



视频编码新技术和新方向

郭勐¹, 聂秀英², 黄更生¹

(1. 中国移动通信有限公司研究院, 北京 100053; 2. 中国信息通信研究院, 北京 100191)

摘要: 以 HEVC 和 AVS2 为代表的新一代视频编码标准, 仍主要采用基于预测/变换的编码框架, 但在编码单元、帧内预测、帧间预测和去块效应滤波等环节均引入了新的编码技术。对这些视频编码新技术进行了总结, 同时分析和探讨了场景视频编码、云编码以及深度学习视频编码等视频技术发展新方向。

关键词: 视频编码; HEVC; AVS2

中图分类号: TN919

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2017246

Recent development and prospective of video coding

GUO Meng¹, NIE Xiuying², HUANG Gengsheng¹

1. China Mobile Research Institute, Beijing 100053, China

2. China Academy of Information and Communications Technology, Beijing 100191, China

Abstract: The next generation video coding standard HEVC(high efficiency video coding) and AVS2 have been developed. The key technologies of video coding were surveyed, specifically including coding unit, intra-prediction, inter-prediction and de-blocking filter. Besides that, the new development directions were analyzed and discussed, such as scene video coding, cloud coding, deep learning video coding, and so on.

Key words: video coding, HEVC, AVS2

1 引言

近年来, 随着网络带宽的提高、智能终端和互联网音视频应用数量的迅猛增长, 视频流量已经成为网络传输的主要流量, 并且仍在移动网络中持续大幅增长。根据思科视觉网络指数 (visual networking index, VNI) 《全球移动数据流量预测 (2016—2021)》报告, 预计 2016—2021 年, 移动视频将增长 8.7 倍, 在移动应用类别中享有最高的增长率; 预计到 2021 年, 移

动视频流量将达到 457 EB, 占总移动流量的 78%。持续增长的视频流量不仅对网络传输带来了巨大挑战, 也带来了海量存储的要求。为了能够传输和存储海量的视频数据, 视频压缩编码作为音视频应用的基础技术在快速发展, 以 HEVC/H.265 为代表的新一代视频编码技术已经完成了标准化^[1-4], 并逐步应用在智能手机、监控摄像机等设备中。本文将对视频编码的新技术进行介绍, 并从应用和技术的角度分析视频编码发展的新方向。

2 视频编码概述

视频编码是利用图像统计特性,去除视频信号中存在的冗余信息的过程,通过视频编码可以实现原始视频数据量的大幅减少。图1截取了视频序列中相邻的两帧。把图像划分为像素块后,可以看到在空间相邻的区域中、时间相邻的区域中,图像像素值相似,通过帧内预测、帧间预测技术可以去除空间冗余和时间冗余。同时,人眼对图像中的边缘、轮廓等信息更为敏感,通过时域到频域的变换,可以把高频的感知冗余做进一步的去除。

经过几十年的技术发展,视频编码基本采用基于块的运动补偿预测、残差变换量化和熵编码方法,形成了统一的技术框架,如图2所示。首先选择帧内编码或者帧间编码;对帧内编码的块进行空间预测,对帧间编码的块进行运动补偿预测;对预测后的残差再做变换和量化,然后进行熵编码,最终得到编码码流。为了减少解码端预测误差积累,在编码器中还包含了一个解码器生成用于帧内或者帧间预测的参考图像。

