



## 交流图片




图像和视频编码课程

2014 年, 第 411-449 页

## 第12章-视频编码标准

大卫·R·布尔

显示更多 ∨

 大纲 |  共享  引用<https://doi.org/10.1016/B978-0-12-405906-1.00012-X>

获取权利和内容

## 摘要

视频格式和**压缩方法**的标准化对视频技术的普遍采用产生了重大影响, 本章对此进行了讨论。它回顾了H.261、H.263和MPEG-2等旧标准的特征, 然后在此基础上全面描述了H.264/AVC和HEVC这两个最新标准的主要特征。它提供了这些标准之间的性能比较。



## 关键词

视频压缩标准; MPEG-2; H.263; MPEG-4; H.264/AVCH; .265/HEVCDiracVP; 9; VC-1

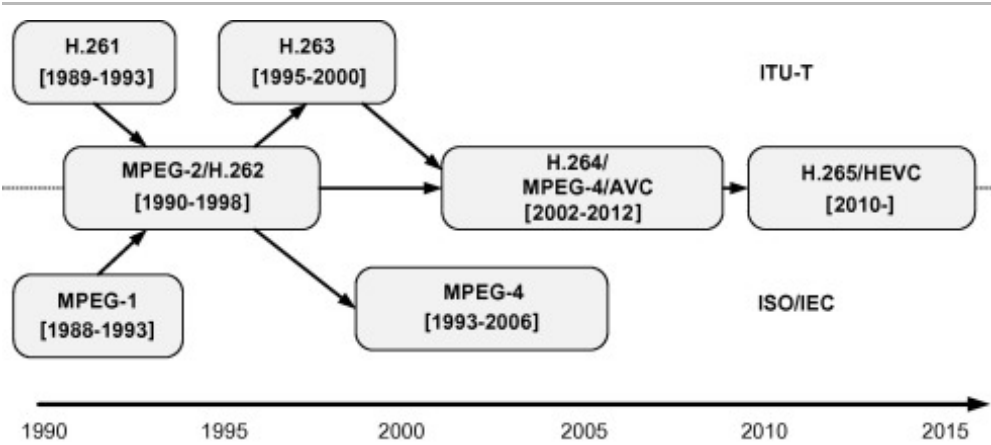
视频格式和**压缩方法**的标准化过程对视频技术的普遍采用产生了重大影响。标准对于互操作性至关重要, 允许来自不同来源的材料通过广泛的网络进行处理和传输, 或存储在广泛的设备上。这种互操作性为用户提供了尽可能广泛的服务。它还降低了制造商的风险, 刺激了研发投入, 并创造了巨大的视频设备市场, 具有批量制造的优势。

从1984年引入H.120到最新的H.265/HEVC编解码器, 视频编码标准每10年将同等视频质量的比特率减半。本章概述了促成这一进展的特点。它无意成为现代编码标准结构和实现的权威指南,

而是使读者能够理解前几章中描述的架构、方法和算法如何在当今常用的编解码器中使用。

### 12.1。标准的必要性和作用

视频编码标准的年表见图12.1（为方便起见，从第一章中重复）。它们的主要特征在以下章节中描述。然而，让我们首先重新审视为什么我们需要标准，并考虑过去30年左右导致一系列成功的编码标准的过程。



[下载：下载全尺寸图像](#)

图12.1。1990年至今的视频编码标准年表。

#### 12.1.1。视频标准化的重点

为了使视频编码标准（或任何其他标准）取得成功，它必须明确满足需求。特别是：它必须优于任何以前的标准，它不得扼杀产品生产者之间的竞争（即它必须允许创新），当然，它必须通过独立于特定的通信网络或存储设备来提供互操作性。在大多数情况下，如果标准向后兼容，它应该是有利的——即它应该能够从以前的标准解码比特流。标准化过程还必须铭记，未来的标准需要向后兼容。

视频编码标准通常定义比特流格式和语法以及解码过程，而不是（大部分）编码过程。图12.2说明了这一点，其中虚线框指示了标准的规范方面。因此，符合标准的编码器是产生兼容的码流的编码器，符合标准的解码器是可以解码兼容的码流的解码器。在这方面，必须强调这样一个事实，即编码器的标准合规性并不能保证质量；制造商仍然面临的挑战是通过创新的低复杂性和高效的编码解决方案来区分其产品。

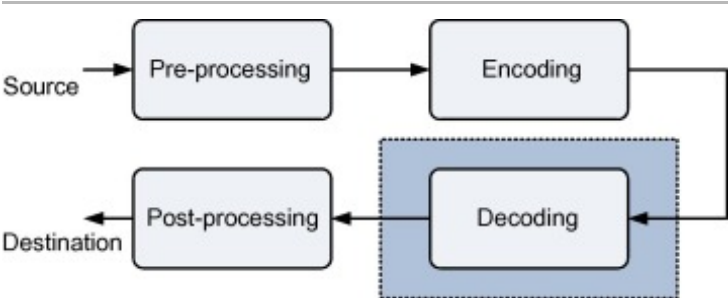


图12.2。标准化的范围。

---

### 12.1.2。标准化过程

早期的视频编码标准由国际电信联盟（ITU）（前CCITT）或国际标准组织（ISO）独立制作，特别是ISO/IEC（国际电工委员会）联合技术委员会1（JTC 1），该委员会处理所有计算机相关活动，包括视频压缩。然而，这两个组织之间的联合方法使最近的标准受益匪浅。

标准化过程[1]遵循明确界定的路径，分为以下阶段：

**需求定义：**这里敲定了标准的范围，定义了标准化过程的要求和目标。

**分歧：**在这一阶段，引入竞争，最初是通过确定关键技术领域（KTA）。然后通常会发出正式的招标书，使来自行业和学术界的专家能够介绍和比较他们的方法和结果，通常与[预先定义的测试数据集进行比较](#)。

**收敛：**此阶段的目的是通过选择在发散阶段贡献的最佳算法，为标准创建一个解决方案。这通常通过使用不断发展的测试模型来实现，例如H.264/AVC中的JM（联合模型）或HEVC测试模型（HM）。该模型通过多次迭代演变，用作后续提案的参考。如果成功，这些将被纳入模型的新改进版本。这个过程一直持续到满足所需的性能规格。

**验证：**在验证过程中，对结果标准及其比特流进行一致性验证。然后为产品的合规性测试定义一致性测试流程。

### 12.1.3。知识产权和许可

当按照国际标准大规模投资生产产品时，制造商和用户正确地期望对专利侵权诉讼有一定的保护。对于大多数标准化编解码器，供应商和用户必须向相关知识产权的所有者支付版税。这主要由一个名为MPEG许可管理局的美国组织（与MPEG无关）处理。MPEG LA管理与MPEG-2、MPEG-4、VC-1和H.264/AVC的专利池相关的专利许可。它还在为HEVC的许可汇集专利。

2005年至2007年间，高通和博通之间关于侵犯H.264专利的争议提交美国法院。相关专利被判定不可执行，因为它们在2003年H.264标准化之前没有向MPEG JVT披露。这种类型的裁决在一定程度上有助于人们对为视频标准化而采用的许可程序的稳健性的信心。

## 12.2。H.120

### 12.2.1.简史

CCITTSG.XV研究小组于1980年开始H.120的工作，并于1984年制定了第一个国际数字视频编码标准，随后于1988年发布了第二版[2]。H.120分别以2.048 Mb/s和1.544 Mb/s处理625/50和525/60电视系统的视频会议应用程序。该标准从未在商业上取得成功，部分原因是它基于不同国际地区

的不同编码策略，但主要是因为它的图片质量（特别是时间质量）不足。

12.2.2. 主要特征

H.120实施了一项有条件的充资战略，根据该战略，每个框架被划分为变化和不变的区域。更改的区域在标准第1部分和第2部分中使用场内DPCM编码，尽管第3部分（在美国使用）使用背景预测和运动补偿场间预测。

 [下载](#)

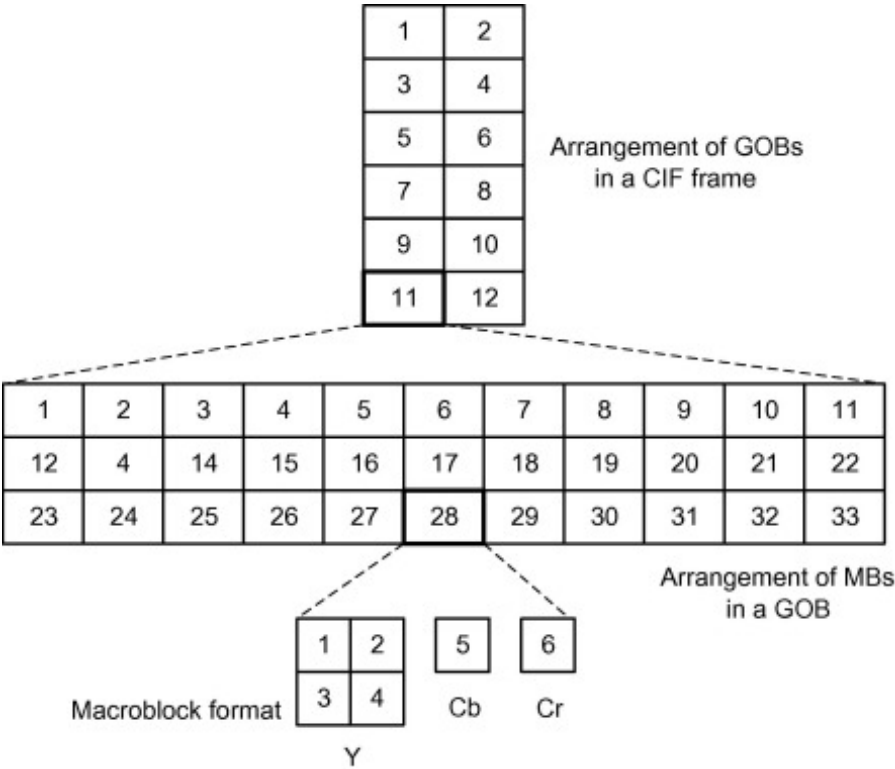
12.3.1. 简史

1989年，H.261 [3]紧随H.120之后，是第一个实现广泛产品采用的视频编解码器。它基于一个  $p \times 64 \text{ kbps}$  ( $p = 1, \dots, 30$ )模型，针对ISDN会议应用程序。H.261是第一个使用变换（离散余弦变换（DCT））、时间DPCM和运动补偿组合的基于块的混合压缩算法。这种架构经受住了时间的考验，因为此后所有主要视频编码标准都基于它。

12.3.2. 图片类型和主要特征

宏块、GOB和帧格式

H.261在图片编码过程中引入了一个基本的层次结构，经过一些修改，今天仍在使用。H.261图片以CIF或QCIF格式表示为一组块（GOB），每组由一些宏块（MB）组成。如第4章所述，宏块由四个空间组成  $8 \times 8$  DCT变换的流形块和2个（子采样）空间  $16 \times 16$  DCT变换色度块。图12.3给出了CIF框架的图片层次结构图。可以看出，GOB中有33MB，给出GOB维度  $176 \times 48$  像素，CIF帧中有12个GOB，给出帧尺寸  $352 \times 288$  像素。



[下载：下载全尺寸图像](#)

图12.3。H.261宏块、GOB和CIF帧格式。

H.261只支持I和P帧——不支持B帧。运动估计和补偿是编码过程不可分割的一部分，其固定的块大小为 $16 \times 16$ 和 $[-15 \dots + 15]$ 像素。

### 编码器控制

H.261在允许在帧间和帧内模式之间切换以及在宏块级别控制量化器步长方面提供了一些灵活性。集成了高级功能，如空间回路滤波器，用于消除高频噪声（系数为0.25、0.5和0.25的三点滤波器）。这允许通过平滑参考和输出帧中的像素来减少预测误差。为了恢复错误，还加入了可选的BCH错误检测校正（511,493）方案。H.261的主要特征汇总于表12.1。

表12.1。H.261的主要特征。

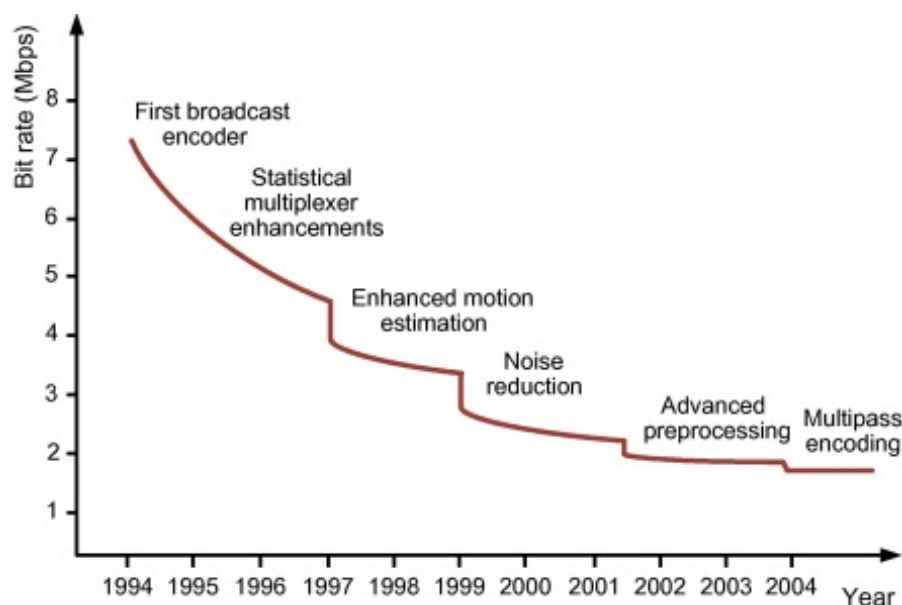
编解码器功能	方法
支持的格式	CIF, QCIF
图片格式	四层（图片、GOB、MB和块）
颜色子采样	4:2:0 YCbCr
框架类型	我, P
编码内转换	阻止DCT
编码间转换	阻止DCT
熵编码	VLC和霍夫曼编码
量化器	制服（DC）和死区（AC）
动议赔偿	选修科目 $[-15 \dots + 15]$
编码控制	间/内和量化器步长的选择
循环过滤器	三触式空间过滤器
错误保护	可选BCH（511, 493）编码

