

H.264/MPEG-4 AVC 视频编码器性能评估

刘党辉^{1,2} 沈兰荪² 强秀丽¹

¹(装备指挥技术学院,北京 101416)

²(北京工业大学信号与信息处理研究室,北京 100022)

E-mail:liudanghui@emails.bjpu.edu.cn

摘要 H.264 和以往的视频标准相比,由于采用了许多新的编码技术,使得压缩效率有了很大的提高。但是,相应的计算复杂度也大大增加了。文章主要讨论这些新的编码技术对压缩效率的影响,以便根据不同的实际应用情况配置编码器,从而在特定硬件平台上达到压缩效率与计算复杂度的最优分配。

关键词 视频编码 压缩效率 计算复杂度

文章编号 1002-8331-(2006)04-0111-03 文献标识码 A 中图分类号 TP311

Performance Evaluation of H.264/MPEG-4 AVC Video Encoder

Liu Danghui^{1,2} Shen Lansun² Qiang Xiuli¹

¹(The Academy of Equipment, Command and Technology, Beijing 101416)

²(Signal and Information Processing Lab, Beijing University of Technology, Beijing 100022)

Abstract: H.264 dramatically improves the video compression efficiency by using many new coding techniques against some previous video coding standards. However, the computation burden also largely increases. The affection of the new coding techniques on compression performance is mainly discussed in this paper so that an optimal selection for the tradeoff between compression efficiency and computation burden can be decided in a real software or hardware implement of the algorithms.

Keywords: video coding, compression efficiency, computation burden

1 引言

人眼视觉信息主要来源于图像/视频。但是,图像/视频的数据量巨大,且又存在着很大的数据冗余,为提高其存储和传输效率,必须采用某种有效的压缩技术。自上世纪 90 年代开始,国际上致力于标准化的工作,先后制定了一系列图像/视频编码标准,如 JPEG、JPEG2000、MPEG-1、MPEG-2、MPEG-4、H.261 和 H.263 等,在图像/视频应用领域中获得了极大的成功。

2003 年 5 月,ITU-T 和 ISO/IEC 联合制定了 H.264/MPEG-4 part 10 AVC (Advanced Video Coding) 视频编码标准^[1]。为叙述方便,以下将其简称为 H.264 标准。H.264 标准具有高效的编码压缩和易于网络传输的特点,以满足视频电话、视频会议、视频存储、视频广播、视频流媒体、数码影院、视频监控、军事侦察等众多应用领域。H.264 标准定义了四个档次,即基本档次、主档次、扩展档次和高级档次。基本档次支持帧内(I 分片)和帧间(P 分片)编码,采用上下文自适应的可变长编码。主档次支持隔行扫描,支持帧内(I 分片)和帧间(P、B 分片)编码,采用加权预测的帧间编码和上下文自适应的算术编码,扩展档次不支持隔行视频和上下文自适应的算术编码,但是增强了不同码流切换的模式和改进的误码恢复功能。高级档次具有主档次的特征,但支持 8/10/12 比特/样值位深,增加了 8x8 整数 DCT 变换,基于感知的量化缩放矩阵以及特定区域的无损压缩等功能,比主档次具有更高的压缩效率,而计算复杂度不会有显著增加。

图 1 给出了 H.264 编解码的结构框图。可以看出,H.264 仍采用了以往视频编码标准的运动估计/补偿+分块 DCT 变换的混合编码框架。但是,和以往的视频标准(如 MPEG-4)相比,在 30dB~35dB 的解码质量范围内,H.264 标准能比 MPEG-4 标准节约 30%~50% 的码率。H.264 标准编码性能的提高,取决于其采用了很多新的编码技术,如:可变块大小的运动预测,四分之一或八分之一精度的运动估计和补偿,多参考帧运动补偿,在环去块效应滤波器,4x4 块整数 DCT 变换,上下文自适应的算术编码和可变长编码,帧/场和宏块自适应编码,等等。但是,相应的计算复杂度也大大增加了^[2-5]。

本文主要讨论这些新的编码技术对编码性能的影响,以便根据不同的应用配置编码器,从而达到编码效率与计算复杂度的最优分配。

2 H.264 编码效率的评估

下面将分别讨论 H.264 标准采用的新的编码技术对编码效率的影响。

2.1 帧内预测编码

在帧内编码中,H.264 采用了帧内预测模式,能够更好地消除图像的空间冗余,提高编码效率。帧内预测的原理主要依据图像中相邻宏块具有相似性的特点,通过已编码宏块来预测当前宏块,然后对当前宏块与预测值的差值进行变换编码。对于亮度,预测分为 4x4 子块模式、16x16 宏块模式,以及直接传

