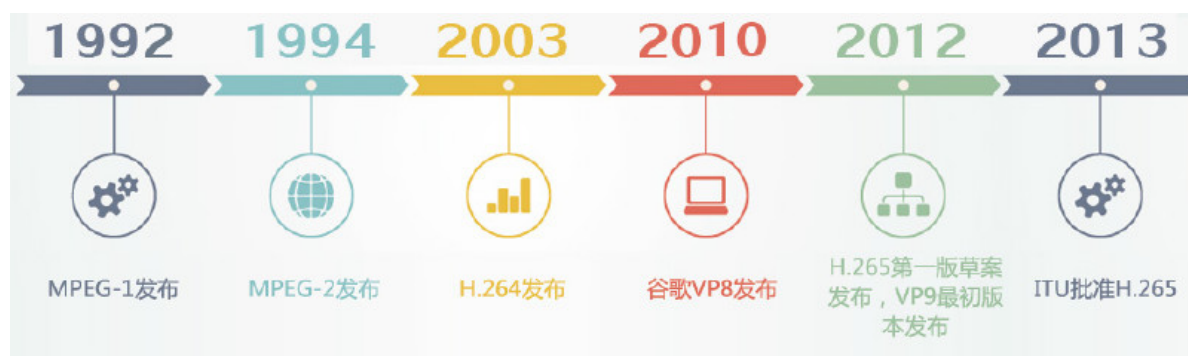


数字视频的超高清潮流奔腾向前，帧率从30 fps向60fps、120fps甚至240fps进发，与此同时，物理媒介日薄西山，内容正通过有形无形的网络在世界各个角落的终端设备上传递。高度密集的数据给带宽和存储带来巨大挑战，当前主流的H.264开始不敷应用，而新一代视频编码标准H.265似乎成为了数字4K时代的“救世主”。

H.265又称为HEVC(全称High Efficiency Video Coding，高效率视频编码，本文统称为H.265)，是ITU-T H.264/MPEG-4 AVC标准的继任者。2004年由ISO/IEC Moving Picture Experts Group(MPEG)和ITU-T Video Coding Experts Group(VCEG)作为ISO/IEC 23008-2 MPEG-H Part 2或称作ITU-T H.265开始制定。第一版的HEVC/H.265视频压缩标准在2013年4月13日被接受为国际电信联盟(ITU-T)的正式标准。

理论上H.265比H.264效率提高30-50%(尤其是在更高的分辨率情形下)，但真的只是这么简单吗？



## H.265的改变

# 基于块的视频编码技术

H.265重新利用了H.264中定义的很多概念。两者都是基于块的视频编码技术，所以它们有着相同的根源，和相近的编码方式，包括：

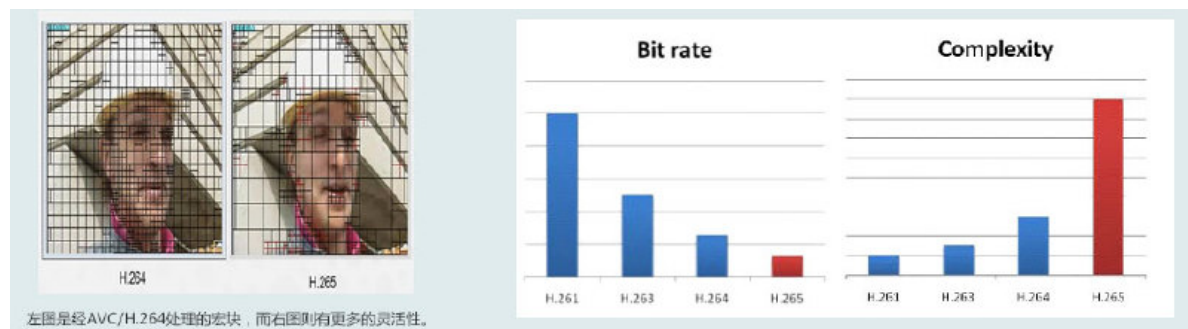
- 1、以宏块来细分图片，并最终以块来细分。
- 2、使用帧内压缩技术减少空间冗余。
- 3、使用帧内压缩技术减少时间冗余（运动估计和补偿）。
- 4、使用转换和量化来进行残留数据压缩。
- 5、使用熵编码减少残留和运动矢量传输和信号发送中的最后冗余。