## 12.4。MPEG-2/DVB

### 12.4.1。简史

1988年，[移动图片专家组](#)（MPEG）成立，于1992年提供一种视频编码算法（MPEG-1），旨在以1.5 Mbps/s的数字存储媒体。随后在1994年推出了MPEG-2[4]，专门针对新兴的数字视频广播市场。十多年来，MPEG-2被纳入所有机顶盒，在真正[支持](#)数字广播革命方面发挥了重要作用。

由于它专注于广播，编码器和解码器的复杂性之间存在公认的不平衡，这导致了对高成本、高复杂性的编码技术进行大量投资，以生产高性能演播室和基于头端的编码器。[图12.4](#)展示了MPEG-2专业编码器在十年的发展中如何提高性能的例子，以及导致该性能一系列步骤变化的创新。



[下载：下载全尺寸图像](#)

图12.4。MPEG-2编码器在1994年至2004年间性能有所改善。

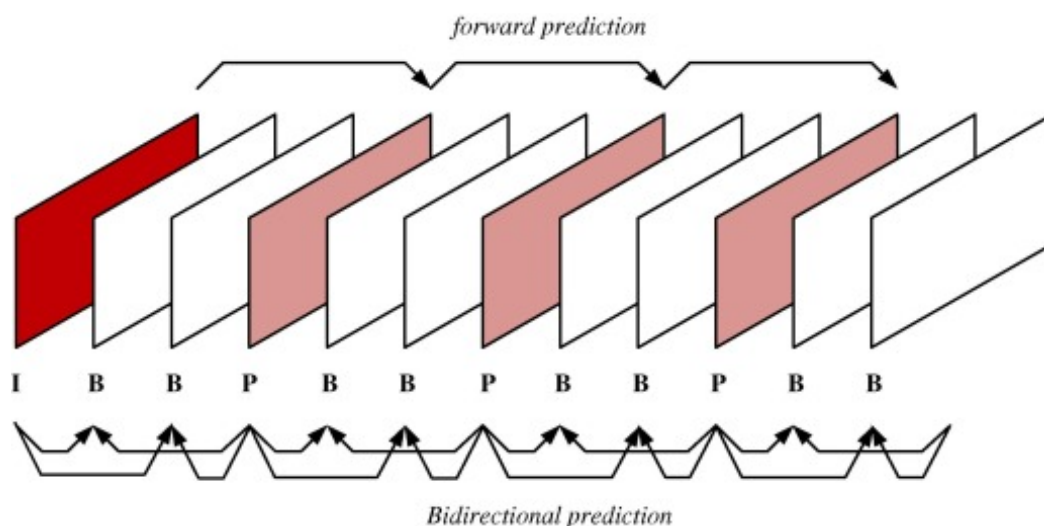
改编自坦德伯格的原版。

### 12.4.2。图片类型和主要特征

MPEG-2 (ISO/IEC-13818:2000) [4]，也称为H.262，是一个通用的视听编码标准，支持从2到30 Mbps的比特率提供一系列应用程序。该标准由四个主要部分组成：13818-1：系统，-2：视频，-3：音频，-4：一致性。它采用混合运动补偿块的DCT架构（类似于H.261），使用 $8 \times 8$ 块和 $16 \times 16$ 基于平移块的运动估计为半像素精度的[宏块](#)。由于MPEG-2专注于数字电视广播，因此必须支持渐进式和交错式图像格式。编码标准主要针对无错误环境定义，并根据应用场景（例如地面、电缆或卫星广播）添加信道编码。MPEG-2被定义为支持从工作室处理（捕获和编辑）到分发和广播交付的广泛应用程序。它与DVB-T、DVB-S和DVB-C广播内容交付标准紧密耦合。

MPEG-2支持三种图片类型：Intra (I)——不参考其他帧（效率最低）、预测 (P) ——基于之前I或P帧的预测编码，以及双向预测 (B) ——从P和/或I帧预测，但不用作进一步预测的基础（效率最高）。它使用[图12.5](#)所示的共和国结构，是第一个引入双向预测B帧的标准。





[下载：下载全尺寸图像](#)

图12.5。MPEG-2共和党结构。

### 12.4.3。MPEG-2配置文件和级别

为了将性能与解码器的能力或容量相匹配，MPEG-2引入了广泛的配置文件和级别。配置文件是整个位流语法的定义子集，配置文件被进一步划分为级别。每个级别为比特流中的参数指定允许值的范围。

MPEG-2支持六个配置文件——简单、主配置文件、信噪比配置文件、空间配置文件、高4:2:2配置文件和多视图配置文件。SNR、空间和高配置文件中包含可伸缩性条款，而简单配置文件和主配置文件仅允许单层编码。它还在每个配置文件中提供了四个可能的级别——低级、主级、高1440级和高级。主电平的参数大致对应于正常电视分辨率，低电平对应于CIF分辨率，High1440和High的值对应于HDTV分辨率。在实践中只使用过两个配置文件：工作室工作和后期制作的4:2:2配置文件和广播电视交付的主要配置文件。表12.2提供了大多数交付应用程序中使用的MPEG-2主配置文件（MP@ML）的规格。

表12.2。MPEG-2 main profile@main level (MP@ML) 边界。

参数	绑定
样品/线路	720
线条/框架	576
帧/秒	30
样本/秒	10,368,000

比特率	15 Mb/s
缓冲器尺寸	1,835,008位
色度格式	4:2:0
宽比	4:3, 16:9平方像素

## 12.5。H.263

### 12.5.1。简史

H.263[5]由ITU-T SG15定义，从1993年开始，目标是以低于64kbps的比特率编码。最初的应用程序重点是PSTN和早期移动无线电应用程序，比特率为10至24kb/s。尽管移动视频起飞速度比预期的要慢，但H.263在会议和监控应用程序以及早期互联网流媒体方面产生了重大影响。特别是，H.263用于为YouTube和MySpace等网站编码Flash Video内容。RealVideo最初基于它，并在几个ETSI 3GPP视频服务中指定了它。该编解码器最初被设计为集成到用于电路交换应用程序的H.324框架中，但已广泛用于H.323（基于RTP的视频会议）和其他IP流包装。

### 12.5.2。图片类型和主要特征

#### 宏块、GOB和帧格式

H.263支持以下图片格式：子QCIF（ $88 \times 72$ ），QCIF（ $176 \times 144$ ），到岸价（ $352 \times 288$ ），4CIF（ $704 \times 576$ ），16CIF（ $1408 \times 1152$ ）全部采用YCbCr 4:2:0格式。与H.261一样，每张图片被划分为一组块（GOB），定义为整数， $k$ ，宏块（MB）行。 $k = 1$ 对于SQCIF，QCIF和CIF， $k = 2$ 对于4CIF，以及 $k = 4$ 为16CIF。H.263 QCIF图片的结构如图12.6所示。

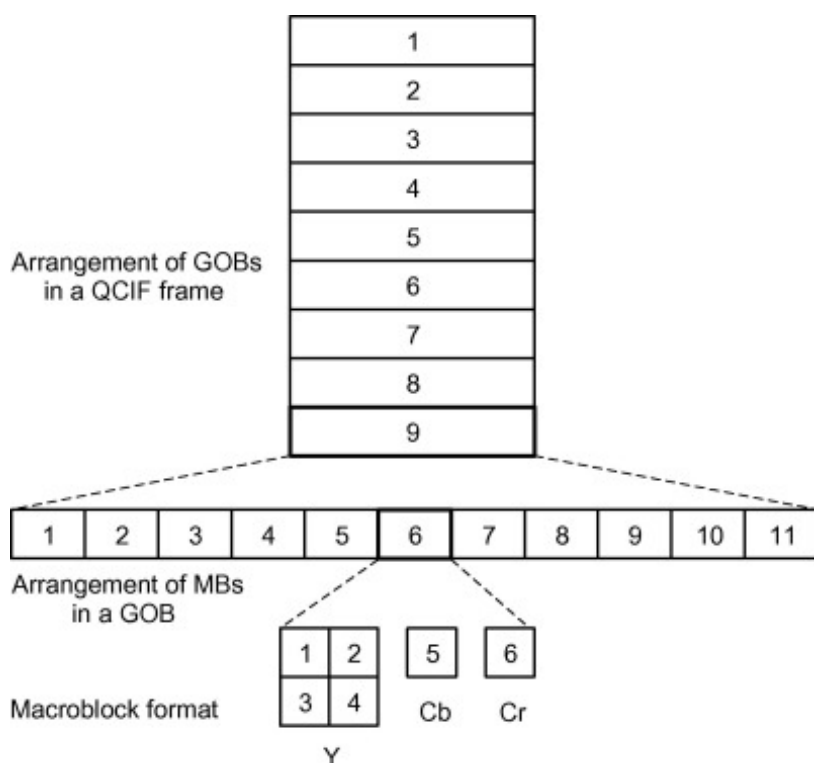




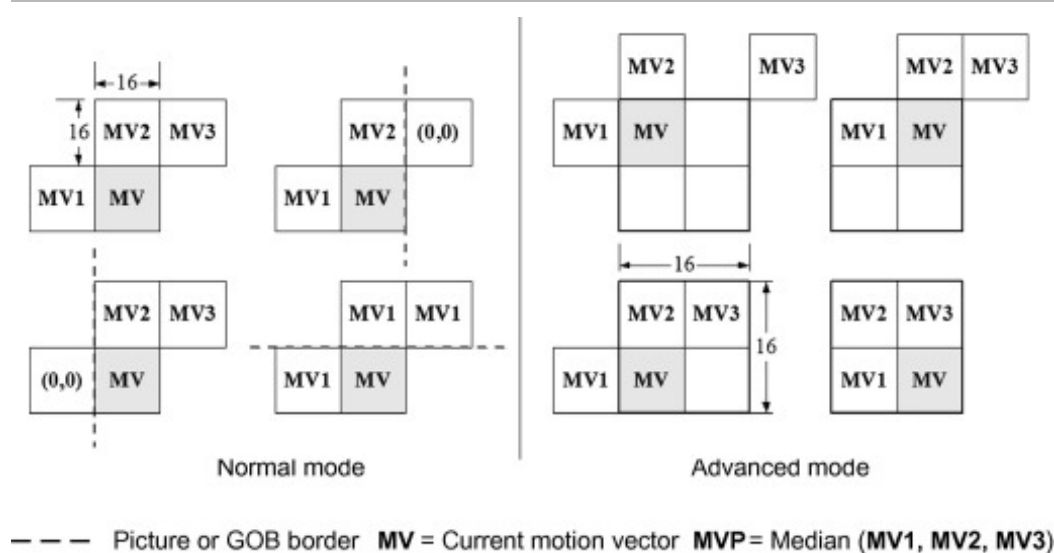
图12.6。H.263图片格式（QCIF）。

H.263的主要特征

尽管H.263使用与H.261和MPEG-2相似的编码架构，但H.263通过集成一套全面的高级功能实现了显著的性能提升。这些汇总在表12.3中。该标准基于混合运动补偿DCT，使用具有量化步长之字形扫描 ( $QP = 1 \dots 31$ )每个MB都可以更改。H.263中使用三维VLC表，其中包含事件{RUN, LEVEL, LAST}，其中LAST信号是块中最后一个非零系数。这避免了对显式EOB符号的需求，并提供了 $8 \times 8$ DCT块。H.263使用半像素精度的运动补偿，运动矢量可以预测编码，如图12.7所示。

表12.3。H.263的主要特征。

编解码器功能	方法
支持的格式	SQCIF, QCIF, CIF, 4CIF, 16CIF
图片格式	四层（图片、GOB、MB和块）
颜色子采样	4:2:0 YCbCr
框架类型	I, P, B, （PB）
编码内转换	块 $8 \times 8$ DCT
编码间转换	块 $8 \times 8$ DCT
量化器	制服（DC）和死区（AC） $QP = 1 \dots 31$
动议赔偿	半像素，无限制向量，重叠。
熵编码	三维VLC、霍夫曼和算术编码
循环过滤器	无
编码控制	内部/内部，量化器步长，以及MB级别的传输/跳过
错误保护	由运输层或BCH-FEC提供



下载：下载全尺寸图像

图12.7。H.263中的运动矢量编码。

就PSNR而言，H.263比H.261最多可以提供3-4分贝的改进。它这样做是因为其半像素运动预测，但也通过包含四种新的可选编码模式：

- **不受限制的运动矢量：**允许运动矢量指向参考图片区域之外，搜索范围扩展到 $[-31.5, +31.5]$ 。
- **基于语法的算术编码：**传统的霍夫曼熵编码器被算术编码器取代，使每个编码符号的字长能够达到分数。
- **高级预测模式：**这包括两个子模式：
  - 每块使用四个MV，即每个块一个 $8 \times 8$ 挡，而不是每个挡一个 $16 \times 16$ 阻止。这使编码器能够更好地处理块内的多个运动。
  - 过量运动补偿（OMC）：在这里，每个像素在 $8 \times 8$ 亮度块被预测为三个预测值的加权之和。这些预测值对应于当前块的向量，以及从块到当前块的顶部、底部、左侧和右侧的四个预测中的两个。选择的两个块是那些最接近计算像素的块——例如，块左上角象限中的像素使用上部和左上的预测。图12.7（右）说明了这一点。
- **PB帧模式：**在PB帧中，两张图片被编码为一个单元——一张P帧从以前的P帧预测，一张B帧，使用双向预测从相邻的P帧中预测。PB帧宏块由12个块组成——预测的P块有6个，B块有6个。PB-帧的好处是，运动矢量不为B-块传输，而是使用P-块的缩放版本来导出，该矢量基于相对于前一个P帧中相应P-块的局部时间活动。这个过程可以在不显著增加比特率的情况下有效提高帧率。