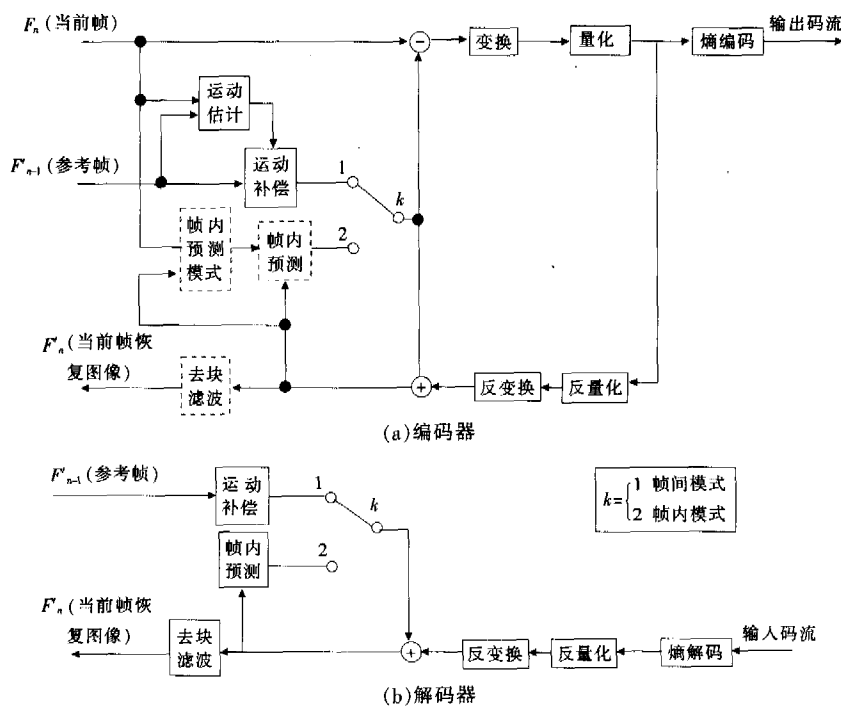


图1 H.264 视频编解码器结构框图

输图像亮度和色度的 I_PCM 帧内编码模式,其中 4×4 子块预测包含 9 种预测模式,16×16 宏块预测包含 4 种预测模式。对于色度,采用 8×8 色度预测,包含的 4 种预测模式与亮度 16×16 宏块预测的 4 种预测模式类似。

H.264 的帧内编码取得了很好的编码效率,基本上能够达到 JPEG2000 的图像质量,而且在对小图像和低码率的条件下,有时超过了 JPEG2000。这主要是因为 H.264 的帧内编码采用了先进的帧内预测模式,它能充分地利用图像的空间相关性,更好地去除了空间冗余信息,同时去块滤波器能有效地抑制块效应问题,使得解码图像具有更高的质量^[9]。

2.2 帧间预测编码

帧间预测主要是利用连续图像序列之间的相关性,通过运动补偿的预测编码方法来消除视频图像的时间冗余。H.264 除了具有原有标准如 H.263、MPEG-4 中的 P 帧、B 帧预测方法外,还增加了许多新的功能:采用不同大小的预测块进行运动估计;采用 1/4(甚至 1/8)像素精度的运动补偿算法;采用多参考帧进行帧间预测编码;去块滤波器消除由块预测误差产生的块效应。

2.2.1 多种子块的帧间预测

H.264 标准对帧间预测的每个 16×16 亮度宏块,根据不同的语法,又可分为 8×16,16×8,8×8,8×4,4×8,4×4 块进行变换编码,色度块为亮度块的 1/4。由于采用不同大小的子块进行帧间预测,使得运动估计模型更接近物体的实际运动,因此运动补偿的精度更高,这种方法比单独 16×16 块的预测方法降低约 16.4% 的码率,其中大约 13.7% 的码率是由于采用了 8×8 大小的预测块。此外,对于高分辨率视频,采用 4×4 的块编码和预测对码率的改善效果更小。

2.2.2 亚像素运动估计的帧间预测

H.264 的亮度运动估计精度可以达到 1/4、1/8 像素,与整数精度的空间预测相比,它可以节约大于 20% 的码率。色度系数采用双线性插值算法,精度可以达到 1/8、1/16 像素。