在视频编码标准方面,ITU-T VCEG 工作组

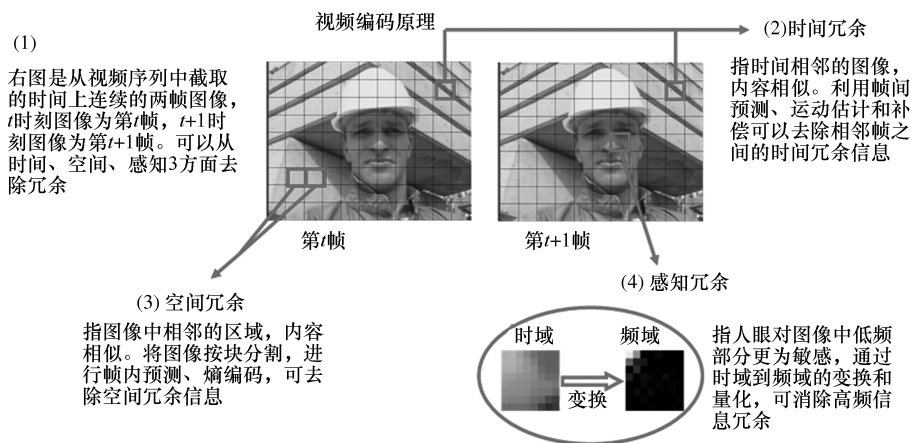


图1 视频冗余信息去除

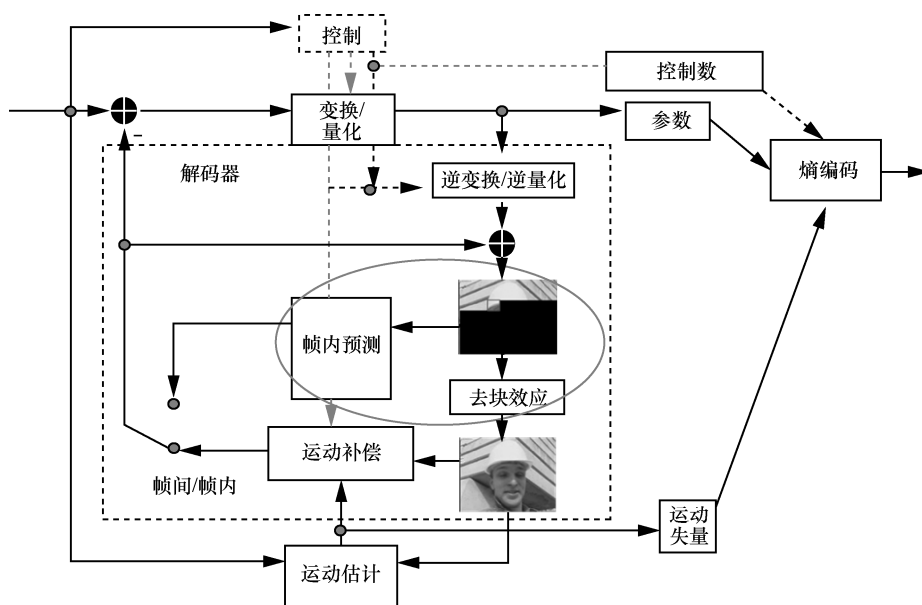


图2 视频编码框架



和 ISO/IEC MPEG 工作组已经制定了 H.26x 系列标准和 MPEG 系列标准, 为视频通话、IPTV 等各类应用的快速普及提供了统一的技术标准。目前, 新一代视频编码 HEVC (high efficiency video coding) 已由 ITU-T VCEG 工和 ISO/IEC MPEG 的联合工作组完成了标准制定, 在 2013 年的 4 月正式发布。HEVC 仍沿用了图 2 的编码框架, 但在帧内预测、帧间预测等编码环节, 引入了更高效的编码算法, 能够更好地去除视频信息冗余。HEVC 主要采用的编码新技术包括灵活的图像块划分、多角度帧内预测、帧内预测子像素插值和样本自适应补偿等, 较之前的 H.264 编码算法, 在获得相同视频质量的前提下, HEVC 编码视频流的码率节省了 50%^[5-9]。在国内, AVS 工作组完成了 AVS2 视频编码标准。AVS2 沿用图 2 的编码框架, 同样引入了灵活的编码块结构、多角度帧内预测和帧间对称预测等技术, 获得了和 HEVC 相近的编码性能^[3,4]。

3 视频编码新技术

在新一代视频编码标准中, 主要引入的新技术有新型编码单元、多角度帧内预测、帧间预测子像素插值和样本自适应补偿等。

3.1 新型编码单元

与以往视频编码采用长度、宽度均为 16 像素的图像块 (称为 16×16 宏块 (macroblock, MB)) 作为基本编码单元不同, HEVC 和 AVS2 均引入了灵活的编码树单元 (coding tree unit, CTU)。CTU 可以是 16×16 像素、32×32 像素或者 64×64 像素,

采用灵活和较大的像素块能够提供高分辨率视频的编码效率。CTU 进一步划分为编码单元 (coding unit, CU), CU 又划分为多个预测单元 (prediction unit, PU) 和多个变换单元 (transform unit, TU)。

PU 是进行预测运算的基本单元, 一个 CU 可以划分为一个或多个 PU, PU 不能超过一个 CU 的大小。图 3 示意了一个 $2N \times 2N$ 的 CU 可划分为 PU 的 8 种方式。HEVC 标准规定帧间预测可以在图 3 所示的 8 种方式里面任意选择, 帧内预测只能选择 $2N \times 2N$ 或者 $N \times N$ 的方式^[9]。

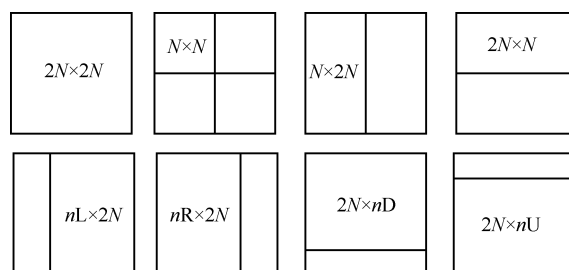


图 3 预测单元划分方式

TU 是进行变换和量化的基本单元, 每个 TU 可以包含多个 PU, 一个 PU 也可以包含多个 TU, TU 不能超过一个 CU 的大小。为了清晰地描述 CTU 和 CU、TU 之间的关系, HEVC 采用了四叉树结构。图 4 为 CU、TU 的四叉树结构关系, 其中实线为 CU 的四叉树分割, 虚线为 TU 的四叉树分割, 编号为各 CU 的编码顺序^[9]。