事实上，视频编解码从MPEG-1诞生至今都没有根本性改进，H.265也只是H.264在一些关键性能上的更强进化以及简单化。

那么问题来了，H.265到底强在哪里？

当你考虑“只是在普通互联网上传输4K内容，还是要实现最好的图像质量”之时，就要先厘清“更多的压缩”和“更好的压缩”这两个概念。如果只是更多的压缩，4K和超高清不一定要保证比今天的1080p或HD做到更好的图片质量。更好的压缩则意味着更聪明的压缩，面对同样的原始素材，更好的压缩会以更好的方式，在不牺牲质量的情况下令数据量减少。更多的压缩很容易，而更好的压缩需要更多的思考和更好的技术，通过更智能的算法来处理图像，在维持质量的同时保持更低的比特率，这正是H.265所要做的。

如何实现更好的压缩，举例来讲，我们通常会发现在很多的图像素材里，如视像会议或者电影的很多场景中，每一帧上的大部分内容并没有改变太多，视像会议中一般只有讲话者的头在动(甚至只有嘴唇在动)，而背景一般是不动的，在这种情况下，我们的做法不是对每一帧的每一个像素编码，而是对最初的帧编码，然后仅对发生改变的部分进行编码。



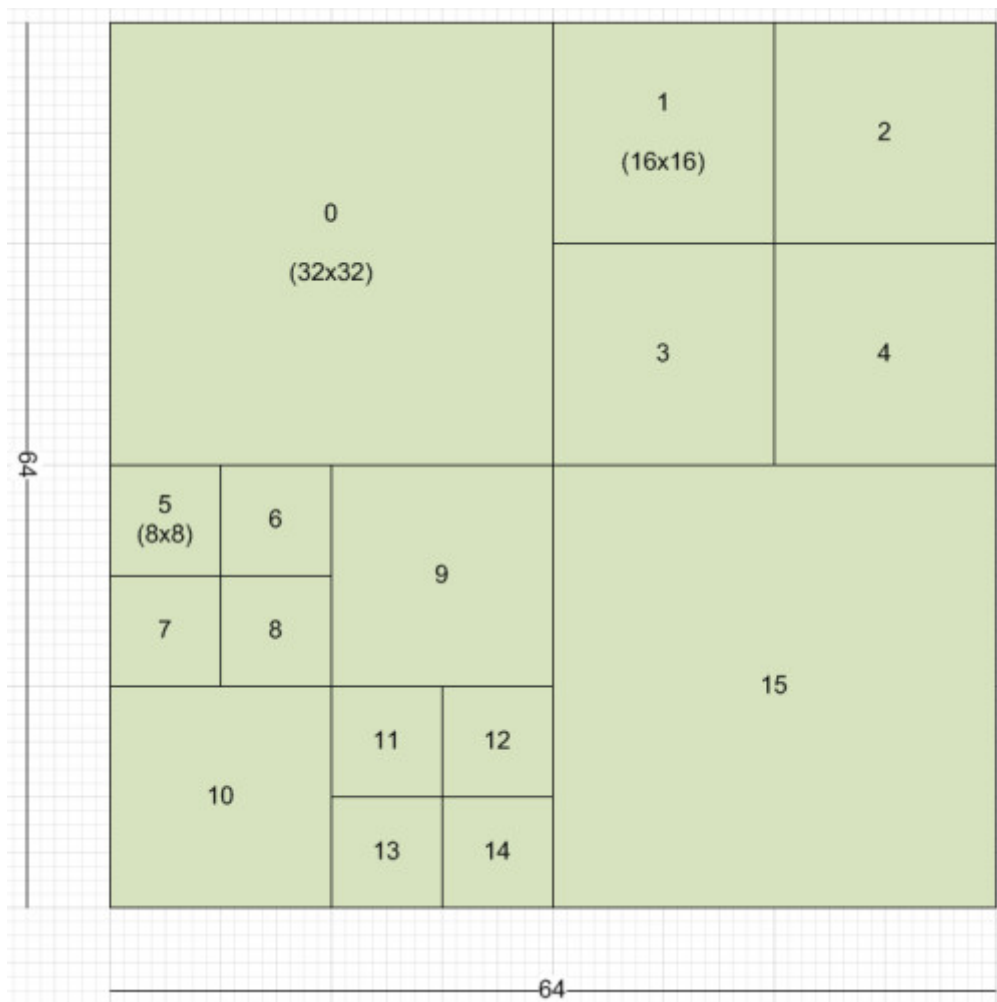
## H.265的主要改变

H.265正从以下几个方面向着“更好的压缩”迈进。

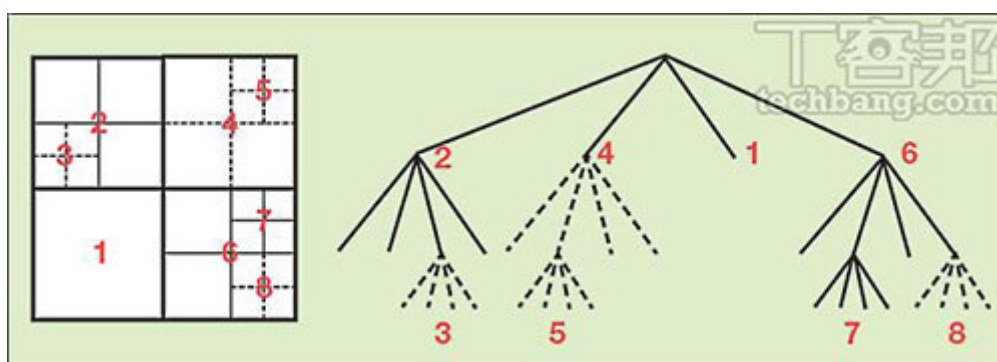
### 图像分区

H.265将图像划分为“树编码单元 (coding tree blocks, CTU) ”，而不是像H.264那样的16×16的宏块。根据不同的编码设置，树编码块的尺寸可以被设置为64×64或有限的32×32或16×16。很多研究都展示出更大的树编码块可以提供更高的压缩效率（同样也需要更高的编码速度）。每个树编码块可以被递归分割，利用四叉树结构，分割为32×32、16×16、8×8的子区域，下图就是一个64×64树编码块的分区示例。每个图像进一步被区分为特殊的树编码块组，称之为切割 (Slices) 和拼贴 (Tiles)。编码树单元是H.264的基本编码单位，如同H.264的宏块。编码树单元可向下分区编码单元 (Coding Unit, CU)、预测单元 (Prediction Unit, PU) 及转换单元 (Transform Unit, TU)。