## H.263扩展（H.263+和H.263++）

H.263版本2，也称为H.263+，于1998年标准化，并通过许多新模式和功能扩展了H.263，进一步

提高了**压缩效率**。H.263++（第3版）在2000年提供了进一步的增强。H.263+中的12种新模式提高了编码增益，提高了错误恢复力，启用可扩展的**比特流**，在图片大小和时钟频率方面引入了灵活性，并提供补充显示能力。其中许多新模式被纳入后来的H.264/AVC标准。

首先，介绍了几种支持错误恢复力的新模式，并总结如下。我们已经**在第11章**中涵盖了其中许多，温格等人也对此进行了审查。[\[6\]](#)。

- **切片结构模式（附件K）**：在此模式下，GOB结构被一组切片取代。由于不允许跨切片边界的预测依赖关系，因此一个切片中的所有宏块都可以独立解码。
- **独立分段解码模式（附件R）**：段边界的作用方式与图片边界相同。段可以是切片、GOB或GOB的集合，并且段的形状必须因帧而相同。随着图像定义部分之间的错误传播被消除，独立片段支持错误恢复力。它们还以类似于MPEG-4中的物体平面的方式启用特效。
- **参考图片选择模式（附件N）**：如**第11章**所述，这允许灵活选择参考图片。也可以将参考图片选择模式应用于单个片段，而不是将完整图片应用于。如果存在反馈通道，此模式支持错误恢复，但也是H.264和HEVC中使用的多个参考框架方法的前身。
- **时间、信噪比和空间可伸缩模式（附件O）**：该模式在H.263标准中引入了分层，以支持向不同功能的终端交付和**拥塞管理**的灵活性。如**第11章**所述，可伸缩的比特流包括基层和增强层，其中基层提供可接受的质量水平，如果可用，其他层可以进一步增强（在改进**信噪比**、改进的时间分辨率或改进空间分辨率方面）。

其他支持增强编码增益的H.263模式有：

- 高级编码内模式（附件一）。
- 修改的无限制运动矢量模式（附件D）。
- 改进的PB框架模式（附件M）。
- 解封过滤模式（附件J）。
- 参考图片重新采样模式（附件P）。
- 分辨率更新模式降低（附件Q）。
- 替代VLC模式（附件S）。
- 修改**量化**模式（附件T）。

最后引入了一种模式来提供额外的比特流信息：

- **补充增强信息模式（附件L）**：这允许包含特定解码器可以解码或不可解码的附加信息。它可以包含色度键控、时间段、图片冻结和图片大小调整等信息。

H.263++为H.263提供了进一步的增强，引入了与MPEG-2相似的9个配置文件和7个级别。配置文件使特定模式组合的形式化，以支持不同的应用程序场景（例如基线和无线）。这些级别使编码器性能与环境约束（例如最大比特率、空间分辨率等）相匹配。它还引入了一些新模式并修订了

其他模式——数据分区切片模式（附件五）、附加增强信息（附件W）和增强参考图片选择模式（附件U）。

## 12.6。MPEG-4

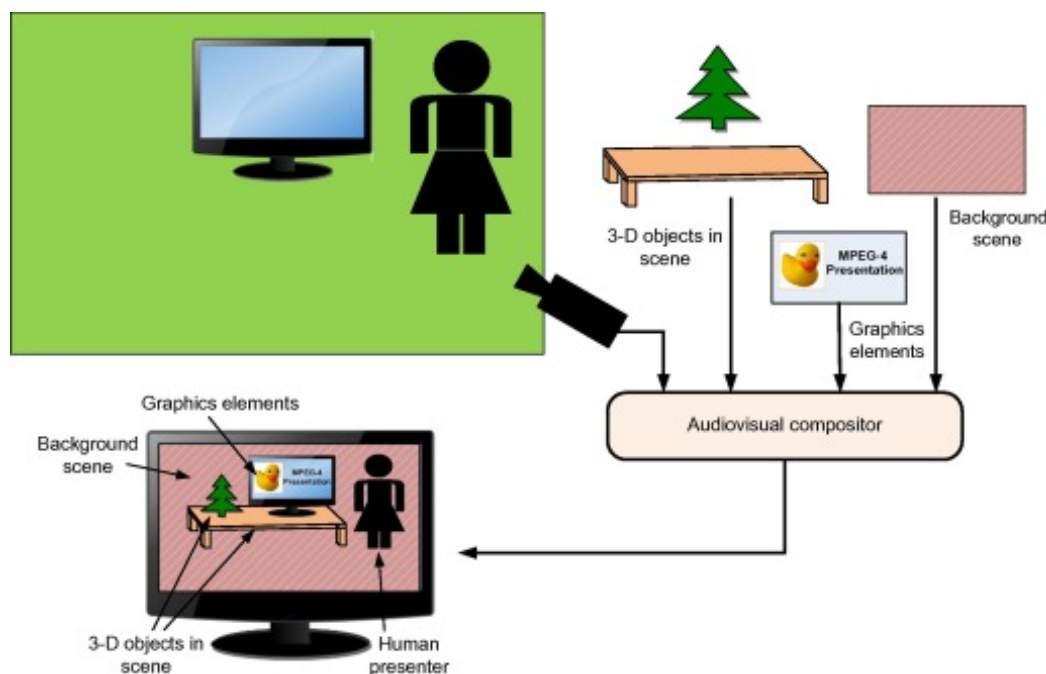
### 12.6.1。简史

MPEG-4[7]是一个非常雄心勃勃的项目，旨在引入使用基于对象和波形的方法的新方法。MPEG-4标准于2000年批准，名为ISO/IEC 14496，可能为MPEG-2提供重大编码改进。然而，总的来说，它太复杂了。比较H.261和MPEG-4的标准文件很有趣。前者长达25页，而MPEG-4仅视觉部分就有500页。在实践中只使用了高级简单配置文件（ASP），这获得了一些牵引力，为当时新兴的数码相机技术奠定了基础。早期的MPEG-4第2部分比MPEG-2或H.263几乎没有实际优势，并迅速被MPEG-4第10部分/AVC（现在称为H.264/AVC）超越。

### 12.6.2。图片类型和主要特征

MPEG-4旨在以比MPEG-2更低的比特率和更高的视频质量对视频进行编码，有效地创建多媒体传输和存储框架。MPEG-4（第2部分）旨在适用于从工作室和电影到手机视频应用程序等各种视频编码场景。在这种情况下，它专注于基于对象和基于波形的编码。

MPEG-4提供了一套视听场景的编码工具，支持任意形状物体的编码，如图12.8所示。这个图中的场景代表二维背景、屏幕上播放的视频、演示者以及桌子和树等三维对象。MPEG-4旨在独立处理此类视听对象及其合成，并允许修改自然和合成（计算机生成）内容。



[下载：下载全尺寸图像](#)

图12.8。MPEG-4视听场景的一个例子。

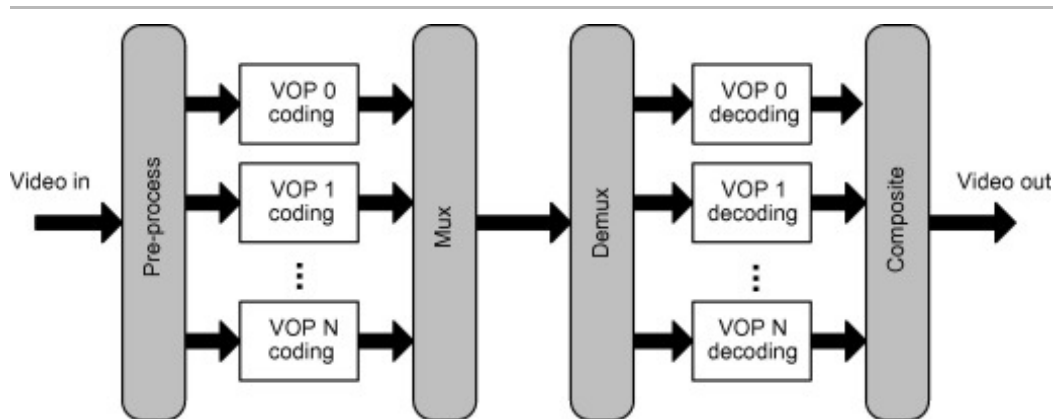
与MPEG-2和H.263++一样，MPEG-4最初使用简单配置文件、核心配置文件和主配置文件定义了一些一致性点。简单和核心配置文件将QCIF和CIF的场景大小为64、128、384kbps和2 Mbps。主配置文件针对CIF、ITU-R 601和HD尺寸，比特率为2、15和38.4 Mbps。该标准还包含对精灵、可伸缩性和错误恢复力的支持。

MPEG-4还首次支持与MPEG-7内容描述方案集成，以促进元数据的使用。易卜拉欣和霍恩[8]提供了MPEG-4自然视频编码的优秀概述。

## 编码框架

虽然我们不会花很多时间去描述MPEG-4的细节，但值得简单地将其目标置于上下文中。

MPEG-4支持将视频内容组合或分解为视觉对象，允许这些内容作为多个视频对象平面（VOP）独立处理（图12.9）。为了描述这些物体平面，有必要对其纹理、形状和运动进行编码。形状编码使用二元掩码或灰度阿尔法通道完成，允许透明度。运动补偿和基于DCT的纹理编码都必须考虑对象边界。错误恢复力由同步标记、数据分区、标头扩展代码和可逆可变长度代码提供。为空间和时间分辨率的增强提供了可伸缩性。MPEG-4还提供了基于对象的可伸缩性，但限制对象形状必须是矩形。



[下载：下载全尺寸图像](#)

图12.9。MPEG-4视频对象平面（VOP）编码的通用表示。

## MPEG-4第2部分高级简单配置文件（ASP）

高级简单配置文件是最常用的MPEG-4配置文件，支持隔行视频、B图、四分之一像素运动补偿和全局运动补偿。随后，H.264/AVC采用了四分之一像素精度，而大多数实现通常不支持全局运动补偿功能。读者会注意到，ASP没有提供很多H.263中没有的东西；事实上，H.263的吸收比ASP广泛得多。

在引入ASP后不久，MPEG-4视觉工作重新聚焦于第10部分，这成为与ITU-T在H.264/AVC旗帜下的一项联合活动。向前迈出的这一重大一步是下一节的主题。

## 12.7。H.264/AVC

### 12.7.1。简史

国际电联-T/SG16/Q6（视频编码专家组）于1998年开始开展一个名为H.26L的项目。VCEG和MPEG（移动图片专家组）于2001年成立了JVT（联合视频团队），旨在建立一个名为H.264/MPEG-4-AVC的联合标准项目。最终草案于2003年3月完成，供正式批准[9]，[可伸缩视频编码](#)（SVC）扩展于2007年完成。

MPEG-4（第10部分）或H.264/AVC是迄今为止最普遍的视频编码标准。这种情况将持续到2015年，届时批量生产和基础设施的改变使H.265/HEVC得以重大转变。与MPEG-2支持数字广播革命一样，H.264/AVC在启用互联网视频、移动服务、OTT服务、IPTV和HDTV方面发挥了关键作用。H.264/AVC是蓝光播放器的必备格式，大多数互联网流媒体网站都使用，包括Vimeo、YouTube和iTunes。它用于Adobe Flash Player和Microsoft Silverlight，并已用于HDTV电缆、卫星和地面广播。

国际电联-T与ISO/IEC合作，于2004年交付了H.264/AVC。我们已经研究了前几章中有助于H.264成功和性能的许多功能，我们总结如下。

### 12.7.2。主要特征

H.264/AVC项目的目的是提供：

- **提高编码效率：**与任何其他视频标准相比，在固定保真的情况下，目标[平均比特率](#)降低50%。
- **错误健壮性：**处理数据包丢失和拥塞管理的工具。
- **网络友好性：**主要目标是移动网络和互联网——通过网络抽象将编码层分离。
- **适应延迟约束：**提供低延迟模式。
- **简单语法规则：**避免过多的可选功能或配置文件配置。

这些都是明智的目标，但最后一个目标很有趣，因为它承认其前身的过度复杂性。

现在让我们更详细地考虑其中一些属性。[表12.4](#)提供了摘要；虽然H.264等复杂编解码器的所有细节都无法在此类表中捕获，但可以看出，许多功能与H.263扩展中的功能相似。有关更多详细信息，请参阅标准文件[9]，或Wiegand等人的概述。[\[10\]](#)。理查森[\[11\]](#)提供了该标准的非常易于理解的介绍。