图 5 对比了 HEVC 和 H.264 在编码单元划分上的不同。HEVC 采用可变的块大小能够更好地表征出图像的纹理和结构特点, CU、PU 和 TU 的分离也使得视频编码更加灵活, 有助于提高视频编码效率。

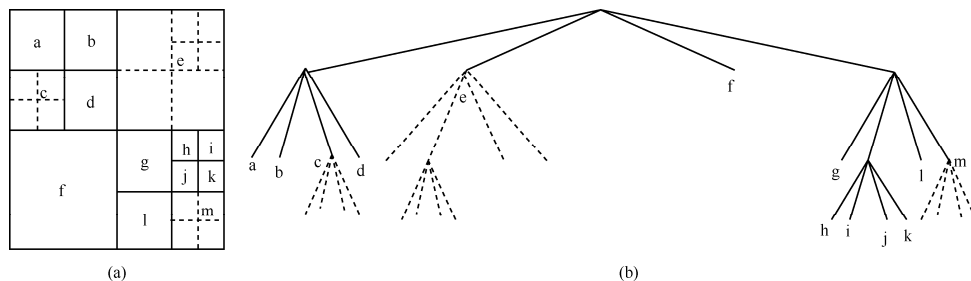


图 4 编码单元划分和四叉树组织

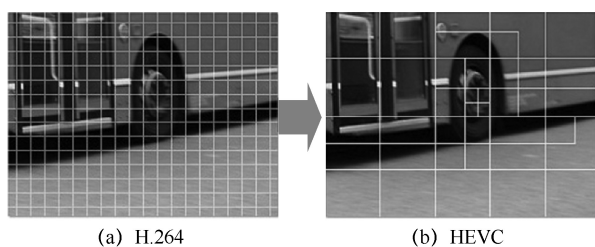


图5 编码单元对比

3.2 多角度帧内预测

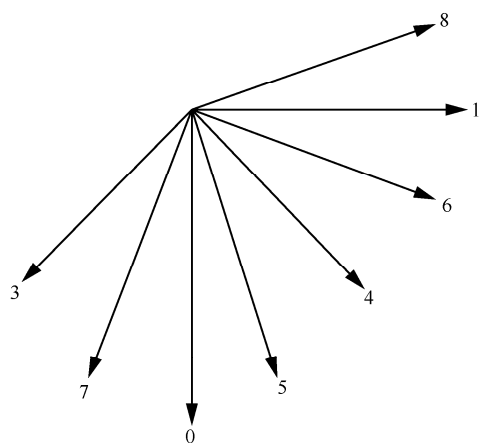
帧内预测利用图像相邻像素之间的相关性,为需要编码的图像块生成预测像素值,预测得到像素值和图像真实像素值的差异很小,也就是大部分预测误差基本为0。在编码端只编码传输预测误差,解码端一方面解码预测误差,一方面生成预测像素值,将二者加和就恢复出了被编码的图像块。因此,帧内预测的关键是能够尽可能地按照图像块的纹理、结构特点,生成和真实像素误差最小的预测像素。HEVC 编码算法较 H.264 等以往编码算法,在预测的角度上有更细的划分,从8种增加到33种。帧内预测更细的角度划分,能够更好地表达图像中的纹理信息和结构信息,从而获得准确的预测值。图6对比了H.264的8种预测方向和HEVC的33种帧内预测方向,HEVC预测方向角度由当前PU的最下面一行或最右

边一列像素与上面参考像素行或者左边参考像素列之间的位移确定,可以看到预测角度分别为 $[-32, -26, -21, -17, -13, -9, -5, -2, 0, 2, 5, 9, 13, 17, 21, 26, 32]/32$ 。同时,HEVC还定义了平面预测(intra_planar)和直流预测(intra_DC)。平面预测使用水平和垂直2个方向的线性插值平均作为当前像素的预测值,适用于纹理平滑、变化趋势一致的图像块。直流预测对当前PU的所有像素使用参考像素的平均值作为预测值,适用于平坦的图像块^[9]。

