

H.264 码流结构解析

1. H.264 简介

MPEG（Moving Picture Experts Group）和 VCEG（Video Coding Experts Group）已经联合开发了一个比早期研发的 MPEG 和 H.263 性能更好的视频压缩编码标准，这就是被命名为 AVC（Advanced Video Coding），也被称为 ITU-T H.264 建议和 MPEG-4 的第 10 部分的标准，简称为 H.264/AVC 或 H.264。这个国际标准已经与 2003 年 3 月正式被 ITU-T 所通过并在国际上正式颁布。为适应高清视频压缩的需求，2004 年又增加了 FRExt 部分；为适应不同码率及质量的需求，2006 年又增加了可伸缩编码 SVC。

2. H.264 编码格式

H.263 定义的码流结构是分级结构，共四层。自上而下分别为：图像层(picturelayer)、块组层(GOB layer)、宏块层(macroblock layer)和块层(block layer)。而与 H.263 相比，H.264 的码流结构和 H.263 的有很大的区别，它采用的不再是严格的分级结构。

H.264 支持 4:2:0 的连续或隔行视频的编码和解码。H.264 压缩与 H.263、MPEG-4 相比，视频压缩比提高了一倍。

H.264 的功能分为两层：视频编码层（VCL, Video Coding Layer）和网络提取层（NAL, Network Abstraction Layer）。VCL 数据即编码处理的输出，它表示被压缩编码后的视频数据序列。在 VCL 数据传输或存储之前，这些编码的 VCL 数据，先被映射或封装进 NAL 单元中。每个 NAL 单元包括一个原始字节序列负荷（RBSP, Raw Byte Sequence Payload）、一组对应于视频编码的 NAL 头信息。RBSP 的基本结构是：在原始编码数据的后面填加了结尾比特。一个 bit “1” 若干比特 “0”，以便字节对齐。



图 1 NAL 单元序列

3. H.264 传输

H.264 的编码视频序列包括一系列的 NAL 单元，每个 NAL 单元包含一个 RBSP，见表 1。编码片（包括数据分割片 IDR 片）和序列 RBSP 结束符被定义为 VCL NAL 单元，其余为 NAL 单元。典型的 RBSP 单元序列如图 2 所示。每个单元都按独立的 NAL 单元传送。单元的信息头（一个字节）定义了 RBSP 单元的类型，NAL 单元的其余部分为 RBSP 数据。

SPS	SEI	PPS	I片	图像定界符	P片	P片
-----	-----	-----	----	-------	----	----

图 2 RBSP 序列举例

RBSP 类型	描述
参数集 PS	序列的全局，如图像尺寸、视频格式等
增强信息 SEI	视频序列解码的增强信息
图像定界符 PD	视频图像的边界
编码片	片的头信息和数据
数据分割	DP 片层的数据，用于错误恢复解码