表12.4。H.264/AVC的主要特征。

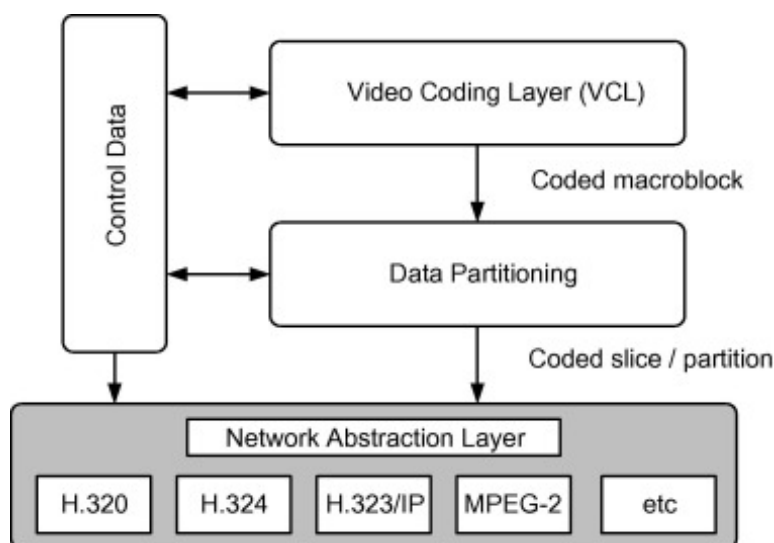
编解码器功能	方法
--------	----



支持的格式	QCIF到UHDTV (4k)
图片格式	四层 (图片、GOB、MB和块)
颜色子采样	4:0:0, 4:2:0, 4:2:2, 4:4:4 YCbCr
分区	灵活的切片结构
框架类型	I, P, B, SP, SI
编码内转换	块 $8 \times 8$ 或者 $4 \times 4$ 整数变换
编码间转换	块 $8 \times 8$ 或者 $4 \times 4$ 整数变换
预测内部	多方向, 多模式
量化器	标量具有对数控制, $QP = 0 \dots 51$
动议赔偿	四分之一像素, 无限制向量, 多个参考帧
ME块大小	$16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8, 8 \times 4, 4 \times 8, 4 \times 4$
熵编码	CAVLC或CABAC
编码控制	多种复杂的编码控制模式——RDO
循环过滤器	环内非线性解块过滤器
错误恢复力	切片、FMO、SI、SP帧、多个参考帧

### 12.7.3. 网络抽象和码流语法

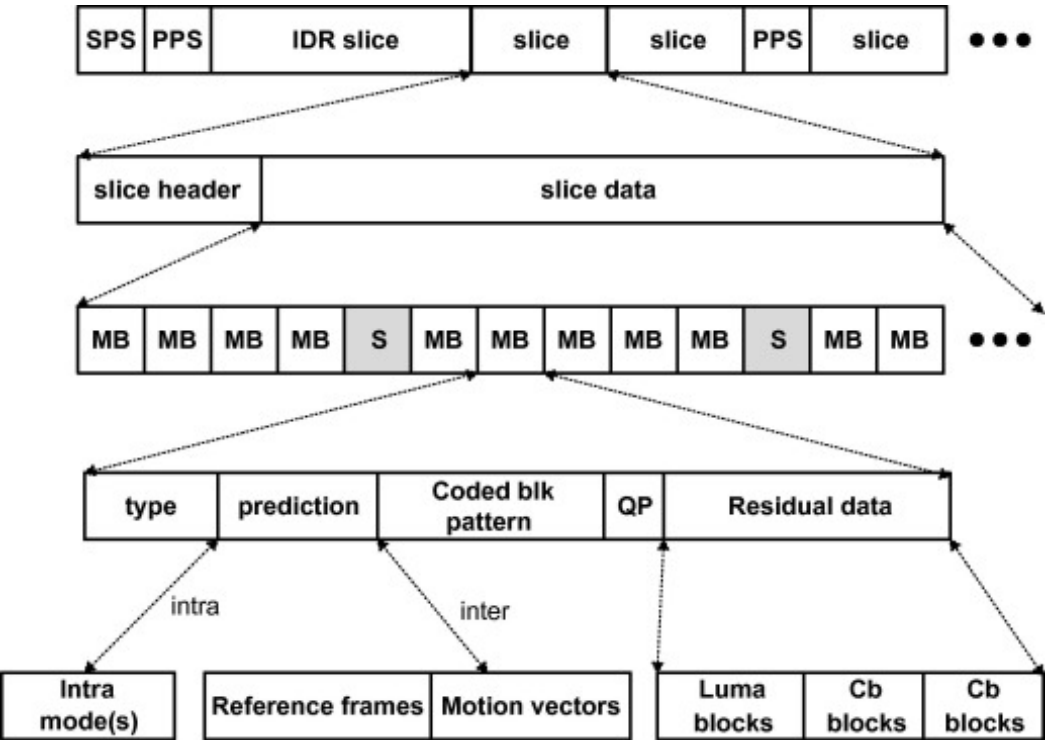
图12.10显示了H.264/AVC编码的分层结构。这确保了编码层独立于网络，依赖于如图所示的协议，以提供映射到适当的视频传输机制。因此，VCL数据可以很容易地映射到传输层，例如用于实时有线或无线互联网的RTP/IP，用于存储和彩信的MP4，用于有线和无线会话服务的H.32X，或用于广播的MPEG-2系统。



[下载: 下载全尺寸图像](#)

图12.10。H.264/AVC层结构。

在H.264/AVC中，[比特流](#)是分层组织的，包括网络抽象层（NAL）和视频编码层（VCL）。编码时，H.264信息表示为一系列NAL单元（NALU）。NALU标头指示NALU的类型，可以是序列参数集（SPS）、图片参数集（PPS）、VCL编码切片数据或补充增强信息（SEI）。SPS包含有关整个序列的信息，如配置文件、电平、帧大小以及对解码器重要的其他特征。PPS包含与帧子集相关的更多本地化信息，例如切片组的数量、熵编码模式、是否使用加权预测和其他初始化参数。SEI涉及对解码不重要但可能有助于缓冲区管理或其他非规范任务的信息。显示分层H.264语法的图表如[图12.11所示](#)。



[下载](#)： [下载全尺寸图像](#)

图12.11。H.264语法。

每个序列都以瞬时解码器刷新（IDR）访问单元开始。这是一个内部编码的图片，它向解码器表明，后续的图片不需要引用它之前的图片。视频切片数据对应于一系列编码[宏块](#)，每个块都包含与解码相关的信息以及实际编码剩余数据。以下是参考[图12.11](#)的MB级别信息：

- **类型**：I, P, B。
- **预测**：I-block模式 ( $16 \times 16, 8 \times 8$  or  $4 \times 4$ )，P块分区 ( $16 \times 16, 16 \times 8, 8 \times 16$  or  $8 \times 8$  (包含子块分区)、B块分区 (如P块)、参考帧的选择和[运动矢量](#)数据。注意：P块和B块可能指示SKIP条件，以表明没有发送剩余数据。

- **编码块模式**：指示存在哪些块系数。
- **量化参数**：MB的QP值。
- **剩余数据**：由美国海关及边境保护局参数指示的编码剩余数据。

#### 12.7.4. 图片和分区

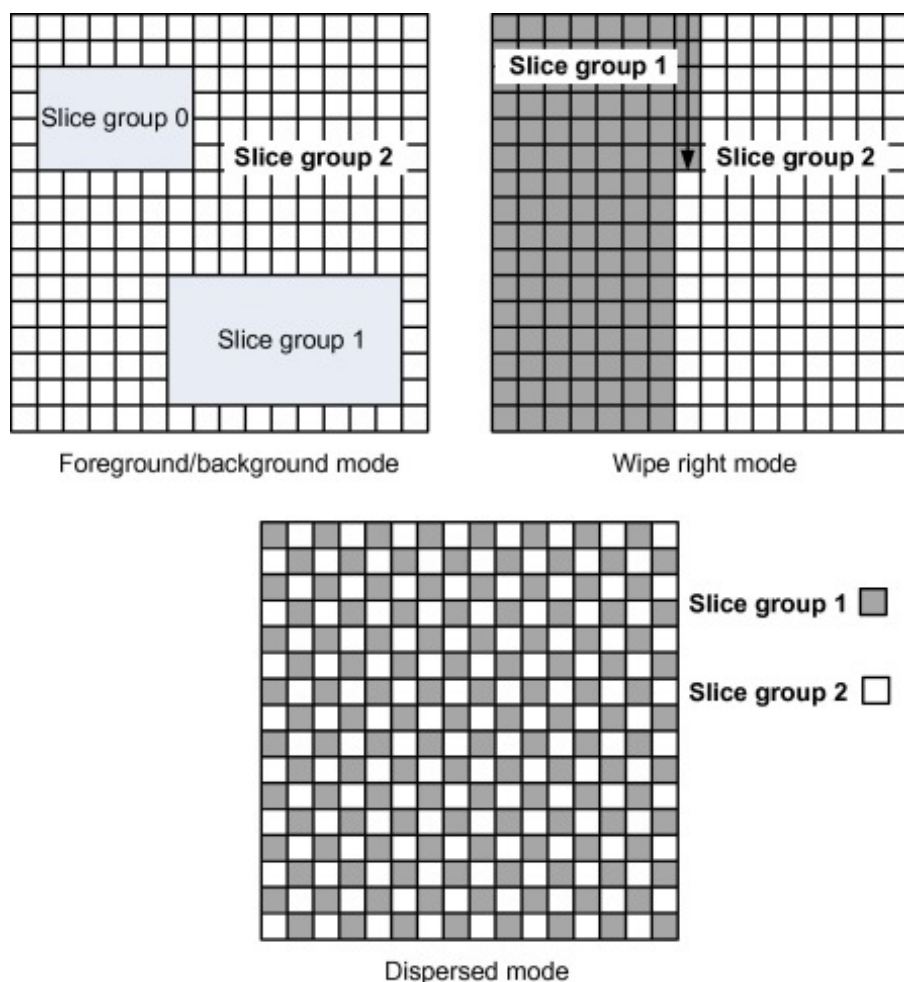
##### 图片类型

H.264/AVC支持以一致的方式处理渐进式和交错格式，作为一个单一单位。PPS中表明了帧是渐进的或交错的，可以使用自适应字段帧编码（将帧编码为单个单元或两个字段），但这不影响通过解码器缓冲区解码图像的方式。

H.264/AVC支持I型帧、P型帧和B型帧。在H.264中，B帧被称为广义B帧，因为它们可以从一个或两个参考帧中预测。H.264/AVC支持从QCIF到4K超高清电视的多种图片格式。

##### 切片和切片组

如第11章所述，切片可以排列成切片组（SG），其中每个SG可以包含一个或多个切片。切片数量由PPS参数发出信号。在多个切片组的情况下，切片的映射安排必须使用0-6范围内的值发出信号。例如，SG类型0是交错的，SG类型1是指分散分配，SG类型2表示前景和剩余情况。根据切片组映射类型，需要额外的语法元素，以使解码器能够将收到的宏块重新映射到编码器使用的切片组。我们在第12.7.3节中介绍了H.264/AVC中使用的各种类型的切片，选定的切片组模式如图12.12所示。



[下载：下载全尺寸图像](#)

图12.12。切片结构示例。

H.264中的切片组支持独立解码，例如，在多个描述编码和[错误隐藏](#)中需要（两者都在[第11章](#)中描述）。开关切片也包含在H.264中。SI和SP切片指示解码器跳转到流中的不同点或流中。这些功能对于使用快进等操作的视频流非常有用。

## 街区

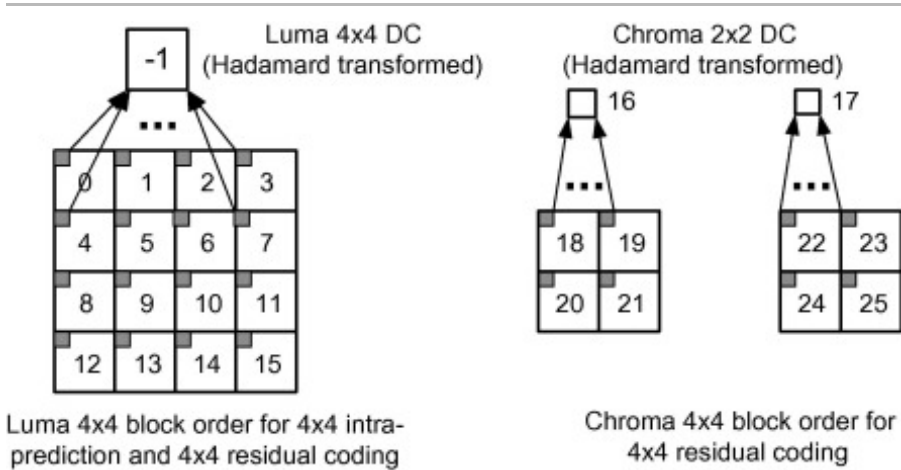
H.264中的[宏块](#)基于一个 $16 \times 16$ 样品排列，包括亮度和[色度](#)残差数据。与H.264/AVC的区别在于，可以根据可用的预测模式以各种方式细分该结构。下面将对此进行更详细的讨论。

### 12.7.5。视频编码层

#### 编码内

- **I\_PCM模式**：这是最基本的H.264模式，并关闭了与标准相关的所有正常预测和转换模式。这不提供任何压缩，但在PCM方法可能比预测转换方法提供更好的速率-失真性能的高质量机制中非常有用。

