解码器sps和pps。SPS 包含的是针对一连续编码视频序列的参数，如标识符 seq\_parameter\_set\_id、帧数及 POC 的约束、参考帧数目、解码图像尺寸和帧场编码模式选择标识等等。PPS对应的是一个序列中某一幅图像或者某几幅图像，其参数如标识符 pic\_parameter\_set\_id、可选的 seq\_parameter\_set\_id、熵编码模式选择标识、片组数目、初始量化参数和去方块滤波系数调整标识等等。

1. 片 slice

一帧图像可编码成一个或者多个片，每片包含整数个宏块（macro block），即每片至少一个宏块，最多时包含整个图像的宏块。

分片的目的：为了限制误码的扩散和传输，使编码片相互间保持独立，所以每个片可以单独解码，不依赖其它片。

片共有5种类型：I片（只包含I宏块）、P片（可包含P和I宏块）、B片（B和I宏块）、SP片（用 于不同编码流之间的切换）和SI片（特殊类型的编码宏块）。

每个分片也包含着头和数据两部分：1、分片头中包含着分片类型、分片中的宏块类型、分片帧的数量、分片属于那个图像以及对应的帧的设置和参数等信息。2、分片数据中则是宏块，这里就是我们要找的存储像素数据的地方。

1. 宏块

宏块是视频信息的主要承载者，因为它包含着每一个像素的亮度和色度信息。视频解码最主要的工作则是提供高效的方式从码流中获得宏块中的像素阵列。

组成部分：一个宏块由一个16×16亮度像素和附加的一个8×8 Cb和一个 8×8 Cr 彩色像素块组成。每个图象中，若干宏块被排列成片的形式。宏块分为I、P、B宏块，I宏块只能利用当前片中已解码的像素作为参考进行帧内预测；P宏块可以利用前面已解码的图像作为参考图像进行帧内预测；B宏块则是利用前后向的参考图形进行帧内预测。

宏块中包含了宏块类型、预测类型、Coded Block Pattern、Quantization Parameter、像素的亮度和色度数据集等等信息。

编码后视频的每一组图像（GOP，图像组）都给予了传输中的序列（SPS）和本身这个帧的图像参数（PPS），SPS与IDR有关系，IDR是一个视频序列的开始。

1. IDR帧和I帧的关系

IDR帧就是I帧，但是I帧不一定是IDR帧，在一个完整的视频流单元中第一个图像帧是IDR帧，IDR帧是强制刷新帧，在解码过程中，当出现了IDR帧时，要更新sps、pps，原因是防止前面I帧错误，导致sps，pps参考I帧导致无法纠正。

1. GOP

GOP是画面组，一个GOP是一组连续的画面。GOP一般有两个数字，如M=3，N=12.M制定I帧与P帧之间的距离，N指定两个I帧之间的距离。增大图片组能有效的减少编码后的视频体积，但是也会降低视频质量，至于怎么取舍，得看需求了

1. IDR

一个序列的第一个图像叫做 IDR 图像（立即刷新图像），IDR 图像都是 I 帧图像。I和IDR帧都使用帧内预测。I帧不用参考任何帧，但是之后的P帧和B帧是有可能参考这个I帧之前的帧的。IDR就不允许这样。

核心作用：

H.264 引入 IDR 图像是为了解码的重同步，当解码器解码到 IDR 图像时，立即将参考帧队列清空，将已解码的数据全部输出或抛弃，重新查找参数集，开始一个新的序列。这样，如果前一个序列出现重大错误，在这里可以获得重新同步的机会。IDR图像之后的图像永远不会使用IDR之前的图像的数据来解码。