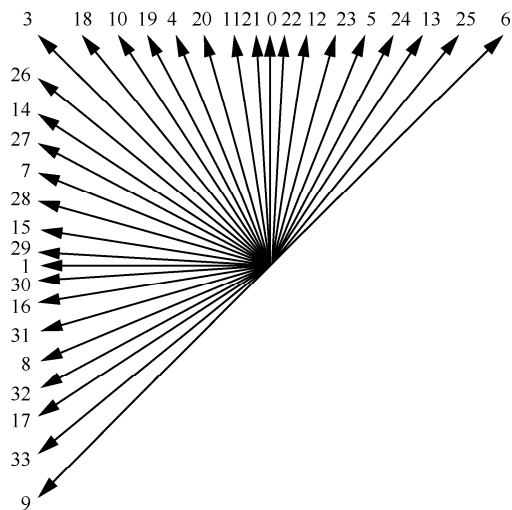
3.3 帧间预测技术

帧间预测利用视频在时间上相邻图像帧的时域相关性,使用已经编码的图像重建帧作为参考帧,通过运动估计和运动补偿对当前编码帧进行预测,计算得到和真实像素值尽可能一致的预测像素值,从而去除时间冗余信息。HEVC 帧间预测编码整体上和 H.264 相似,主要是在子像素插值和运动矢量编码方面进行了改进。

在运动估计和运动补偿中,预测得到的运动矢量可以指向整像素点,也可以指向1/2或者1/4子像素点。对于整像素位置,可以使用参考帧像素值直接进行运动补偿;对于1/2或者1/4像素点,则需要根据整像素点进行插值运算,得到1/2或1/4像素点



(a) H.264的8种预测角度



(b) HEVC的33种预测角度

图6 帧内预测角度对比



值。在 H.264 中, 首先利用 6 抽头的滤波器对整像素点滤波并舍入得到 $1/2$ 像素点插值, 再平均整像素点和相邻的 $1/2$ 像素点值得到 $1/4$ 像素点插值。HEVC 对此进行了改进, 首先使用分离的两个滤波器, 分别进行 $1/2$ 像素点和 $1/4$ 像素点插值, 一方面简化了子像素插值过程, 另一方面避免了 H.264 计算 $1/4$ 像素点过程中的舍入操作, 提高了预测精度。同时, HEVC 采用了更长的滤波器来进行子像素插值, 对 $1/2$ 子像素采用 8 抽头滤波器, 对 $1/4$ 子像素采用 7 抽头滤波器, 能够获得更高的预测精度。

在帧间预测编码中, 每个 PU 都有对应的运动矢量, 为了进一步压缩码流, 运动矢量同样需要进行预测和编码。在 H.264 中采用了空域中相邻块的运动矢量中值作为运动矢量的预测值, 然后对运动矢量预测误差进行编码。HEVC 对运动矢量编码进行了改进, 可以把空域中相邻块的运动矢量作为预测候选运动矢量, 也可以把参考帧中对应位置的时域相邻块的运动矢量作为预测候选运动矢量。图 7 说明了在空域和时域上可以采用的多个候选运动矢量, 其中 C_0 、 C_1 为参考帧时域相邻块的候选运动矢量, A_0 、 A_1 、 B_0 、 B_1 、 B_2 为空域相邻块的候选运动矢量。通过先进运动矢量预测技术, HEVC 提高了运动矢量预测的准确性, 进一步减少了编码码率^[9]。

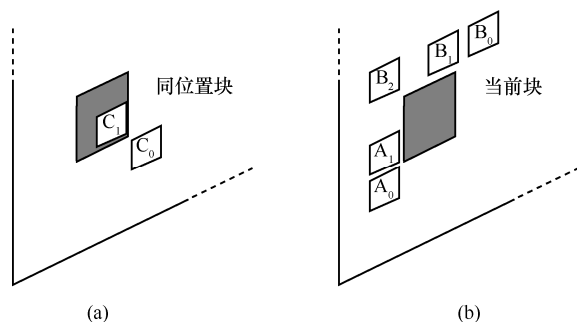


图 7 运动矢量候选

进行运动估计和补偿时, 每个 PU 都有一组运动参数, 为了进一步地压缩编码参数, HEVC

对运动参数编码引入了 Merge 模式, 也就是将运动参数相同或相近的一些 PU 合并起来形成一个区域, 为这个区域只编码和传输一次运动参数。图 8 (a) 为需要编码的图像, 其中白色箭头示意了图像中主要运动的类似钟摆的部分; 图 8 (b) 为经过运动估计的 CU 的四叉树划分, 可以看到对于具有相同或相近运动参数的静止区域, 被划分为多个 PU; 图 8 (c) 为经过 Merge 合并后可以共享运动信息的区域划分, 这些区域共用了运动信息, 减少了编码码流^[9]。