- **预测内部：** H.264/AVC受益于一系列预测内部模式的引入。当本地细节水平很高时，特别是边缘等定向功能，这些功能会带来显著的好处。H.264支持内部预测 $4 \times 4$ ,  $8 \times 8$ （仅限高知名度），以及 $16 \times 16$ 街区。给定块的编码是参照位于当前块左侧和上方的先前编码块的样本子集执行的。[第9章](#)详细讨论了这些模式。
- **变换：** 在内部\_的情况下 $4 \times 4$ 模式，预测后使用核心转换的**剩余信号** $4 \times 4$ [第9章](#)所述的整数变换。对于内\_的情况 $16 \times 16$ 模式，16种模式 $4 \times 4$ 如[图12.13](#)所示，残余块被转换，16个直流系数中的每个都使用 $4 \times 4$  **哈达玛变换**。对转化色度残差的直流系数也采取了类似的方法，但这次 $2 \times 2$ 应用哈达马变换。



[下载：下载全尺寸图像](#)

图12.13。预测残差和编码块排序的哈达马变换。

### 编码间

- **可变块大小：** 我们在[第9章](#)中看到了使用多个块大小进行帧的预测编码如何在处理复杂纹理和运动方面提供显著的好处，比使用固定块大小节省高达15%的比特率。为了方便起见，H.264/AVC中可用的块大小范围在[图12.14](#)中再次显示，[图12.15](#)给出了典型块分解的例子。然而，为了最大限度地受益于这种灵活性，在速率-失真优化方面，编码过程必须增加很大的复杂性（见[第10章](#)）。

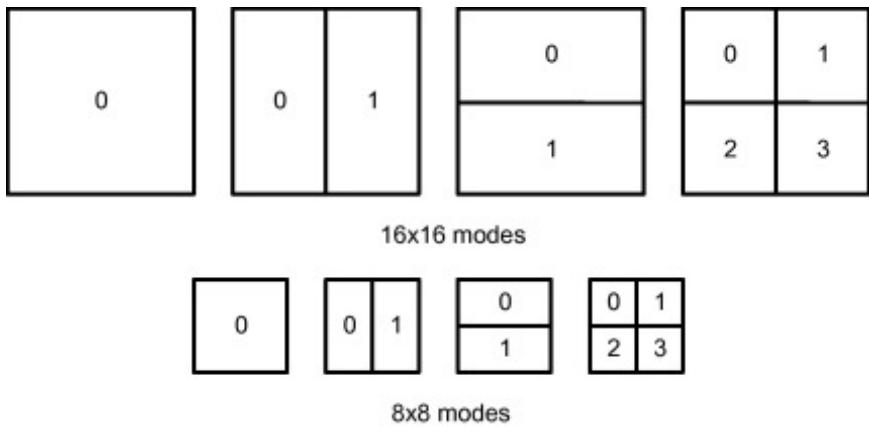
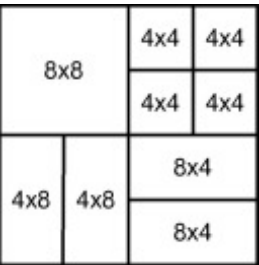


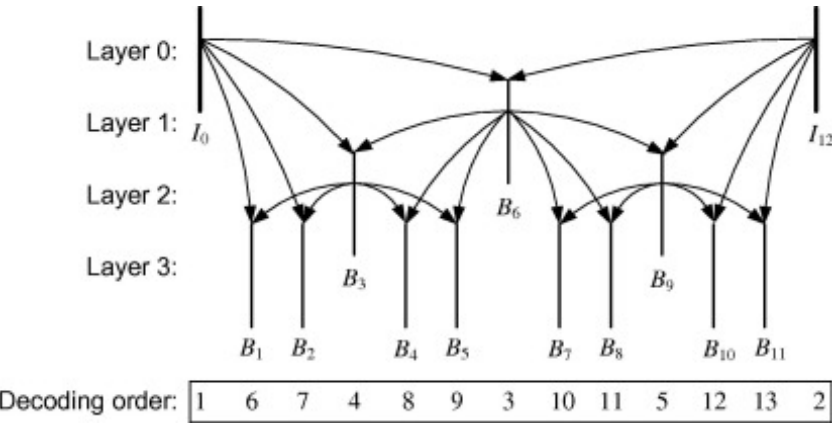
图12.14。H.264/AVC支持的可变块大小。顶部： $16 \times 16$ 模式1-24。底部： $8 \times 8$ 模式1-24。



[下载：下载全尺寸图像](#)

图12.15。示例H.264/AVC宏块分区用于互编码。

- **变换：** H.264/AVC采用 $4 \times 4$ 类DCT整数变换（第5章和第9章，第5 and 章第9章）用于基线和主配置文件中的剩余编码。除此之外，使用 $2 \times 2$ 在色度信号的情况下，哈达马德变换在直流值上。对于高调，该标准还包括一个 $8 \times 8$ 转换，因为这通常更适合空间分辨率较高的内容。
- **多个参考帧：** H.264参考缓冲区中最多可以缓冲16帧或32个字段（取决于配置文件）。参考缓冲区不必包含16个最新帧；只要编码器和解码器保持同步状态，可以根据它们的效用更新帧。有关MRF-ME的更多详细信息，请参阅第9章。
- **预测结构：** H.264编码机制使时间预测方法更加灵活。例如，在低延迟、低内存约束的情况下，一个IPPP...I结构最合适，如基线剖面所提供。如果有些延迟是可容忍的，加上B帧和多个参考帧的可用性，那么可以实现更好的压缩性能。图12.16中有一个例子，其中施加了分层预测结构。使用这种方法已经证明[12]，如果以可控方式增加量化参数，使用图层，则可以实现更有效的压缩性能。



[下载：下载全尺寸图像](#)



图12.16。H.264中的分层B-图片。

- **加权预测：**这是一个有用的工具，在H.264中引入，可以加权/抵消B帧中的预测样本值。这在加速运动或场景切割时提供编码增益。
- **四分之一像素预测：**H.264受益于[亚像素](#)运动补偿的可用性，从而达到四分之一像素的准确性。据估计，在整数像素精度的情况下，这可以节省高达20%的费用。第9章全面描述了H.264/AVC子像素运动估计和补偿过程。H.264/AVC使用五点滤波器进行半像素插值，使用简单的两点滤波器进行局部细化，以达到四分之一像素精度。

## 解封操作

H.264解块过滤器已被评估为将相同主观质量的编码比特率降低高达10%。提高了解码画面的主观视觉和客观质量。这种高度内容自适应的[非线性过滤器](#)可以去除阻塞工件，并且不会不必要地模糊视觉内容。其结构和操作在[第9章](#)中进行了描述。

## 变长编码

- **Exp-Golomb编码：**这是一种高度结构化和简单的可变长度编码方法，在H.264中几乎普遍用于除变换系数以外的所有符号。
- **CAVLC：**上下文自适应VLC是一种相对低复杂度但有效的方法，用作变换系数熵编码的基线和扩展剖面。编码器使用局部上下文，根据局部统计信息在不同VLC表中进行选择，在这种情况下，相邻块中的非零系数数。
- **CABAC：**CABAC是一种更[复杂](#)、基于[算法](#)编码的方法，仅在H.264主图和高剖面图中支持。它使用本地上下文来调整其条件概率。与CAVLC相比，CABAC的比特率降低了10到20%，这取决于内容类型和[量化水平](#)。

有关上述所有熵编码方法的更多详细信息，请参阅[第7章](#)。

## 编码器控制

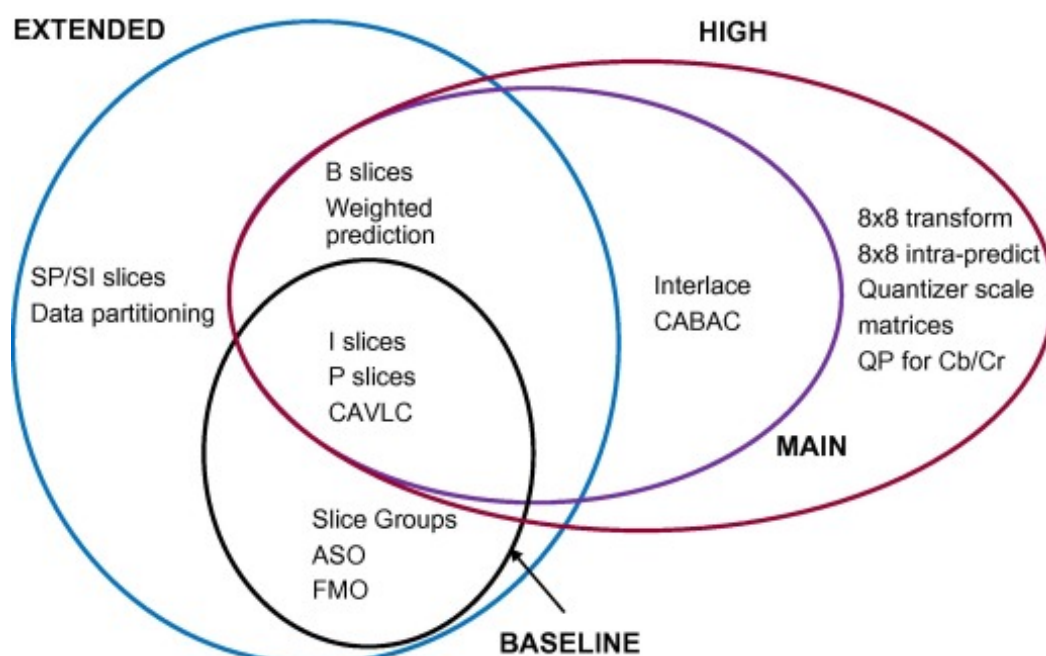
编码器控制是H.264/AVC的非规范部分。然而，这是从标准中获得最佳性能的关键因素，因为如果没有良好的RDO，包含的许多高级功能（例如块大小选择）将无法节省大量费用。编码器控制的目标是选择视频信号的哪些部分应该使用哪些方法和参数设置进行编码。通常采用拉格朗日优化方法，其中失真度量， $D$ ，基于平方差和比率之和， $R$ ，包括与决策相关的所有位，包括标头、运动、侧面信息和转换数据。有关H.264/AVC中使用的RDO方法的更多详细信息，请参阅[第10章](#)。

### 12.7.6.概况和级别

H.264/AVC始于相对较少的配置文件和级别。然而，自推出以来，该数字大幅增加，以反映其在日益广泛的应用程序中的受欢迎程度，以及对处理新兴格式的更高质量配置文件的需求。该标准

现在支持21个配置文件，从最简单的受限基线配置文件（CBP）和基线配置文件（BP）到高4:2:2配置文件（Hi422P-用于工作室工作）和高4:4:4预测配置文件（Hi444PP）。主配置文件（MP）通常用于标准定义（4:2:0）DVB广播。扩展配置文件（XP）用于流媒体，高配置文件（HiP）用于蓝光播放器和DVB HDTV传输。

H.264还支持一系列关于内存和计算能力的解码器约束，以每秒宏块的数量为单位，指定了解码比特率、解码图片缓冲区大小、图像大小和处理速度等参数。图12.17显示了H.264配置文件的精选内容，参考文献中提供了所有配置文件和级别的良好摘要。[13]。

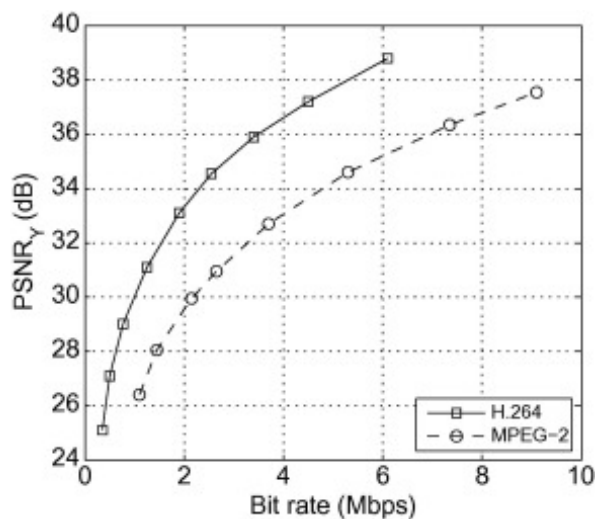


下载：下载全尺寸图像

图12.17。精选的H.264配置文件。

### 12.7.7。表演

H.264/AVC相对于以前标准的性能摘要。[14]。图12.18显示了娱乐应用程序的RD配置文件示例。这表明H.264/AVC和MPEG-2之间的相同PSNR值节省了大约50%的比特率。Ohm等人[15]提供了包括HEVC在内的最新比较（见表12.7）。



[下载：下载全尺寸图像](#)

图12.18。标准定义娱乐内容的典型H.264/AVC（MP）性能与MPEG-2相比。

### 12.7.8。可伸缩扩展

2007年批准了H.264/AVC（SVC）的可扩展扩展。如第11章所述，比特流可伸缩性有助于在动态信道条件下提供优雅的性能降级。SVC包含五个额外的配置文件，为可伸缩性提供了分层方法。这创建一个嵌入式比特流，提供信噪比、空间和/或时间可伸缩性，其中数据包可以有序地丢弃，以管理拥塞或终端限制。从图12.16可以看出，使用分层B帧是提供时间可伸缩性的非常有用的工具。有关SVC的更多详细信息，请参阅参考文献[16]。

### 12.7.9。多视图扩展

最后，为了适应三维立体视频内容和广播的预期繁荣，2009年批准了多视图扩展（MVC），其中包括另外三个H.264配置文件。顾名思义，MVC支持从多个相机获取的内容编码。其主要重点是立体（3-D）视频，但也支持自由视角和多视角三维应用程序、电视和多视角三维电视。立体高轮廓用于三维蓝光设备。有关MVC的更多信息，请参阅Refs。[17]，[18]。