每个编码树单元内包含1个亮度与2个色度编码树块，以及记录额外信息的语法元素。一般来说影片大多是以YUV 4:2:0色彩采样进行压缩，因此以16 × 16的编码树单元为例，其中会包含1个16 × 16的亮度编码树区块，以及2个8 × 8的色度编码树区块。



编码单元是H.265基本的预测单元。通常，较小的编码单元被用在细节区域（例如边界等），而较大的编码单元被用在可预测的平面区域。



### \*转换尺寸\*

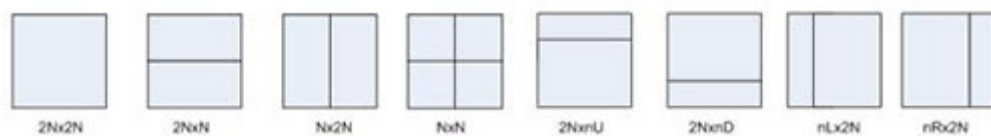
每个编码单元可以四叉树的方式递归分割为转换单元。与H.264主要以4x4转换，偶尔以8x8转换所不同的是，H.265有若干种转换尺寸：32x32、16x16、8x8和4x4。从数学的角度来看，更大的转换单元可以更好地编码静态信号，而更小的转换单元可以更好地编码更小的“脉冲”信号。



### **\*预测单元\***

在转换和量化之前，首先是预测阶段（包括帧内预测和帧间预测）。

一个编码单元可以使用以下八种预测模式中的一种进行预测。



即使一个编码单元包含一个、两个或四个预测单元，也可以使用专门的帧间或帧内预测技术对其进行预测，此外内编码的编码单元只能使用 $2N \times 2N$ 或 $N \times N$ 的平方划分。间编码的编码单元可以使用平方和非对称的方式划分。

**帧内预测：**HEVC有35个不同的帧内预测模式（包括9个AVC里已有的），包括DC模式、平面（Planar）模式和33个方向的模式。帧内预测可以遵循变换单元的分割树，所以预测模式可以应用于 $4 \times 4$ 、 $8 \times 8$ 、 $16 \times 16$ 和 $32 \times 32$ 的变换单元。