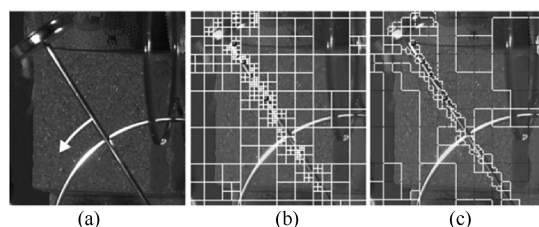


图 8 运动参数编码 Merge 模式

3.4 样本自适应补偿

由于视频编码均把像素块作为基本的编码处理单元, 在相邻的像素块边界处会出现像素值不连续的现象, 这种现象被称为块效应。为了消除块效应, 提高编码质量, H.264 等编码算法均定义了去块效应滤波 (de-blocking filter) 算法对重建图像的块边界进行平滑。HEVC 在采用去块效应滤波的同时, 另外加入了样本自适应补偿 (sample adaptive offset, SAO) 来增强图像质量。SAO 在去块效应滤波之后, 对重建像素块中的像素值进行分类, 并加上相应的偏移值, 达到降低图像失真的效果。SAO 分为边缘补偿 (edge offset, EO) 和条带补偿 (band offset, BO)。图 9 示意了边缘补偿的水平、垂直、 135° 和 45° 的 4 个方向^[9]。在选定边缘补偿的方向后, 继续根据图 10 的分类,

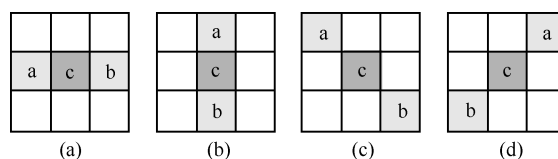


图 9 边缘补偿的 4 个方向

进行像素值的偏移, 对由于量化造成的像素点局部拐点、峰值点和谷值点加上一个适当的相反的补偿量, 用来消除编码引入的偏移^[9]。图 10 中, 类别 1 指 $c < a$ & $c < b$, 类别 2 指 $(c < a \& \& c = b) \vee (c = a \& \& c < b)$, 类别 3 指 $(c > a \& \& c = b) \vee (c = a \& \& c > b)$, 类别 4 指 $c > a \& \& c > b$, 类别 0 指以上都不是。条带补偿将图像像素值范围划分为 32 等份, 每个等份是一个“band(条带)”。条带补偿首先计算每个条带中原始像素点和重建像素点之间的平均值误差(补偿值), 并选择连续的 4 个偏移最大的条带进行补偿(把补偿值加到每个重建像素上)。图 11 中虚线表示图像原始像素值, 圆点表示图像重建像素值, 条带补偿将对 4 个条带中的重建像素值加上补偿值, 使之更接近原始图像像素值^[9]。

4 视频编码新方向

当前, HEVC 和 AVS2 新一代视频编码标准已制定完成, 其主要技术仍采用了基于预测/变换的编码框架。为了能够满足视频应用持续增长的需求, 视频编码技术还在快速发展, 出现了一些新的发展方向。

4.1 场景视频编码

虽然标准化的视频编码算法能够在各种类型的视频上均取得较好的压缩效果, 但在一些特定应用中仍有较大的提升空间。例如, 在视频监控应用中, 一般来说摄像机采集到图像的背景是保持不变的, 如果能够获取到视频监控场景的特征, 并在编码过程中使用, 将能够大幅提高压缩效率。AVS2 标准已经面向智能监控应用, 提出了场景视频编码, 使

