【H.264/AVC视频编解码技术详解】三. H.264简介

blog.csdn.net/shaqoneal/article/details/52081021

《H.264/AVC视频编解码技术详解》视频教程已经在"CSDN学院"上线,视频中详述了H.264的背景、标准协议和实现,并通过一个实战工程的形式对H.264的标准进行解析和实现,欢迎观看!

"纸上得来终觉浅,绝知此事要躬行",只有自己按照标准文档以 代码的形式操作一遍,才能对视频压缩编码标准的思想和方法有 足够深刻的理解和体会!

链接地址:H.264/AVC视频编解码技术详解

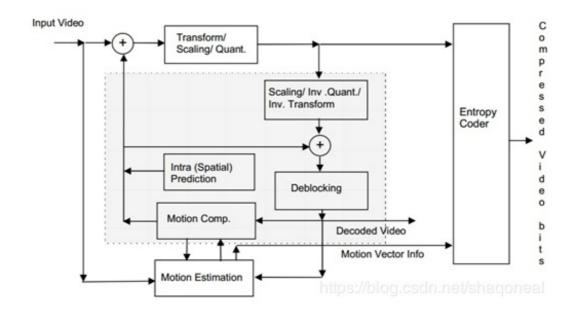
本节视频免费

一. H.264视频编码标准

H.264视频编码标准是ITU-T与MPEG合作产生的又一巨大成果,自颁布之日起就在业界产生了巨大影响。严格地讲,H.264标准是属于MPEG-4家族的一部分,即MPEG-4系列文档ISO-14496的第10部分,因此又称作MPEG-4/AVC。同MPEG-4重点考虑的灵活性和交互性不同,H.264着重强调更高的编码压缩率和传输可靠性,在数字电视广播、实时视频通信、网络流媒体等领域具有广泛的应用。

二. H.264视频编码方法简介

在整体的编码框架方面,H.264依然采用了与前期标准类似的结构,即块结构的混合编码框架。其主要结构图如下图所示:



在H.264进行编码的过程中,每一帧的H图像被分为一个或多个**条带**(slice)进行编码。每一个条带包含多个**宏块**(MB,Macroblock)。宏块是H.264标准中基本的编码单元,其基本结构包含一个包含16×16个亮度像素块和两个8×8色度像素块,以及其他一些宏块头信息。在对一个宏块进行编码时,每一个宏块会分割成多种不同大小的子块进行预测。帧内预测采用的块大小可能为16×16或者4×4,帧间预测/运动补偿采用的块可能有7种不同的形状:16×16、16×8、8×16、8×8、8×4、4×8和4×4。相比于早期标准只能按照宏块或者半个宏块进行运动补偿,H.264所采用的这种更加细分的宏块分割方法提供了更高的预测精度和编码效率。在变换编码方面,针对预测残差数据进行的变换块大小为4×4或8×8(仅在FRExt版本支持)。相比于仅支持8×8大小的变换块的早期版本,H.264避免了变换逆变换中经常出现的失配问题。

H.264标准中采用的熵编码方法主要有上下文自适应的变长编码CAVLC和上下文自适应的二进制算数编码CABAC,根据不同的语法元素类型指定不同的编码方式。通过这两种熵编码方式达到一种编码效率与运算复杂度之间的平衡。

同前期标准类似,H.264的条带也具有不同的类型,其中最常用的有I条带、P条带和B条带等。 另外,为了支持码流切换,在扩展档次中还定义了SI和SP片。

• |条带:帧内编码条带,只包含|宏块;

● P条带:单向帧间编码条带,可能包含P宏块和I宏块;

● B条带:双向帧间编码条带,可能包含B宏块和I宏块;

视频编码中采用的如预测编码、变化量化、熵编码等编码工具主要工作在slice层或以下,这一层通常被称为**"视频编码层"(Video Coding Layer, VCL)。相对的,在slice以上所进行的数据和算法通常称之为"网络抽象层"(Network Abstraction Layer, NAL)**。设计定义NAL层的主要意义在于提升H.264格式的视频对网络传输和数据存储的亲和性。

为了适应不同的应用场景,H.264也定义了三种不同的档次:

- 基准档次(Baseline Profile):主要用于视频会议、可视电话等低延时实时通信领域;支持I条带和P条带,熵编码支持CAVLC算法。
- 主要档次(Main Profile):主要用于数字电视广播、数字视频数据存储等;支持视频场编码、B条带双向预测和加权预测,熵编码支持CAVLC和CABAC算法。