**帧间预测：**针对运动向量预测，H.265有两个参考表：L0和L1。每一个都拥有16个参照项，但是唯一图片的最大数量是8。H.265运动估计要比H.264更加复杂。它使用列表索引，有两个主要的预测模式：合并和高级运动向量（Merge and Advanced MV.）。

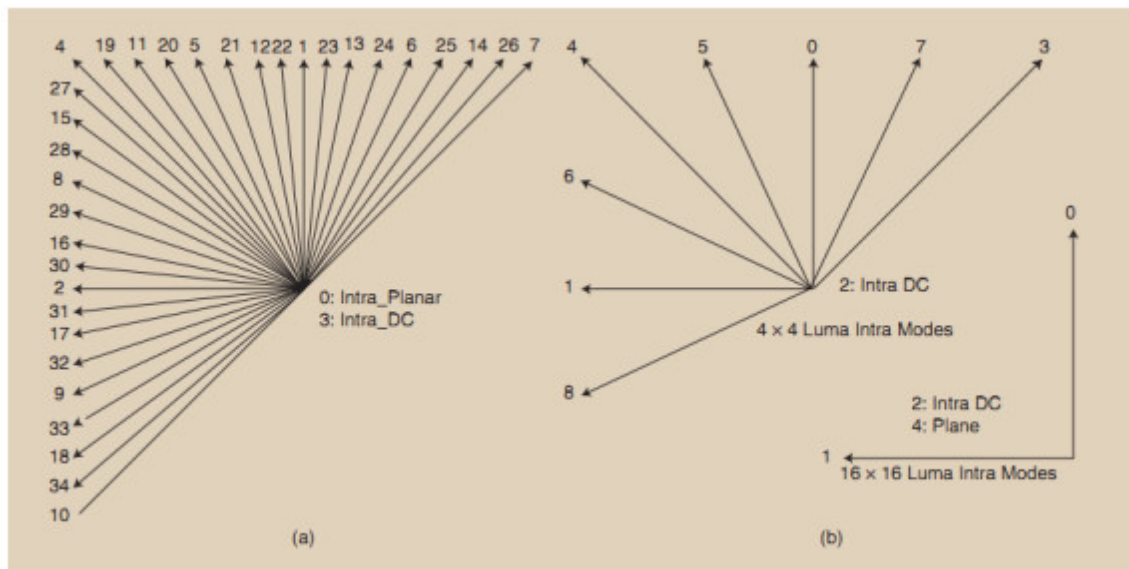


FIGURE 8. Luma intraprediction modes of (a) HEVC and (b) H.264/AVC.

在编码的过程，预测单元是进行预测的基本单元，变换单元是进行变换和量化的基本单元。这三个单元的分，使得变换、预测和编码各个处理环节更加灵活，

### \*去块化\*

与H.264在4×4块上实现去块化所不同的是，HEVC的只能在8×8网格上实现去块。这就能允许去块的并行处理（没有滤波器重叠）。首先去块的是画面里的所有垂直边缘，紧接着是所有水平边缘。与H.264采用一样的滤波器。