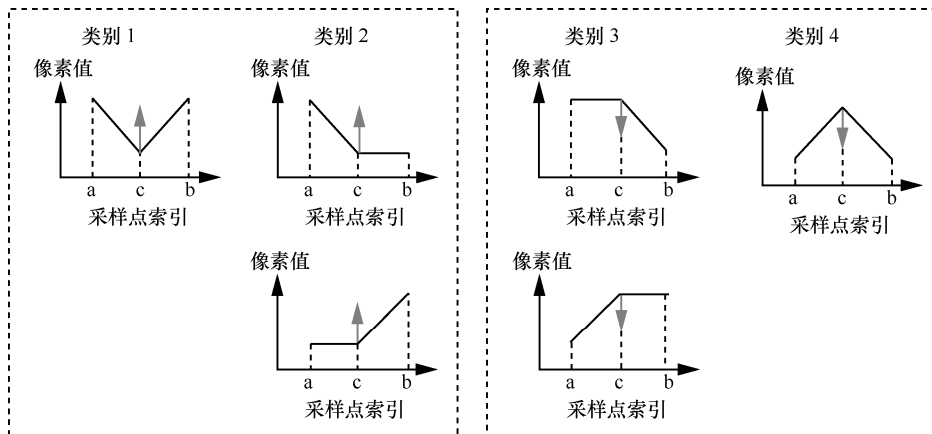


图 10 边缘补偿的分类

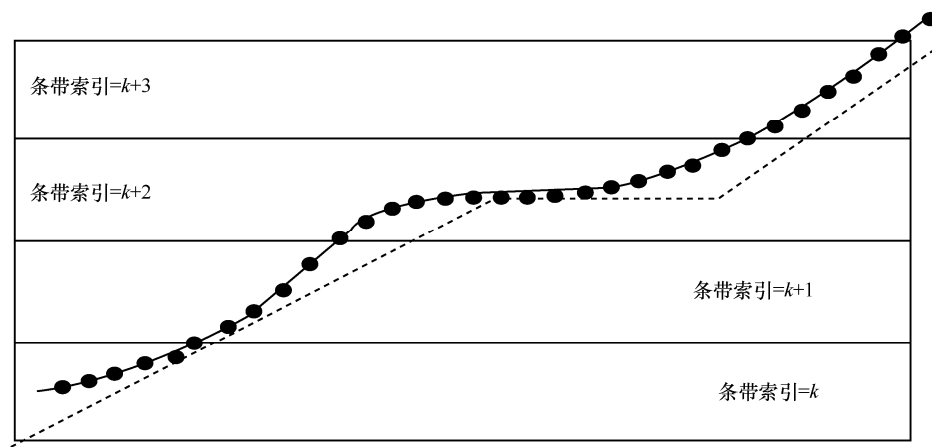


图 11 条带补偿示例



得整体压缩效率提高了大约 1 倍。AVS2 的场景视频编码引入了场景图像，在标准中为 G 图像或者 GB 图像。G 图像是只使用帧内预测解码的场景图像，应被输出；GB 图像也是只使用帧内预测解码的场景图像，但不被输出。图 12 中说明了场景图像如何作为时间上的长期参考帧，用于预测^[11]，且量化参数较小。虽然场景图像使用较小的量化系数，产生了较多的编码长度，但是场景图像编码质量会较高，能够提供更高的预测精度，降低后续编码帧的编码长度。在编码复杂度方面，场景视频编码未改变编码框架，仍采用预测、变换和量化的编码技术，保持了现有编码技术相近的时间复杂度，编解码器只增加了存储场景图像的缓冲存储^[11]。

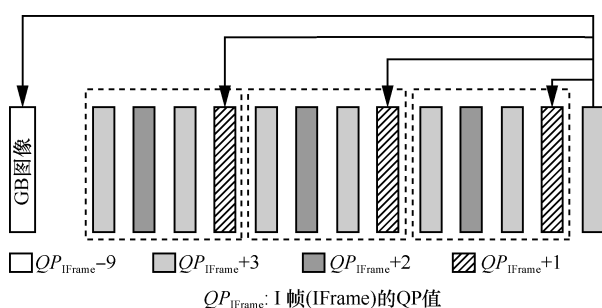


图 12 AVS2 的场景编码

目前场景编码在 H.264 或 H.265 等编码标准中尚未引入，但在实际产品中，已出现了针对

监控视频场景不变、图像静止的编码优化技术，对静态场景降低帧率、码率来实现进一步的压缩。除了视频监控应用，还有视频会议、课堂教学视频、桌面云等一系列应用，都有着视频场景不变或切换不频繁的特点，结合这些应用的特点，对场景加以分析和利用，以便获取更高压缩比，是视频编码技术发展的方向之一^[12]。

4.2 云编码

随着互联网应用的发展，在云端保留了大量的图像以及视频，这些图像和视频可以用作参考，用于预测，从而实现云编码，提高编码效率。

云编码对云端的大量图像和视频数据中提取共性的对象，形成共享的视觉词典。图 13 描述了云编码的技术特征，当新的图像或视频需要进行压缩时，首先分析图像中的典型编码对象，和云端的视觉对象进行对比，记录视觉词典的索引后，只编码有差异的部分^[12]。当前已有研究工作尝试了采用 SIFT 方法提取图像特征，在云平台进行特征匹配以实现图像压缩编码^[13]。