## 12.8。H.265/HEVC

### 12.8.1。简短的背景

HEVC，正式名称为ISO/IEC MPEG-H Part 2和ITU-T H.265，是ISO/IEC MPEG和ITU-T VCEG之间合作的最新成果。该标准于2013年4月获得批准[19]，随着专利许可框架的到位和几个HEVC支持产品的宣布，它即将对视频行业产生重大影响。与AVC相比，HEVC具有高达50%的压缩效率提高的潜力，并将在一定程度上解决移动视频的爆炸，以及支持广播和下载服务的需求，其质量和体验水平不断提高。

HEVC的目标是与H.264/AVC相同的应用程序组合，但特别侧重于降低比特率以提高视频分辨率，支持并行处理以及损失恢复力和易于与适当传输机制集成。2012年12月出版的IEEE视频技术

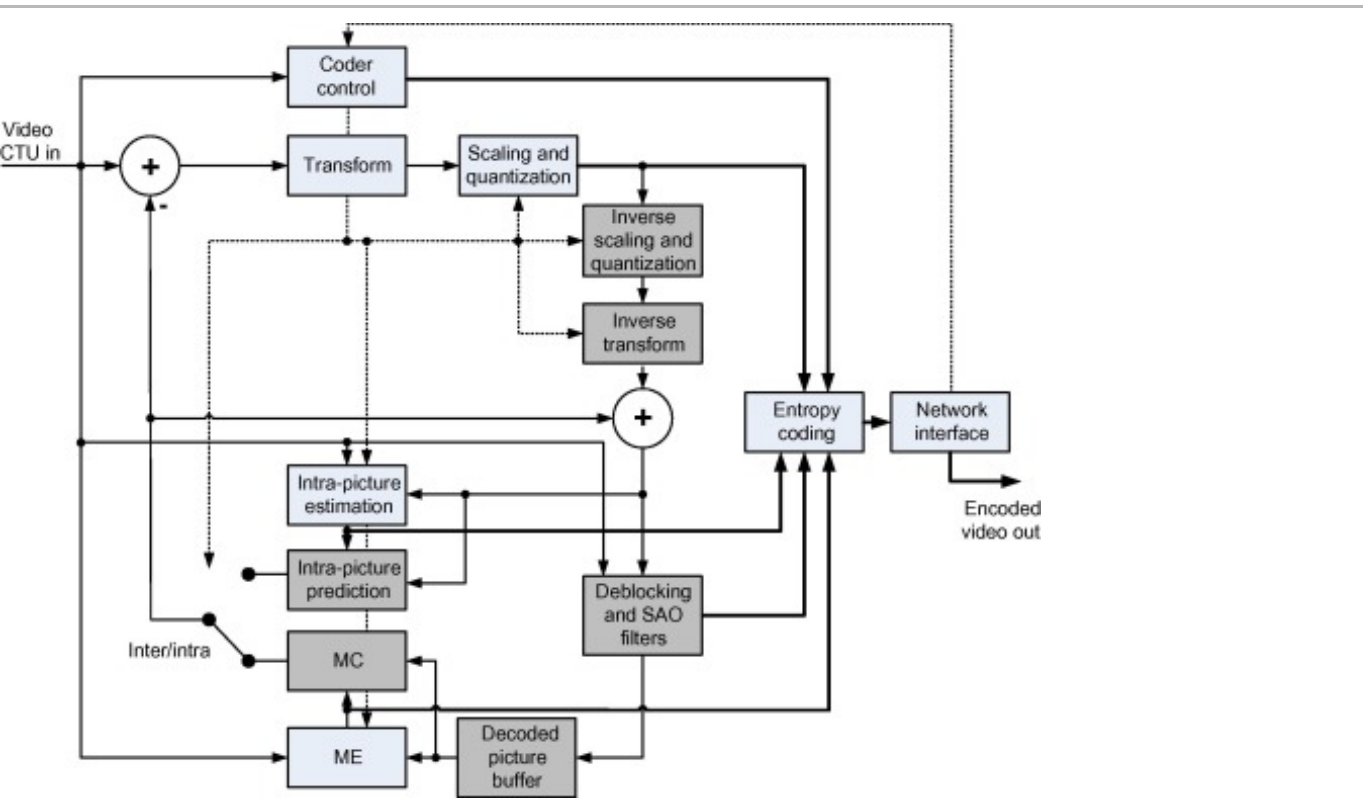
电路和系统交易特刊上，特别是Sullivan等人[20]的概述和Ohm等人的性能比较，对HEVC功能的范围进行了出色的介绍。[15]。有关完整细节，读者应查阅标准文档[19]。

### 12.8.2. 主要特征

正如我们稍后将看到的，HEVC在保留图像质量的同时，与现有编解码器相比，将比特率要求降低了大约一半。它使用自H.261以来为我们服务良好的混合方法来做到这一点，但增加了几个新的和重要的功能。虽然其中一些增加了编解码器的复杂性，但与H.264/AVC相比，许多编解码器实际上提供了简单性，并方便了并行处理，与最近的硬件开发兼容。HEVC和以前的标准一样，可以根据应用要求来权衡计算复杂度、压缩率、对错误的鲁棒性以及处理延迟时间，并且这一范围的操作点由其配置文件和级别结构定义。HEVC的一个主要重点是下一代高清电视显示器和采集系统，该系统具有逐步扫描帧、更高的帧率和高达UHDTV的分辨率。

在以下章节中，我们更仔细地研究了HEVC的视频编码和句法结构。读者将看到，H.264/AVC中提取了大量HEVC技术，省略了一些H.264功能，并酌情添加了其他新的和修改的功能，以提高性能。HEVC的主要新功能与 (i) 其块结构有关，其中使用称为编码树单元 (CTU) 的超级宏块，其中CTU的大小可以高达 $64 \times 64$ 样本；(ii) 使用这些CTU使用水树结构将这些CTU划分为可变大小的编码、转换和预测块的方式；(iii) 显著增强的内预测，支持35种模式；(iv) 使用示例自适应偏移 (SAO) 与预测间循环中的去块过滤器一起减少带状和环状伪影。

HEVC编解码器的架构说明了其中许多功能，如图12.19所示。以下各节将更详细地对此进行描述。



下载：下载全尺寸图像

图12.19。HEVC视频编码器架构。

### 12.8.3。网络抽象和高级语法

#### NAL单位

HEVC的高级语法保留了H.264/AVC中使用的一些元素，以提供网络抽象和错误恢复力的基础。NALU仍然被用作传输机制的基本接口，HEVC支持63种NALU（VCL和非VCL）类型[19]，[20]。

#### 切片结构

切片启用独立解码，并协助[同步](#)和错误鲁棒性。HEVC中的切片结构在很大程度上是从H.264/AVC中提取的，读者请参阅第9章和第11, [章, 第, 9 and 章第11章](#)，以进行更详细的描述。

#### 参数集

使用与H.264/AVC相同的参数集，辅之以新的视频参数集（VPS）。VPS可以包括元数据来描述编码层的一般特征，例如在未来的可伸缩扩展中可能有用。NAL单元头中的信息也支持识别时间层。

#### 参考图片集和参考图片列表

管理解码[图片缓冲区](#)（DPB）的内容很重要，以确保编码器和解码器之间的同步。保留的图片集称为参考图片集（RPS）。HEVC支持两个参考图片列表，RPL0和RPL1。在单向预测的情况下，可以从任一列表中选择参考图片，在[双向预测](#)的情况下，每个列表必须有一张图片。HEVC进行了一些改进，使RPS的识别比H.264/AVC更健壮。

### 12.8.4。图片和分区

HEVC支持非常广泛的图片大小和帧速率，如[表12.5](#)中的级别示例所示。HEVC支持的基本图片类型与H.264/AVC中相同，但IDR帧的修改版除外，现在称为清洁随机访问（CRA）图片。

表12.5。HEVC级别示例。

等级	最大亮度采样率	最大亮度图片率	最大比特率（kbps）	示例解决方案
1	552,960	36,864	128	176 × 144@15
3	16,588,800	552,960	6000	720 × 576@30
5	267,386,880	8,912,896	160000 <sup>a</sup>	3,840 × 2160 @60
6.2	4,278,190,080	35,651,584	800万 <sup>a</sup>	7680 × 4320 @120

<sup>a</sup>



高层规格，其他是主层。

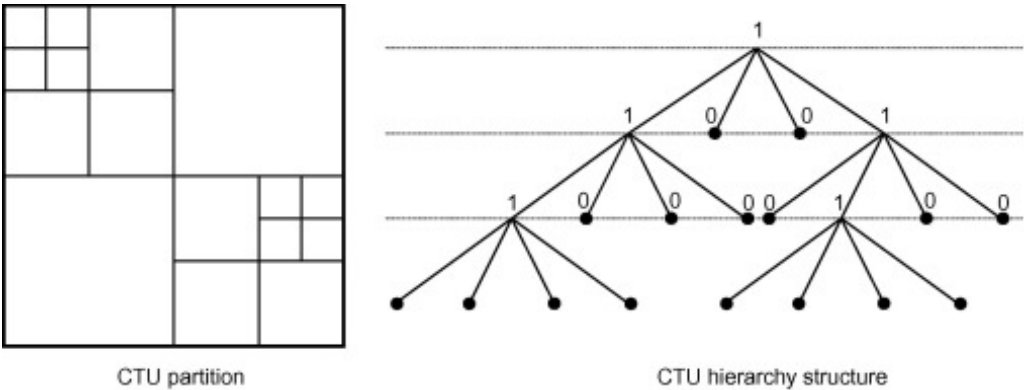
HEVC编码单元（CU）最多可以包含三个块（流明和色度样本阵列）以及解码这些块所需的任何信息。每个CU被细分为预测单元（PU），这些单元使用内部或内部预测进行编码。每个PU内都会发出适当的预测参数（内部或内部）信号。然后使用剩余四叉树（RQT）将CU划分为转换单元（TU），其中对剩余信息进行编码。如上所述，这些单元都包含组成亮度和色度成分的编码块（CB）、预测块（PB）和转换块（TB）。参考文献更详细地描述了这种结构。[\[21\]](#)。HEVC图片的分解将在下面更详细地描述。

### 编码树单元（CTU）和编码树块（CTB）

HEVC用一个 $L \times L$  CTU在哪里 $L = 16, 32, 64$ 。这为高分辨率格式提供了更有效的编码。CTU由亮度和色度CTB组成。

### 四叉树CTU结构

使用四叉树分解，CTB可以分为编码单元（CU），由亮度和色度编码块（CB）组成，块大小较小。最大编码单元（LCU）的大小是 $64 \times 64$ 。CTU的分解方式在标准中没有具体说明（它不是规范的），但仍然是编码器RDO处理的重要组成部分。[图12.20](#)给出了CTU分区的一个例子。每个CU进一步分为变换单元（TU）和预测单元（PU）。这种方法为编码器提供了更大的灵活性，可以根据视频内容优化块分区。



[下载：下载全尺寸图像](#)

图12.20。HEVC中的CTU分区——一个例子。

### 预测单元（PU）

HEVC支持用于编码内部和编码之间的更多种类的块分解，从 $4 \times 4$ 到 $64 \times 64$ 。

### 转换单元（TU）

预测残差，无论是来自编码内部还是编码间，都使用类似于DCT的整数变换进行转换，其大小来自 $4 \times 4, 8 \times 8, 16 \times 16$ ，以及 $32 \times 32$ 。对于预测内残留的情况， $a4 \times 4$ 整数正弦变换也可用，因为这在



相关性较低的残差上提供了更好的性能。

### 随机接入点（RAP）和清洁随机接入（CRA）图片

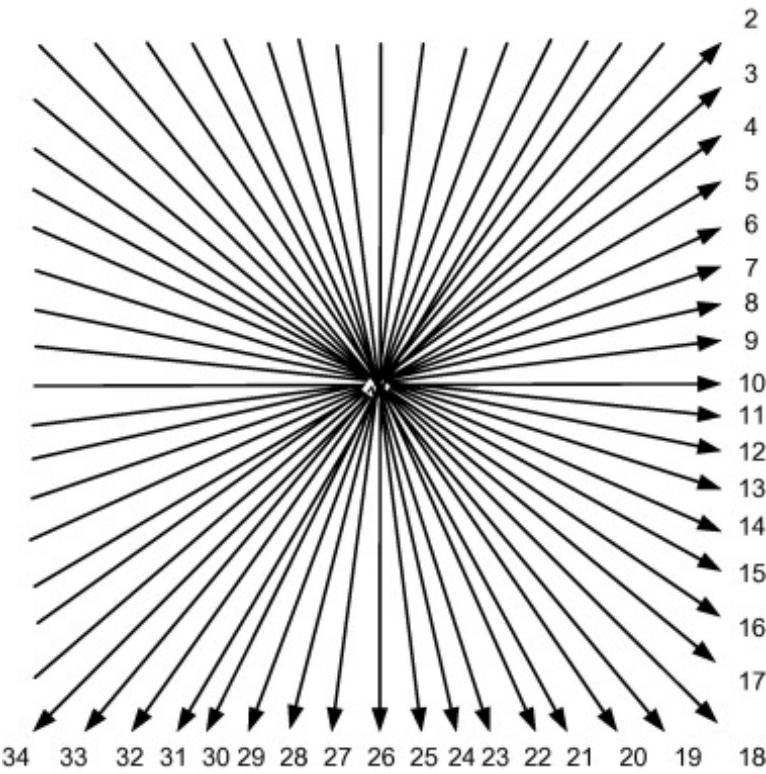
CRA图片的概念与H.264/AVC中的IDR帧相同。CRA通过体现独立编码的图片来推广这一点，其中以下帧不依赖于CRA图片之前收到的帧。CRA图片可以放置在需要随机访问的地方，称为RAP。