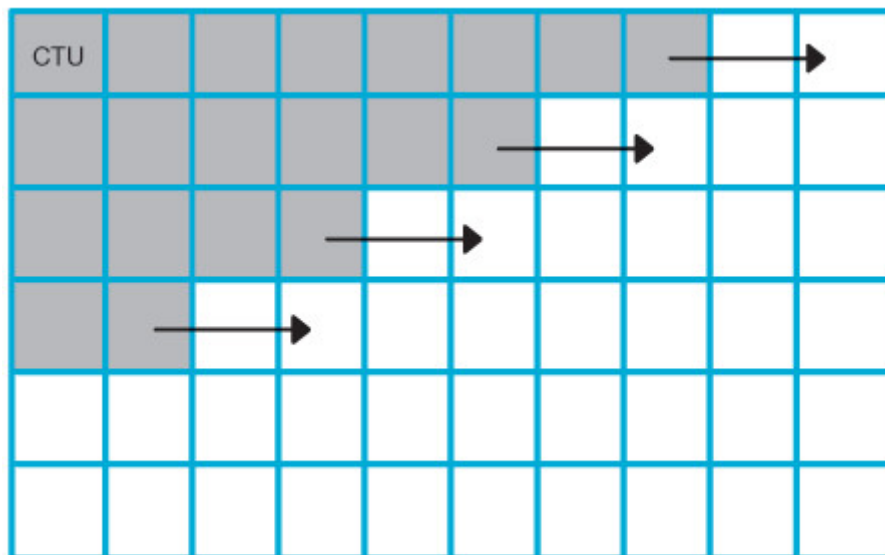
采样点自适应偏移（Sample Adaptive Offset）

去块之后还有第二个可选的滤波器，叫做采样点自适应偏移。它类似于去块滤波器，应用在预测循环里，结果存储在参考帧列表里。这个滤波器的目标是修订错误预测、编码漂移等，并应用自适应进行偏移。

### 并行处理

由于HEVC的解码要比AVC复杂很多，所以一些技术已经允许实现并行解码。最重要的为拼贴和波前（Tiles and Wavefront）。图像被分成树编码单元的矩形网格（Tiles）。当前芯片架构已经从单核性能逐渐往多核并行方向发展，因此为了适应并行化程度非常高的芯片实现，H.265 引入了很多并行运算的优化思路。





总而言之，HEVC将传统基于块的视频编码模式推向更高的效率水平，总结一下就是：

- 可变量的尺寸转换（从4×4 到32×32）
- 四叉树结构的预测区域（从64×64到4×4）
- 基于候选清单的运动向量预测。
- 多种帧内预测模式。
- 更精准的运动补偿滤波器。
- 优化的去块、采样点自适应偏移滤波器等。

## KEY CODING FEATURE COMPARISON



Tool	AVC	HEVC
Basic Coding Unit	16x16 Macroblock (MB) 16x32 "Super" MB for Interlaced Coding	8x8, 16x16, 32x32, 64x64
Temporal (Inter) Prediction	Square, Symmetric Rectangular	Square, Symmetric and Asymmetric Rectangular
Spatial (Intra) Prediction	9+4 maximum modes	33+2 maximum modes
Transform Size	4x4, 8x8, two-stage 16x16	4x4, 8x8, 16x16, 32x32 plus non-square versions
Transform Type	DCT	DCT or DST
In-Loop Filtering	Deblocking	Deblocking, SAO
Entropy	CABAC or CAVLC	

关键编码特征比

依马狮传媒  
www.imaschina.com

较

## · H.265所面临的挑战 ·

与之前从H.261到H.264的其他标准相比，H.265的显著改善不仅表现在帧间压缩领域，还表现在帧内压缩方面。由于可变量的尺寸转换，H.265在块压缩方面有很大的改善，但是增加压缩效率的同时也带来了一些新挑战。

视频编码是一个复杂的问题，对于内容的依赖性很高。众所周知，有静态背景的和高亮的低动态场景可以比高动态、黑场的图片进行更多的压缩。所以对于像H.264这样的现代化编解码器来说首要解决的是最困难的场景/情境。例如，有细节的关键帧、高动态的“勾边（crisp）”图像、黑暗区域的慢动态、噪声/纹理等。

H.265在帧内编码方面效率更高，所以细节区域可以被编码得更好，在平滑区域和渐变区域也是如此。与H.264相比，H.265的运动估计和压缩更有效，而且在伪影出现前可以在更低的比特率上操作。好消息是，H.265产生的伪影更加“平滑”，质量的降低也非常协调，即便对非常激进的分辨率/比特率编码时，也观感良好。

然而，正如硬币的两面，当处理黑暗区域的慢动态和噪声/纹理两种问题时，H.265的优势也会变成弱势。黑暗区域和噪声/纹理要求更精确的高频保留和更小的色阶变化。这通常被称之为编码的心理优化。

由于H.264使用小的转换，可以轻松将量化误差变成特征/细节，虽然与原始内容不同，但是感觉上“近似”。接近原生频率范围的误差生成可以通过小的边界转换来阻止，因此也更加可控。而更大转换的H.265要使用这种方式则会更加复杂。

H.265编码视频的存储依然是个问题，即使蓝光光盘协会正在寻求一个能够在蓝光光盘上存储4K视频的解决方案。只有至少达到100GB容量的光碟才能存储H.264编码的蓝光4K电影。而另一方面，即使H.265编码和芯片部件已经准备就绪，但是仍然缺少支持4K内容的存储和重放解决方案，并且能够兼容现有的蓝光标准。这也是H.265发展中的一个主要挑战。