2.2.3 多参考帧模式的帧间预测

与 H.263 和 MPEG 标准不同,H.264 采用多个参考帧进行帧间预测,通过选择最佳参考帧进行运动补偿和预测过程,典型地,这比单参考帧方法可以节省 5.7% 的码率。并且有利于码流的错误恢复。图像编码顺序不是基于时间的图像显示顺序,而是基于图像之间的依赖关系的。解码器得到图像序列后,通过重排,以便按照顺序进行显示。

2.2.4 在环去块滤波

由于基于块的 DCT 变换和运动补偿算法,不可避免地会产生方块效应。为了在低码率情况下提高图像质量,H.264 标准中定义了去块滤波器。滤波器是以 16×16 像素大小的宏块为单位,对整个重建图像进行滤波。滤波器去除了预测误差产生的块效应,从而获得了更好的主观图像质量。典型地,采用去块滤波器能缩减 5%~10% 的码率,然而仍能保持与不采用滤波器时相同的图像恢复质量。

2.3 熵编码

H.264 采用两种方法进行熵编码:VLC 编码和 CABAC 编码算法。对于量化系数之外的所有语法元,采用简单的 Exp-Golomb 编码方法,而对于量化系数,则采用上下文自适应的 VLC 编码(CAVLC)。相对于采用单一的 VLC 表,编码效率得到了改善。而且,采用基于上下文自适应的二进制算术编码算法(CABAC),能够充分利用上下文信息和算术编码的优点,使得编码后的平均码长更逼近图像的信息熵,达到最佳的编码效率。相比采用 CAVLC,采用 CABAC 算法编码,可以节约 5%~15% 的码率,然而计算复杂度却几乎没有增加。

2.4 自适应帧场和自适应宏块编码

对于隔行扫描视频,由于运动对象或相机的运动,上下场中的相邻行之间的冗余特性降低,因而单独编码每一场也许更有效。因此,H.264 对每一帧可以自适应采用三种编码方式:(1)帧模式,即合并上下两场为一个编码帧;(2)场模式,即独立编码每帧的上下两场;(3)合并上下两场为一个编码帧,但是把

上下相邻的两个宏块重组成一个场对或帧宏块对。前两项的自适应选择对应的就是自适应帧场编码,第三项对应的是自适应宏块编码。比起仅仅采用帧模式,采用自适应帧场编码模式可以节约码率 16%~20%(如对于“Canoa,Rugby”等视频序列)。而比起自适应帧场编码,采用自适应宏块编码模式可以进一步节约 14%~16%的码率(如对于“Mobile and Calendar,MPEG-4 World News”等视频序列)。

2.5 率失真模型

如果在编码的过程中保持固的定编码参数,就不可能产生最优的编码率失真性能。但是,要获得压缩率和解码质量之间的最优平衡也是很困难的。在“离线”的编码方式下,一般采用“两步法”来获得近似的最优率失真性能。H.264 参考软件采用了基于 Lagrangian 的优化算法。实际上,编码器的优化实际上至少是关于码率、率失真和计算复杂度这三者之间的一个优化问题。例如,对于“Office”视频序列,比起 MPEG-4 的简单档次,采用 H.264 的基本档次,在 PSNR 为 35dB 时,节约码率 22.8%。同样条件下,对于“Carphone”视频序列,H.264 比 MPEG-4 节约码率高达 36%。这个编码增益很高,可能是由于拍摄的视频质量较高的原因。

3 H.264 计算复杂度的评估

虽然 H.264 标准能大大改善编码效率,但是计算复杂度也大大提高了。由于计算复杂度和特定编码模式的编码效率与视频内容有着极大的相关性,因此,对于特定的应用,应该选择可能的编码模式来适应编码器对视频编码效率和处理平台的要求。下面以“Violin”视频序列(QCIF 格式,每秒 25 帧)为例说明选择不同编码模式对编码性能的影响,其中采用的量化参数固定为 36,仅对其前 25 帧进行编码。对 H.264 编码器采用如下两种配置^[3]:

(1)基本配置:I/P 帧编码,各种块预测,一个参考帧,去块滤波器,CAVLC 熵编码,无率失真优化。

(2)高级配置:I/P/B 帧编码,各种块预测,5 个参考帧,去块滤波器,CABAC 熵编码,有率失真优化。

表 1 给出了编码性能的比较,可以看出亮度 PSNR 几乎是相同的,然而码率和编码时间的差异是明显的。

表 1 H.264 可选模式的编码性能的比较
(“Violin”视频序列,QCIF 格式 25 帧/s)