● 扩展档次(Extended Profile):主要用于网络视频直播与点播等;支持基准档次的所有特性,并支持SI和SP条带,支持数据分割以改进误码性能,支持B条带和加权预测,但不支持CABAC和场编码。

三. H.264标准采用的编码工具

在H.264中使用的编码技术主要有以下类型:

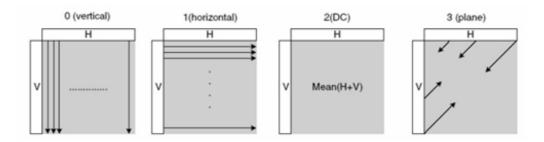
帧内预测

H.264中采用了基于像素块的帧内预测技术。主要可分为以下不同类型:

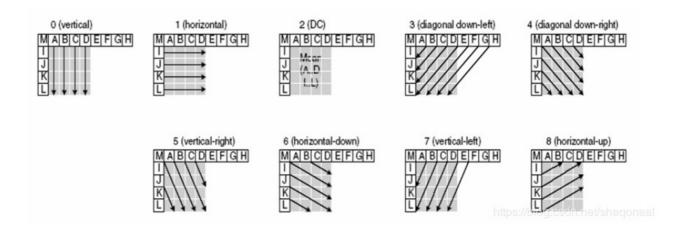
16×16大小的亮度块:4种预测模式4×4大小的亮度块:9种预测模式

● 色度块:4种预测模式,同16×16的亮度块

16×16亮度块和色度块的4种预测模式如下图:



4×4亮度块的9种预测模式如下图表示:



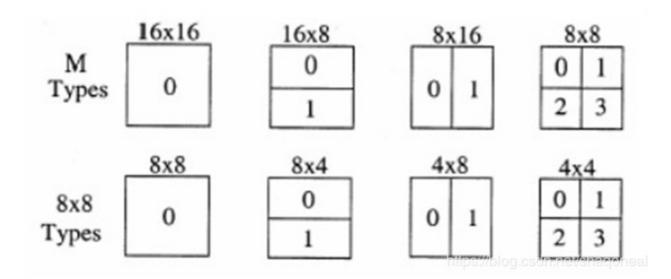
帧间预测

H.264中的帧间预测方法采用了基于块的运动估计和补偿方法,其主要特点有:

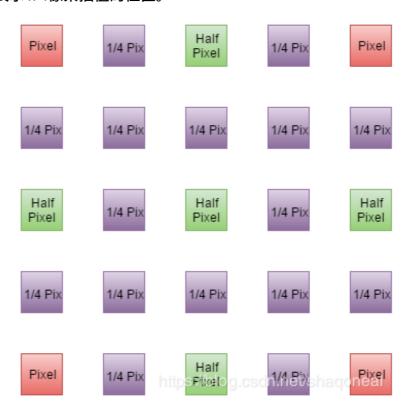
- 多个候选参考帧;
- B帧作为参考帧;
- 任意参考帧排序;

- 多种运动补偿像素块形状,包括16×16、16×8、8×16、8×8、8×4、4×8和4×4像素;
- 1/4像素(亮度)的亚像素插值;
- 对交错视频的基于帧或场的运动估计;

进行帧间预测的宏块进行子宏块分割的方式如图所示:



亚像素插值的表示如下图。其中红色点表示图像中的整像素点的位置,绿色点表示两个正像素之间的1/2像素插值的位置,紫色点表示1/4像素插值的位置。



交错视频编码

针对隔行扫描的视频,H.264专门定义了用于处理此类交错视频的算法。

- PicAFF: Picture Adaptive Frame Field——图像层的帧场自适应;
- MBAFF:MacroBlock Adaptive Frame Field——宏块层的帧场自适应;

变换和量化编码

H.264的变换编码创新性地采用了类DCT的整数变换,有效降低了运算复杂度。对于基础版 H.264,变换矩阵为4×4;在FRExt扩展中,还支持8×8的变换矩阵。

H.264的量化算法仍然采用标量量化方法。

无损熵编码算法

H.264标准针对不同的语法元素指定了不同的熵编码算法,主要有:

- UVLC(Universal Variable Length Coding):主要采用指数哥伦布编码;
- CAVLC(Context Adaptive Variable Length Coding):上下文自适应的变长编码;
- CABAC(Context Adaptive Binary Arithmetic Coding):上下文自适应的二进制算数编码;

其他技术

除了上述的核心算法之外,H.264还定义了包括去块环路滤波器、SI/SP帧、码率控制等多种技术。