## · 那么问题又来了，今天，我们应该如何对待H.264和H.265呢？ ·

当你考虑“只是在普通互联网上传输4K内容，还是要实现最好的图像质量”之时，就要先厘清“更多的压缩”和“更好的压缩”这两个概念。如果只是更多的压缩，4K和超高清不一定要保证比今天的1080p或HD做到更好的图片质量。更好的压缩则意味着更聪明的压缩，面对同样的原始素材，更好的压缩会以更好的方式，在不牺牲质量的情况下令数据量减少。更多的压缩很容易，而更好的压缩需要更多的思考和更好的技术，通过更智能的算法来处理图像，在维持质量的同时保持更低的比特率，这正是H.265所要做的。

如何实现更好的压缩，举例来讲，我们通常会发现在很多的图像素材里，如视像会议或者电影的很多场景中，每一帧上的大部分内容并没有改变太多，视像会议中一般只有讲话者的头在动(甚至只有嘴唇在动)，而背景一般是不动的，在这种情况下，我们的做法不是对每一帧的每一个像素编码，而是对最初的帧编码，然后仅对发生改变的部分进行编码。

## · H.265 PK VP9 ·

在H.265大步向前的同时，谷歌VP8的继任者VP9也已推出，同样在VP8的基础上号称编码效率提高50%，支持8K内容。VP9是一个开源和免费的规格，是WebM架构的一部分。谷歌已经在Chrome浏览器和YouTube中整合支持VP9。

与H.265在表面上类似，它同样可以抓取64×64个超级块。但265不同的是，它不一定是平方形式的，所以它可以以64×32或4×8的块来采样，实现更大的效益。但另一方面，它只有10个预测模式来重建它们。

两者都很大程度上简化了现有这些格式，尽管实现了相近的文件尺寸，有初步的报告认为，H.265有更高的图像质量，而VP9对于流媒体来说更加可靠。H.265更大的预测模型实现了边缘可视化，而VP9实施更严格的编码规则，似乎可以让流媒体更加连贯和可靠。

H.265与VP9的比较有一点类似于HDMI与DisplayPort的比较。后者以版权免费的方式去争取一席之地，但是前者的无处不在的应用意味着它会有更广泛的行业支持。这也是之前H.264轻松打败VP8的原因。

与此同时，第三个压缩格式也在规划之中，Xiph.Org基金会开发了“Daala”，虽然它还比较遥远，但是Xiph称其将是性能超越H.265和VP9的新一代规格。

## —— H.265的未来 ——

高像素数量导致需要更复杂的编解码器来最小化带宽需求。持续连接PC或TV，平滑处理4K信号的最小码流是20Mbit/s，例如Netflix要求用户的互联网连接至少提供持续的25Mbit/s带宽量。20到25Mbit/s代表带宽的巨大改善，原生的、非压缩的4K视频需要在60Mbit/s的带宽上才会有好的表现。

对于大多数的行业应用来说，H.265就是解决这一问题的答案之一，但是也要付出一定代价：显著增加的算法复杂性据说需要10倍目前2K部署所用H.264编解码器的计算能力来支撑，而提供这种能力所需的硅也远非一个简单的商品条目。

很多制造商希望在上游芯片和IC技术供应商的努力之下，解决成本和功能不平衡的问题，让H.265快速取代H.264。就目前来看，H.265在广电领域已经有比较好的发展，但是否也会成为专业应用领域的主流规范还存有疑问。因为安防监控领域等专业领域不仅受制于上述挑战，而且还要看终端用户。对于项目化的专业用户和需要监控的一般消费者而言，平安城市、交通检测和银行监控这类专业用户需要更加稳定和可靠的系统。他们中大多数已经在使用现有的技术，对于是否采用H.265还心存犹豫，这就需要更长的验证周期。

另一方面，中小企业和家庭、商店用户等消费者需要低安装成本，因此更加倾向于采用新技术。基于这个原因，H.265可能首先在中小企业应用中获得成功，并在消费市场获得认可。如果H.265标准快速成熟，其压缩效率比H.264提升50%，它就能够节省20%的投资，保证更高的性能和更替的网络和系统建设成本。