云编码不仅能够从已有海量图像视频数据中获取视频编码的先验知识，提高编码效率，还能够利用云端丰富的计算资源，降低终端编码复杂度，是值得深入研究的视频编码方向之一。

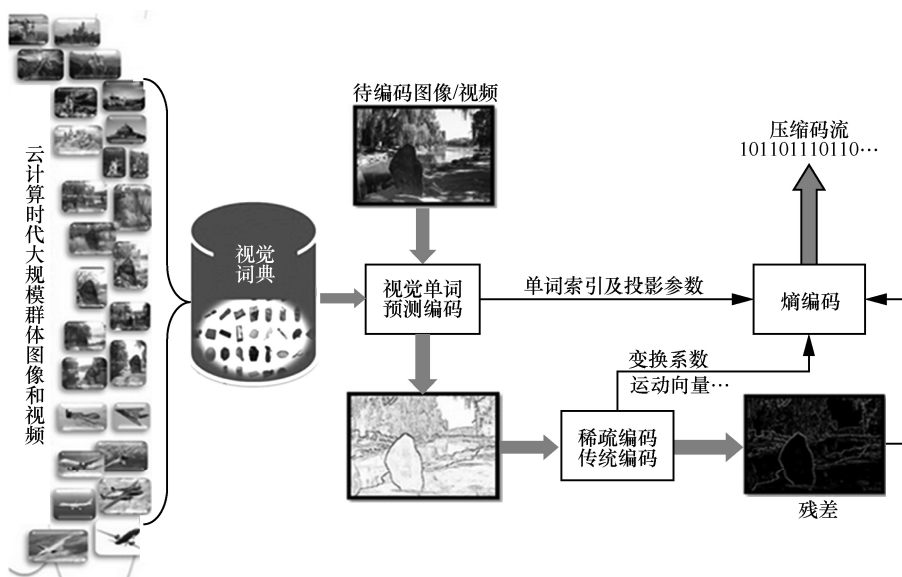


图 13 云编码技术特征

4.3 深度学习视频编码

随着计算能力和数据量的提升,深度学习成为当前人工智能研究的热点之一。通过大批量数据训练,深度网络模型能够自动发现和提取图像特征,已普遍应用在人脸识别、图像分割、目标检测等任务上,并取得了优于传统手工特征方法的检测和识别性能。

在图像编码领域,Google和MIT的研究人员也开始研究采用深度学习进行图像和视频压缩编码^[14,15]。Google在2016年提出了一种基于神经网络的有损图像压缩方法,采用了基于递归神经网络(RNN)的编码和解码器以及二值化和熵编码神经网络,在柯达图像集合上取得了优于JPEG算法的压缩性能^[14]。MIT提出了生成压缩的概念,采用生成网络(generative network)进行编码压缩。图14对比了生成压缩和JPEG2000的效果,MIT在论文中特别提出和传统图像视频压缩算法依赖手工设计编码器相比,生成神经网络具有自动特征提取的优势,将是视频编码的重要发展方向^[15]。

4.4 立体视频编码

为满足3D电视、全景视频的需求,H.264等视频编码标准制定了多视点视频编码的算法。近年来,随着VR的发展,立体视频编码仍是视频编码的研究热点之一^[16],HEVC也开展了面向多视点和深度信息的编码研究。立体视频编码时,不仅需要去除某一视点内的空间和时间冗余,还需要去除不同视点间的空间和时间冗余,同时还需要对深度图进行压缩编码。相比对每个视点视频独立编码传输,立体视频编码能够显著降低码率,随着VR设备和应用的增长,立体视频编码也将逐步得到应用。

5 结束语

HEVC和AVS2是当前最新制定的视频编码标准,在互联网、物联网和可穿戴等领域都有广泛应用^[17-19]。HEVC和AVS2虽然仍采用了基于预测/变换的编码框架,但引入了很多先进的编码技术。在编码单元上,可以是 16×16 像素、 32×32 像素或者 64×64 像素,并能够进一步灵活划分为预测单元和变换单元;在帧内预测环节,可以有33种预测方向;在帧间预测环节,采用了更高精度的子像素插值和高效的运动矢量编码;在环路滤波环节,新增加了样本自适应补偿。本文首先介绍了视频编码的通用框架,随后总结了这些编码新技术。最后,本文对视频编码技术进行了展望,提出了场景视频编码、云编码、深度学习视频编码等可能的发展新方向。

参考文献:

- [1] 沈燕飞,李锦涛,朱珍民,等. 高效视频编码[J]. 计算机学报, 2013, 36(11): 2340-2355.
SHEN Y F, LI J T, ZHU Z M, et al. High efficient video coding[J]. Chinese Journal of Computers, 2013, 36(11): 2340-2355.
- [2] SULLIVAN G J, OHM J R, HAN W J, et al. Overview of the high efficiency video coding (HEVC) standard[J]. Transactions on Circuits and Systems for Video Technology, 2012, 22(11): 1649-1688.
- [3] 马思伟. AVS 视频编码标准技术回顾及最新进展[J]. 计算机研究与发展, 2015, 52(1): 27-37.
MA S W. History and recent developments of AVS video coding standards[J]. Journal of Computer Research and Development, 2015, 52(1): 27-37.
- [4] WANG S S, LUO F L, MA S W. Overview of the second generation AVS video coding standard(AVS2)[J]. ZTE Communications, 2016, 1(1): 3-11.
- [5] 朱秀昌,李欣,陈杰. 新一代视频编码标准——HEVC[J]. 南京邮电大学学报(自然科学版), 2013, 33(3): 1-11.
ZHU X C, LI X, CHEN J. Next generation video coding stan-