配置方式	平均亮度 PSNR	码率/Kbps	时间/s
基本配置	29.06	45.9	40.4
基本配置+最小 8x8 块编码预测模式	29.00	46.6	33.9
基本配置+5 个参考帧模式	29.12	46.2	157.2
基本配置+率失真优化模式	29.18	44.6	60.5
基本配置+“IBPBP...”编码模式	29.19	42.2	55.7
基本配置+CABAC 模式	29.06	44.0	40.5
高级配置	29.57	38.2	180.0
高级配置+1 个参考帧模式	29.42	38.8	77.0

“基本配置”需要 40s 来编码这个序列,码率为 46Kbps(不包括第一个 I 帧的码流)。当仅仅用 8x8 或更大的运动补偿块来缩减编码时间(大约 6s)时,编码码率增大。当用 5 个参考帧时,编码时间大约增加了到 4 倍。当采用率失真优化模式,编码时间增加了大约 50%;采用 B 帧提高了编码效率,但是编码时间也增加了大约 50%。CABAC 增加了编码效率,然而并没有增加编码时间。

“高级配置”的编码时间是“基本配置”的编码时间的四倍多,但是节约了大约 17%的码率。通过使用一个参考帧,编码时间可以大大降低。

这些结果表明:对于这个视频序列,最有效的优化是采用 CABAC 和 B 帧编码,在编码效率显著提高的同时,编码时间没有提高很多。在这个例子中,采用多参考帧模式大大增加了计算时间,然而编码效率仅有很小的改善。但是,应该注意到:(1)对于不同的视频序列,将得到不同的输出结果,不同模式所起的作用也是不同的;(2)使用的测试软件并没有进行优化。

4 H.264 的实现

在 2002 年 9 月阿姆斯特丹的 IBC 会议上,VideoLocus 公司演示了开发的 H.264 编解码器,是基于 Pentium4 平台计算机上实现的,采用 FPGA card 做为硬件加速,用于执行运动估计,帧内预测、模式选择统计和视频的预处理。2002 年 10 月,UBVideo 公司在一台 800MHz 的 Pentium3 笔记本电脑上实现了 H.264 的基本档次,对于 CIF 格式,编码帧率可达 49fps,解码帧率可达 105fps,编解码器同时运行可达 33fps。与参考软件相比,码率提高了大约 10%^[4]。

其他很多公司,如 Deutsche Telekom, Broadcom, Nokia, Motorola, the Heinrich Hertz Institute (HHI) 等目前都在开发 H.264 的实时解决方案。在 Pentium4 平台上运行的软件已可以对 CIF 格式的视频进行实时解码和 20fps 的编码。HHI 的解码器在 ARM922 处理器上对于 QCIF 格式可达到每秒 25 帧的解码速度^[5]。清华大学在 TI 公司的 NVDK C6416 开发环境下,对于 QCIF 视频序列,也达到了每秒 50~60 帧的解码速度^[6]。在该平台上,也实现了 H.264 主档次下的编解码器,在 QCIF 格式下达到了实时的效果。

5 结束语

H.264 的高编码效率可以提供新的应用领域和商业机会,将适用于有线、卫星、DSL 和 Cable Modem 等媒介的数字广播;光、磁设备的数字存储,如 DVD 等;通过 ISDN、Ethernet、LAN、DSL、无线和移动网络以及上述混合网络传输的对话型业务、视频点播和多媒体服务等,数码影院和视频监控、军事应用等领域。H.264 的成功应用将取决于对其运行平台的选择,对编解码器的精细设计和对编码参数的有效选择。本文通过对 H.264 编码效率和计算复杂度进行评估,从而为特定应用的编解码器的设计提供了有效的指导信息。(收稿日期:2005 年 7 月)

参考文献

- Advanced video coding[S].ISO/IEC 14496-10,ITU-T Recommendation H.264,2003
- Thomas Wiegand,Gray J Sullivan et al.Overview of H.264/AVC video coding standard[J].IEEE Trans on CSVT,2003;13(7):560~576
- Iain E C Richardson.H.264 and MPOEG-4 video compression.The Robert Gordon University,Aberdeen,UK,2004
- 沈兰荪,卓力.小波编码与网络视频传输[M].科学出版社,2005
- 计文平,郭宝龙,丁贵广.新一代视频编码标准的研究[J].计算机应用与软件,2003;(12)
- 安维嵘,张旭东.H.264 视频解码器在 C6416 DSP 上的实现[J].电子技术应用,2004;(9):46~48