### 12.8.5。视频编码层（VCL）

与H.264类似，HEVC使用混合编码结构，将运动补偿与块变换和熵编码相结合。几项创新已被纳入HEVC，这些创新为性能提供了显著收益，这些创新将在下面描述。

### 编码内

与H.264/AVC类似，相邻解码块的边界像素值可以用作空间预测的基础。虽然H.264/AVC只支持8种预测模式，但在HEVC中，这已显著增强，有35种模式，包括33种定向模式（图12.21）以及表面拟合和直流预测。有关计算这些模式预测的更多详细信息，请参阅参考文献。[19]。模式0中的平面振幅是根据周围块中参考像素值得出的水平和垂直斜率计算的。直流模式（1）的计算方式与H.264/AVC相同——基于块边界参考值的平均值。



[下载：下载全尺寸图像](#)

图12.21。HEVC预测模式（模式0=平面模式，模式1=直流）。

HEVC还扩展了参考和边界样品平滑的概念。这些操作通过在预测之前过滤参考样本（来自周围

块) 或边界样本 (在预测块内) 来帮助减少剩余和重建边缘。

当使用模式内时, 预测块 (PB) 大小与编码块 (CB) 相同, 但最小CB大小除外, 其中进一步分区为四个PB象限 $4 \times 4$ 允许。预测后, 残差使用正向整数变换、量化和熵编码。

编码间

交叉预测和编码的方法类似于H.264/AVC。多个参考帧存储在解码图像缓冲区中, 其使用方式与H.264/AVC相似, 支持单预测和双预测模式。与H.264/AVC一样, 可以使用加权预测。HEVC中包含的一些修改如下:

- **亚像素运动补偿:** 四分之一像素精度用于估计时间预测期间的运动。局部亮度插值为半像素分辨率由八点滤波器提供, 而不是H.264/AVC中的六点滤波器。然后, 七点过滤器用于四分之一样本插值, 而不是H.264/AVC中使用的两点过滤器。这些是可分离的 (与H.264不同) 不使用中间四舍五入。这提高了精度, 简化了预测架构。用于插值的**滤波系数**在方程 (12.1) 中给出, (12.2) 。有关这些用途的更多详细信息, 请参阅参考文献。[19], [20]。

$$h_h[i] = \{-1, 4, -11, 40, 40, -11, 4, -1\} / 64$$

(12.1)

$$h_q[i] = \{-1, 4, -10, 58, 17, -5, 1\} / 64$$

(12.2)

- **运动矢量预测:** 高级运动矢量预测模式 (AMVP) 用于初始化和推导可能的运动矢量候选模型。此外, 还包含改进的跳过模式, 以及在相邻块之间合并MV的能力。
- **编码块划分为预测块:** 在相互预测的情况下, 亮度和色度CB可以分为一个、两个或四个PB (后者仅适用于最小的CB大小) 。这些分区选项如图12.22所示。

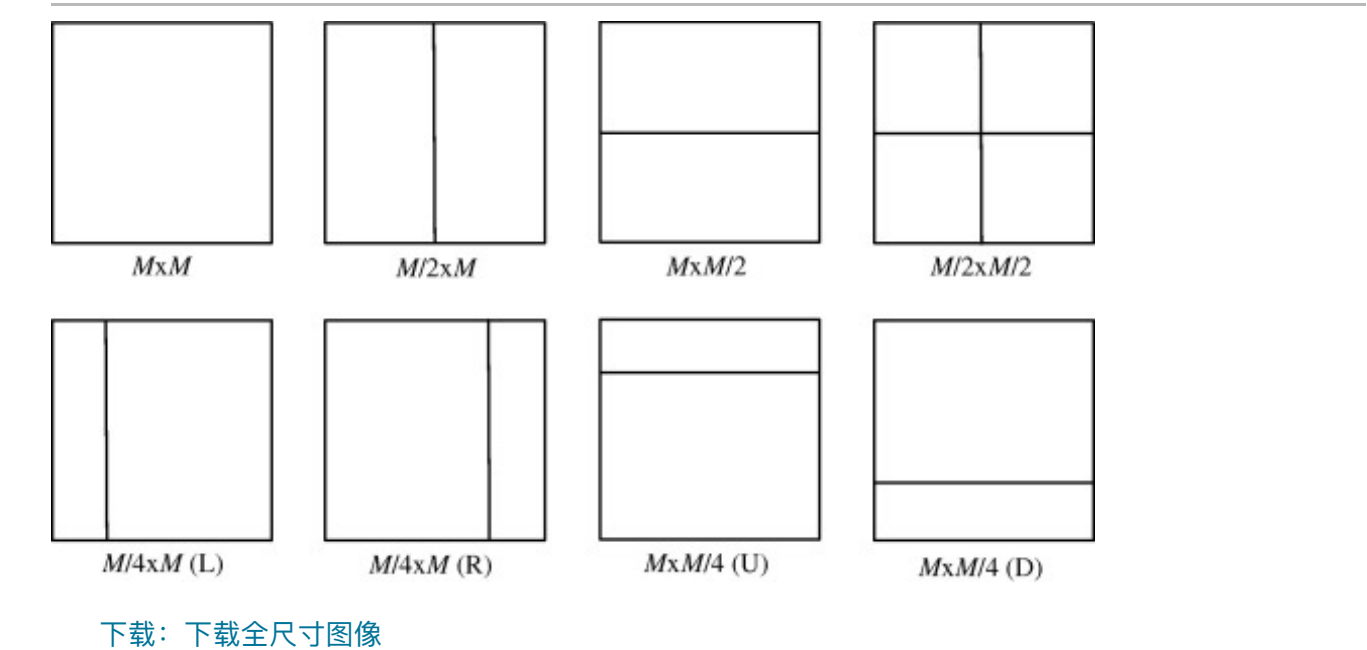


图12.22。用于相互预测PB的CB分区。

## HEVC中的变换

预测后的剩余数据使用可分离整数变换进行编码，其结构取决于预测模式和块大小。HEVC支持的变换块大小为： $4 \times 4$ ,  $8 \times 8$ ,  $16 \times 16$ ，以及 $32 \times 32$ 。至于H.264/AVC，变换基函数来自DCT。变换被定义为 $32 \times 32$ 通过对该较大的矩阵进行子采样，可以获得基函数矩阵和较小的变换大小。例如， $16 \times 16$ 矩阵派生自 $32 \times 32$ 通过从第0行、第2行、第4行中提取前16个值来矩阵。类似的子采样用于生产 $8 \times 8$ 和 $4 \times 4$ 变换矩阵。例如， $8 \times 8$ 变换如下所示。

$$\mathbf{A}_{8 \times 8} = \begin{bmatrix} 64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 \\ 89 & 75 & 50 & 18 & -18 & -50 & -75 & -89 \\ 83 & 36 & -36 & -83 & -83 & -36 & 36 & 83 \\ 75 & -18 & -89 & -50 & 50 & 89 & 18 & -75 \\ 64 & -64 & -64 & 64 & 64 & -64 & -64 & 64 \\ 50 & -89 & 18 & 75 & -75 & -18 & 89 & -50 \\ 36 & -83 & 83 & -36 & -36 & 83 & -83 & 36 \\ 18 & -50 & 75 & -89 & 89 & -75 & 50 & -18 \end{bmatrix} \quad (12.3)$$

在编码内的情况下 $4 \times 4$ 残差，使用另一种整数变换——来自DST。据观察，由于误差值随着与用于预测它们的边界样本的距离而增加，内部残差的统计量更适合DST的基础函数。此 $4 \times 4$ 因此，变换矩阵是：

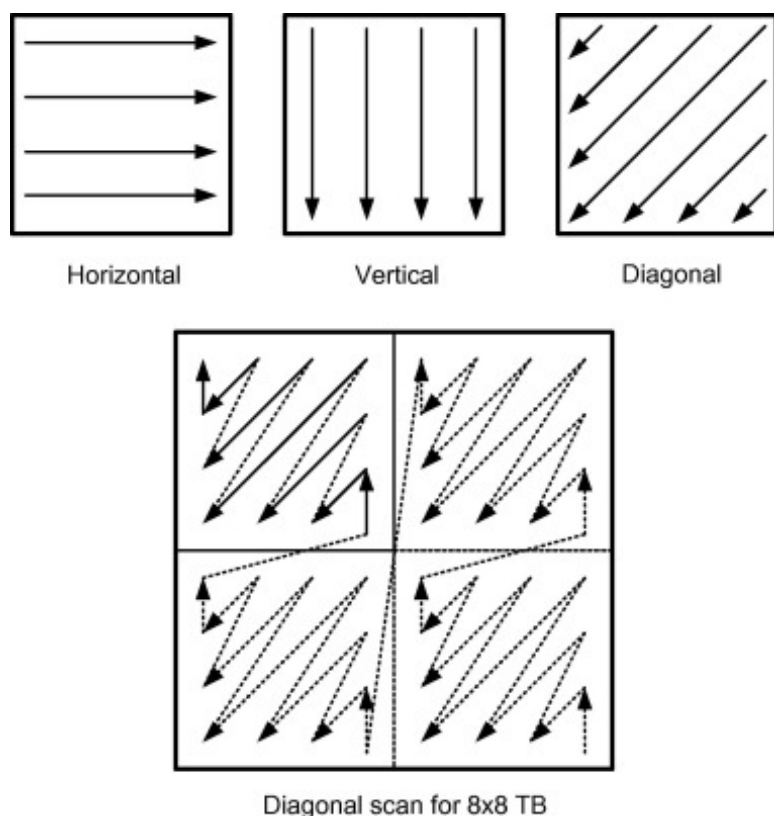
$$\mathbf{A}_{4 \times 4I} = \begin{bmatrix} 29 & 55 & 74 & 84 \\ 74 & 74 & 0 & -74 \\ 84 & -29 & -74 & 55 \\ 55 & -84 & 74 & -29 \end{bmatrix} \quad (12.4)$$

## HEVC中的量化

HEVC采用与H.264相似的量子化方法转换系数。它使用均匀的量化，每个块大小都有一系列缩放矩阵。

### 系数扫描

HEVC TB[22]中系数的扫描基于使用 $4 \times 4$ 所有结核病大小的块。代表每个系数组的三个系数扫描选项（CG：一组连续16个系数）如图12.23所示。在块内的情况下，使用一种称为模式依赖系数扫描（MDCS）的新方法，其中选择最接近预测模式内部方向的模式。对于编码块间，只使用对角线模式。对于较大的块间，系数从最小有效到最大有效，如图12.23所示，对于 $8 \times 8$ 结核病。与H.264/AVC和之前标准中使用的之字形扫描不同，这种简化方法使扫描与系数上下文更好地匹配。在编码效率方面，据报道，与H.264/AVC相比，增长了约3.5%[22]。



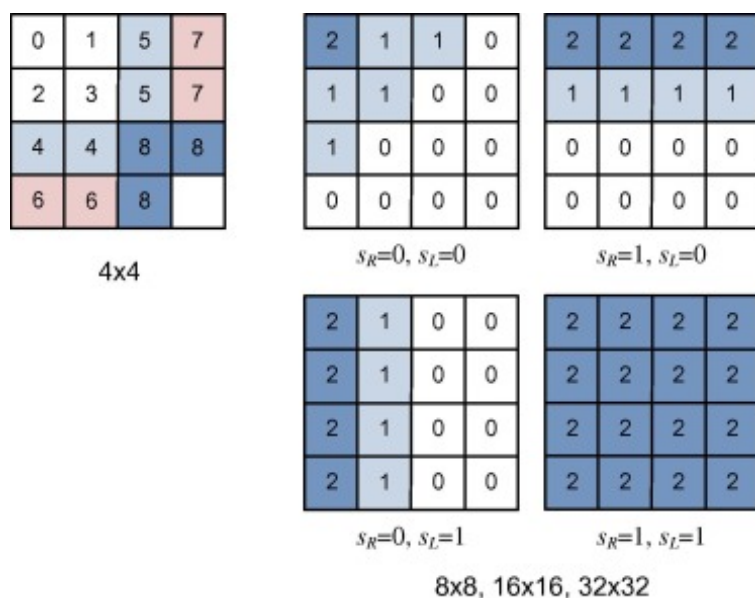
[下载：下载全尺寸图像](#)

图12.23。HEVC中的系数扫描。

HEVC相对于AVC的进一步修改与扫描中最后一个有效系数的信号传递有关。正如我们所看到的，对最后一个有效系数发出信号，通过消除对尾随零进行显式编码运行的必要性，减少了编码垃圾箱的数量。在HEVC中，最后一个显著系数的位置被明确编码，后面跟着标志，以指示显著系数。

## 背景和意义

与H.264/AVC一样，HEVC使用编码块标志（CBF）来表示结核病的重要性，该标志指示TB是否包含任何非零系数。CBF的上下文取决于四叉树分解中块的水平。HEVC中的显著性地图的编码方式利用显著性地图的**稀疏性**，使用两次传递。首先编码系数组的意义，然后编码其中系数组的意义。HEVC中的系数意义上下文 $4 \times 4$ 和 $8 \times 8$ 块如图12.24所示。意义标志根据上下文模型为最后一个系数和直流电之间的每个系数编码，该上下文模型取决于其位置和邻居的重要性。在 $4 \times 4$  TB，系数根据频率和分布进行分组，其中分布相似的系数被分组在一起。



[下载：下载全尺寸图像](#)