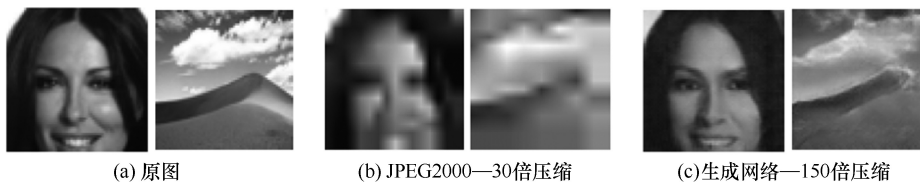


图14 生成压缩效果对比



- ard—HEVC[J]. Journal of Nanjing University of Posts and Telecommunications(Natural Science), 2013, 33(3): 1-11.
- [6] 周作成, 贾克斌. 新一代视频编码标准 HEVC 关键技术及其应用[J]. 北京工业大学学报, 2014, 40(10): 1489-1496.
- ZHOU Z C, JIA K B. Key technologies and applications of next generation video coding standard HEVC[J]. Journal of Beijing University of Technology, 2014, 40(10): 1489-1496.
- [7] 赵耀, 黄晗, 林春雨, 等. 新一代视频编码标准 HEVC 的关键技术[J]. 数据采集与处理, 2014, 29(1): 1-10.
- ZHAO Y, HUANG H, LIN C Y, et al. Key techniques in new video coding standard HEVC[J]. Journal of Data Acquisition and Processing, 2014, 29(1): 1-10.
- [8] 朱秀昌, 刘峰, 胡栋. H.265/HEVC 视频编码新标准及其扩展[M]. 北京: 电子工业出版社, 2016.
- ZHU X C, LIU F, HU D. H.265/HEVC video coding standard[M]. Beijing: Publishing House of Electronics Industry, 2016.
- [9] SZE V, BUDAGAVI M, SULLIVAN G J. High efficiency video coding (HEVC) algorithms and architectures[M]. Heidelberg: Springer, 2014.
- [10] YANG J Y, DONG S W, TIAN Y H, et al. Introductions to AVS2 scene video coding techniques[J]. ZTE Communications, 2016(1): 50-53.
- [11] ZHANG X G, LIANG L H, HUANG Q, et al. An efficient coding scheme for surveillance videos captured by stationary cameras[C]//2010 SPIE Conference on Visual Communications and Image Processing, July 11, 2010, Huangshan, China. Bellingham: SPIE Press, 2010: 77442A-1-10.
- [12] 黄铁军. AVS2 标准及未来展望[J]. 电视技术, 2014, 38(22): 7-10.
- HUANG T J. AVS2 standard and prospective[J]. Video Engineering, 2014, 38(22): 7-10.
- [13] YUE H J, SUN X Y, WU F, et al. SIFT-based image compression[C]//2012 IEEE International Conference on Multimedia and Expo, July 9-13, 2012, Los Alamitos, USA. New Jersey: IEEE Press, 2012: 473-478.
- [14] TODERICI G, VINCENT D, JOHNSTON N, et al. Full resolution image compression with recurrent neural networks[J]. arXiv: 1608.05148v118.
- [15] SANTURKAR S, BUDDENY D, SHAVIT N. Generative compression[J]. arXiv:1703.01467v2.
- [16] 黎洪松, 程福林, 艾新宇. 视频编码技术的若干新进展[J]. 计算机应用研究, 2016, 33(8): 2246-2254.
- LI H S, CHENG F L, AI X Y. Survey on some new advances in video coding technology[J]. Application Research of Computers, 2016, 33(8): 2246-2254.
- [17] 魏峥, 施唯佳, 祝谷乔. 互联网视频中多屏互动技术的应用[J]. 电信科学, 2014, 30(5): 27-32, 39.
- WEI Z, SHI W J, ZHU G Q. Multi-screen interaction technologies on internet streaming video[J]. Telecommunications Science, 2014, 30(5): 27-32, 39.
- [18] 毛彤, 周开宇. 可穿戴设备综合分析及建议[J]. 电信科学, 2014, 30(10): 134-142.
- MAO T, ZHOU K Y. Discussion about the influence of wearable device on mobile operators' service[J]. Telecommunications Science, 2014, 30(10): 134-142.
- [19] 魏颖琪, 林玮平, 李颖, 等. 物联网智能终端技术研究[J]. 电信科学, 2015, 31(8): 146-152.
- WEI Y Q, LIN W P, LI Y, et al. Study on key technologies of intelligent IoT device[J]. Telecommunications Science, 2015, 31(8): 146-152.

[作者简介]



郭勤 (1981-), 男, 博士, 中国移动通信有限公司研究院高级工程师、项目经理, 主要研究方向为音视频编码、视频云和人工智能等。



聂秀英 (1965-), 女, 中国信息通信研究院教授级高级工程师, 担任 IEC TC1 中国专家、SAC/TC232 委员、CCSA TC1 WG3 组长等, 主要从事数据通信、多媒体通信、互联网、现代服务业等领域研究和标准制定工作。获得国家科学技术进步奖二等奖 1 项, 省部级科学技术进步奖多项。出版著作两部, 发表论文 60 多篇, 向 ITU-T 提交文稿 20 多篇。



黄更生 (1962-), 男, 博士, 中国移动通信有限公司研究院业务所所长、主任研究员, 中国移动技术咨询委员会业务领域专家委员。