序列结束符	表明下一图像为 IDR 图像
流结束符	表明该流中已没有图像
填充数据	哑元数据，用于填充字节

表 1 RBSP 描述

4. H.264 码流结构图

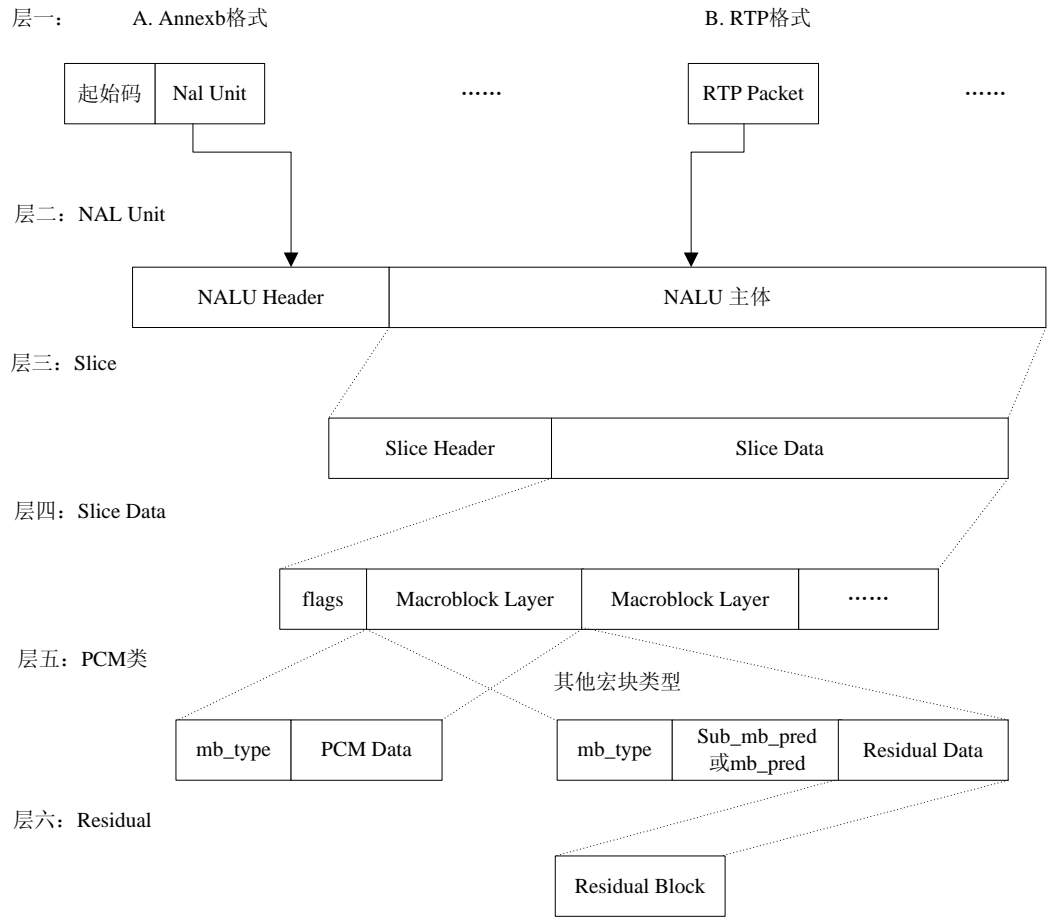


图 3 H.264 码流分层结构

起始码：如果 NALU 对应的 Slice 为一帧的开始，则用 4 字节表示，即 0x00000001；否则用 3 字节表示，0x000001。

NAL Header: forbidden_bit, nal_reference_bit（优先级），nal_unit_type（类型）。

脱壳操作：为了使 NALU 主体不包括起始码，在编码时每遇到两个字节（连续）的 0，就插入一字节 0x03，以和起始码相区别。解码时，则将相应的 0x03 删除掉。

NALU 类型：

Nal_unit_type	NAL 类型	C
0	未使用	
1	不分区、非 IDR 图像的片	2, 3, 4
2	片分区 A	2
3	片分区 B	3
4	片分区 C	4
5	IDR 图像中的片	2, 3
6	补充增强信息单元(SEI)	5

7	序列参数集	0
8	图像参数集	1
9	分界符	6
10	序列结束	7
11	码流结束	8
12	填充	9
13~23	保留	
24~31	未使用	

表 2 nal_unit_type 语义

sub_mb_pred 和 mb_pred: 运动补偿或者是帧内预测。

5. H.264 解码

NAL 头信息的 nal_reference_idc (NRI) 用于在重建过程中标记一个 NAL 单元的重要性，值为 0 表示这个 NAL 单元没有用预测，因此可以被解码器抛弃而不会有错误扩散；值高于 0 表示 NAL 单元要用于无漂移重构，且值越高，对此 NAL 单元丢失的影响越大。

NAL 头信息的隐藏比特位，在 H.264 编码器中默认为 0，当网络识别到单元中存在比特错误时，可将其置为 1。隐藏比特位主要用于适应不同种类的网络环境（比如有线无线相结合的环境）。

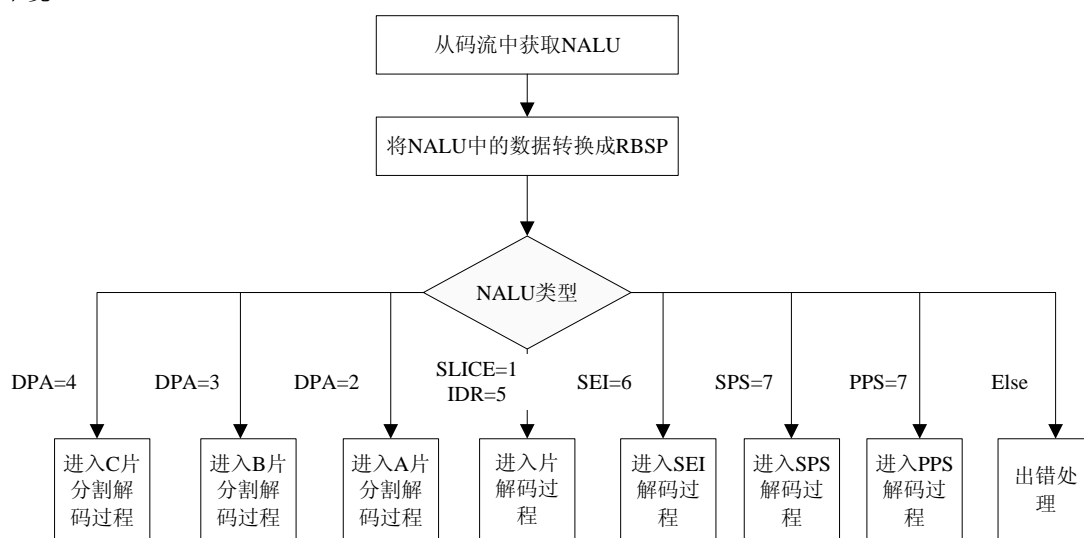


图 4 NAL 单元解码

NAL 单元解码的流程为：首先从 NAL 单元中提取出 RBSP 语法结构，然后按照如图 4 所示的流程处理 RBSP 语法结构。输入的是 NAL 单元，输出结果是经过解码的当前图像的样值点。

NAL 单元中分别包含了序列参数集和图像参数集。图像参数集和序列参数集在其他 NAL 单元传输过程中作为参考使用，在这些数据 NAL 单元的片头中，通过语法元素 pic_parameter_set_id 设置它们所使用的图像参数集编号；而相应的每个图像参数集中，通过语法元素 seq_paramter_set_id 设置他们使用的序列参数集编号。

6. 各分层结构的语法元素参考 G50 标准。