图12.24。意义标志上下文 $4 \times 4$ 和 $8 \times 8$ HEVC转换块。

对于 $8 \times 8$ ,  $16 \times 16$ 和 $32 \times 32$ HEVC中的结核病，上下文建模再次基于位置，但根据图12.24所示的模板（右）。四个模板中有一个是根据 $s_R$ 和 $s_L$ 在哪里 $s_R$ 和 $s_L$ 分别对应右侧和下部子块的编码子块标志。

## 使用CABAC进行可变长度编码

如第7章所述，CABAC是一种强大的算法编码方法，在H.264/AVC主图和高剖面图中得到支持。由于其性能优势，HEVC在所有熵编码目的都采用了CABAC。与H.264/AVC一样，HEVC中的CABAC使用本地上下文来调整其条件概率。CABAC通过以下方式改进了HEVC：

1. 更好的上下文建模：基于HEVC的树结构方法的利用。
2. 减少上下文编码的垃圾箱数量：虽然上下文的数量少于H.264/AVC，但其性能更好。
3. 引入增强型旁路模式：更广泛地使用旁路，以减少需要使用CABAC上下文处理的数据量，从而增加吞吐量。

## 循环过滤器

HEVC在更新其解码图像缓冲区之前包括两个循环过滤阶段：解块过滤器（DBF）和示例自适应偏移（SAO）过程。

- **去阻塞过滤器**：HEVC中的去阻塞操作与第9章中描述的H.264/AVC非常相似。它略微简化，使其更易于并行处理。DBF适用于所有TU和PU边界，除非它们与图片或切片边界重合。在 $4 \times 4$ 仅当块边界与 $8 \times 8$ 边界。这简化了处理，通常性能没有明显的下降。就HEVC而言，只有三种解块优势，而不是五种H.264/AVC。
- **样本自适应偏移**：SAO是HEVC中的一项新功能，在DBF之后调用，目的是通过减少带状和响

铃工件来提高图像质量。这是一个非线性调整，根据在编码器上创建（并传输到解码器）的查找表，使用信号振幅的直方图分析为样本添加偏移量。有两种SAO模式：边缘偏移模式和带偏移模式。第一个在CTB的基础上使用四个梯度算子之一来比较给定方向的样本和相邻值。根据样品相对于其邻居的相对值，可以产生五种条件。例如，如果样本，考虑水平方向的情况 $p$ 有一个左邻 $n_0$ 和右邻 $n_1$ ，然后根据表12.6对区域进行分类。在第二种情况下，带偏移，直接根据样本振幅应用偏移量。将全振幅范围分为32波段，并使用传输偏移量修改连续四个波段的样品值。这有助于减少平滑区域的带状伪影。

表12.6。SAO边缘分类。

EdgeIdx	条件	案例
0	案件例外 (1...4)	单调
1	$p < n_0$ 和 $p < n_1$	当地分钟。
2	$p < n_0$ 和 $p = n_1$ 或者 $p < n_1$ 和 $p = n_0$	边缘
3	$p > n_0$ 和 $p = n_1$ 或者 $p < n_1$ 和 $p = n_0$	边缘
4	$p > n_0$ 和 $p > n_1$	本地最大值。

12.8.6。概况和级别

HEVC版本1指定了三个配置文件，主图、主图10和主静止图。

主要配置文件

主配置文件支持每个样本8位的位深度和4:2:0（和4:0:0）色度采样，采用前面小节中描述的功能。

主要10配置文件

Main 10配置文件支持高达10位的位深度和高达4:2:2的色度采样。在最近的测试中，已证明Main 10配置文件优于Main Profile，在相同的PSNR值下，比特率降低约5%。

主静止图片个人资料

HEVC还像H.264一样，包括静态图像编码配置文件，这是主配置文件的子集。基于MOS和PSNR分数的性能比较表明，HEVC仍然的图像编码在基于PSNR的比特率下降约20%和MOS分数的30%方面，仍然比JPEG 2000优于JPEG 2000。同样，它的表现也比JPEG好40%和60%。

级别

与之前的标准一样，HEVC还支持一系列级别，这些级别施加参数限制，限制解码器的性能。目



前有13个与HEVC相关的级别。一些例子见[表12.5](#)。

### 12.8.7。扩展

在编写本报告时，正在开发一些HEVC范围扩展和五个额外的配置文件。这些包括主12、主4:2:210、主4:2:212、主4:410、主4:4:4 12。从标题中可以明显看出颜色采样和位深度的扩展范围。

### 12.8.8。HEVC的性能比最近的标准有所提高

Ohm等人[\[15\]](#)提供了编码标准的优秀比较，其中报告了HEVC和以前标准的全面比较。我们在[表12.7](#)中转载了他们的结果，指出交互式应用程序的改进程度更高。[根据经验](#)，我们之前曾表示，自1984年引入H.120以来，视频编码标准每10年为等效视频质量提供一半的比特率。[表12.7](#)的结果证明了这一点。

表12.7。娱乐应用视频编码标准对比使用PSNR测量的等客观质量值显示[平均比特率节省](#)。

视频标准	相对比特率节省 (%)		
	H.264/AVC	MPEG-4	MPEG-2
HEVC议员	35.4	63.7	70.8
H.264/AVC HP	X	44.5	55.4
MPEG-4 ASP	X	X	19.7
H.263 HLP	X	X	16.2

## 12.9。其他事实标准和专有编解码器

还有一些其他视频编解码器。虽然这些值得一提，但我们在这里不会详细介绍。

### 12.9.1。VP9

也许最重要的是谷歌的VP9。VP9是VP8的继承者，开源且免版税。VP8由On2 Technologies创建，该公司于2010年被谷歌收购。同样归谷歌所有的YouTube使用Adobe Flash Player和H.264/AVC，但据报道，有人转向VP9。2013年，谷歌Chrome网页浏览器发布了支持VP9解码的版本。VP9也可能在某些基于云的应用中受到青睐。

### 12.9.2。VC-1

微软WMV 9编解码器是专有编解码器，但在2006年被采用为国际标准SMPTE 421M或VC-1。VC-1是蓝光设备识别的格式。

### 12.9.3。RealVideo

RealVideo包括一系列由RealNetworks为其玩家开发的专有视频压缩格式。目前的版本rv40被认为基于H.264/AVC。

### 12.9.4。狄拉克

狄拉克由英国广播公司研究与开发公司开发，作为主流标准的免版税替代品。Dirac 采用小波压缩而不是 DCT 等块变换，并为 HDTV 格式和其他格式提供良好的性能。仅限I帧的版本Dirac Pro已被BBC内部用于演播室和外部广播应用程序，并于2010年被SMPTE标准化为VC-2。还有一个被称为薛定 ?? 的便携式版本。

## 12.10。摘要

在本章中，我们研究了所有主要视频[压缩标准](#)的主要功能，从H.120到H.261、H.263和MPEG-2，到当前的H.264/AVC标准和新的H.265/HEVC编解码器。尽管这些在很大程度上建立在同一混合运动补偿变换架构上，但过去三十年左右的持续创新平均每10年为同等视频质量提供一半的比特率。虽然这是令人印象深刻的进展，但对大量高质量视频内容的需求仍在增长，将继续需要新的方法来满足未来的需求。下一章将考虑一些潜在的候选解决方案。

[Recommended articles](#)

[Citing articles \(0\)](#)

## 参考文献

- [1] 美国。奥布  
参考模型方法论——协同创建视频编码标准的工具  
IEEE会议记录，83（2）（1995年），第139-150  
[谷歌学术](#)
- [2] CCITT/SG XV，使用主群传输视频会议编解码器，rec.H.120，CCITT（现为ITU-T），1989年。  
[谷歌学术](#)
- [3] 国际电信联盟-电信（ITU-T），建议H.261，视听服务视频编解码器，px64 kbit/s，版本1，1990年；版本2，1993年。  
[谷歌学术](#)
- [4] ITU-T和ISO/IEC JTC 1，移动图片和相关音频信息的通用编码——第2部分：视频，ITU-T Rec.H.262和ISO/IEC 13818-2（MPEG-2视频），1994年版本1。  
[谷歌学术](#)
- [5] ITU-T，用于低比特率通信的视频编码，ITU-T Rec.H.263，版本1，1995版本2，1998版本3，2000年。  
[谷歌学术](#)
- [6] 美国。温格，G。诺尔，J。奥特，F。科森蒂尼

## H.263+中的错误恢复力支持

IEEE 视频技术电路和系统交易, 8 (7) (1998年), p.867-877

[在Scopus中查看记录](#) [谷歌学术](#)

- [7] ISO/IEC JTC 1, 视听对象编码——第2部分: 视觉, ISO/IEC 14496-2 (MPEG-4 Visual), 版本1, 1999年, 版本2, 2000, 版本3, 2004年。  
[谷歌学术](#)
- [8] T. 易卜拉欣, C. 霍恩  
**MPEG-4自然视频编码: 概述**  
信号处理图像通信, 15 (2000年), 第365-385页  
[文章](#)  [下载PDF](#) [在Scopus中查看记录](#) [谷歌学术](#)
- [9] ITU-T和ISO/IEC JTC 1, 通用视听服务高级视频编码, ITU-T Rec.H.264和ISO/IEC 14496-10 (AVC), 版本1, 2003年版本2, 2004年版本3, 版本4, 2005, 版本5, 2006年版本7, 8, 2007年版本9, 10, 11, 2009年版本12, 13, 2010年版本14, 15, 2011年版本16, 2012年。  
[谷歌学术](#)
- [10] T. 维冈德, G. 沙利文, G. Bjøntegaard, A. 阿杰·卢特拉  
**H.264/AVC视频编码标准概述**  
IEEE 视频技术电路和系统交易, 13 (7) (2003), p.560-576  
[在Scopus中查看记录](#) [谷歌学术](#)
- [11] 我。理查森  
**H.264高级视频压缩标准**  
(第二版), 威利 (2010)  
[谷歌学术](#)
- [12] H. 施瓦茨, D. 马普, T. 维冈德  
**分层B图与MCTF解析**  
多媒体和博览会国际会议 (2006年), p.1929-1932  
[谷歌学术](#)
- [13] <[http://en.wikipedia.org/wiki/H.264/MPEG-4\\_AVC](http://en.wikipedia.org/wiki/H.264/MPEG-4_AVC)>, 2013年 (2013年8月访问)。  
[谷歌学术](#)
- [14] J. 奥斯特曼, J. 博曼斯, P. 列表, D. 马普, M. Narroschke, F. 佩雷拉, T. 斯托克哈默, T. 韦迪  
**使用H.264/AVC进行视频编码: 工具、性能和复杂性**  
IEEE Circuits and Systems Magazine, 4 (1) (2004年), p.7-28  
[在Scopus中查看记录](#) [谷歌学术](#)
- [15] J.-R. 欧姆, G. 沙利文, H. 施瓦茨, T. 谭, T. 维冈德  
**视频编码标准编码效率比较——包括高效视频编码 (HEVC)**  
IEEE 视频技术电路和系统交易, 22 (12) (2012), p.1669-1684

- [16] H. 施瓦茨, D. 马普, T. 维冈德  
**H.264/AVC标准的可扩展视频编码扩展概述**  
IEEE Transactions on Circuits and Systems for Video Technology, 17 (9) (2007), p.1103-1120  
[CrossRef](#) [在Scopus中查看记录](#) [谷歌学术](#)
- [17] P.Merkle, A. 斯莫利奇, K. Müller, T. 维冈德  
**多视图视频编码高效预测结构**  
IEEE Transactions on Circuits and Systems for Video Technology, 17 (11) (2007), pp.1461-1473  
[CrossRef](#) [在Scopus中查看记录](#) [谷歌学术](#)
- [18] A.维特罗, T. 维冈德, G.J.沙利文  
**H.264/MPEG-4 AVC标准的立体声和多视图视频编码扩展概述**  
IEEE会议记录, 99 (4) (2011年), p.626-642  
[CrossRef](#) [在Scopus中查看记录](#) [谷歌学术](#)
- [19] ITU-T SG 16 WP 3和ISO/IEC JTC 1/SC 29/WG 11 ISO/IEC 23008-2的视频编码联合协作小组 (JCT-VC) 和ITU-T建议H.265, 高效视频编码 (HEVC), 2013年1月。  
[谷歌学术](#)
- [20] G.沙利文, J.-R.哦, W. 韩, T. 维冈德  
**高效视频编码 (HEVC) 标准概述**  
IEEE Transactions on Circuits Systems, Video Technology, 22 (12) (2012), pp.1648-1667  
[谷歌学术](#)
- [21] 我。金, J. Min, T. 李, W. 韩, J. 公园  
**HEVC标准中的块分区**  
IEEE 视频技术电路和系统交易, 22 (12) (2012), p.1697-1706  
[CrossRef](#) [在Scopus中查看记录](#) [谷歌学术](#)
- [22] J. 鞋底, R. 乔希, N. 阮, T. 纪, M. 卡尔切维奇, G. 克莱尔, F. 亨利, A. 杜纳斯  
**HEVC中的变换系数编码**  
IEEE 视频技术电路和系统交易, 22 (12) (2012), p.1765-1777  
[CrossRef](#) [在Scopus中查看记录](#) [谷歌学术](#)

版权所有©2014爱思唯尔有限公司。保留一切权利。





[远程访问](#)

[购物车](#)

[广告](#)

[联系和支持](#)

[条款和条件](#)

[隐私政策](#)

我们使用cookie来帮助提供和增强我们的服务，并定制内容和广告。继续即表示您同意**使用cookie**。  
版权所有?2021爱思唯尔B.V.或其许可方或贡献者。ScienceDirect ®是爱思唯尔B.V.的注册商标。  
ScienceDirect ®是爱思唯尔B.V.的注册商标。