

H. 264 中块分割模式选择算法研究^{*})

郭建军 戴 葵 成 运 王志英

(国防科技大学计算机学院 长沙 410073)

摘 要 H. 264 是一种新提出来的视频压缩标准。本文首先介绍了 H. 264 中所采用的多宏块分割模式技术,然后对其中比较耗时的块分割模式选择模块进行测试分析,在此基础上提出了一种块分割模式选择算法,并对该算法进行了测试,实验结果表明:与 JM8.2 编码方案相比采用新的算法后的运行时间加速比可达到 4.5 以上,峰值信噪比只是略有下降。

关键词 视频压缩, H. 264, 码率, 模式选择, 帧类型

Research on Block Partition Mode Selection Algorithm in H. 264

GUO Jian-Jun DAI Kui CHENG Yun WANG Zhi-Ying

(School of Computer Science, National University of Defense Technology, Changsha 410073)

Abstract H. 264 is a newly proposed video compression standard. The multi-macroblock partition mode technology adopted in H. 264 is introduced first, then some tests and analysis on the block partition mode selection module which is time-consuming are carried out, and based on the analysis of the test results, a new block partition mode selection algorithm is proposed. Experimental results indicate that the new algorithm can achieve a speedup of above 4.5 compared with JM8.2 and brings only a little degradation of PSNR.

Keywords Video compression, H. 264, Bitrate, Mode selection, Frame type

1 前言^[1]

多媒体技术的迅猛发展及其网络应用使得数据量巨大的视频图像信息处理技术越来越成为多媒体通信研究的热点问题。H. 264 视频编码标准是由 ITU 和 ISO 联合组成的联合视频专家组(JVT)负责制定并于 2003 年 3 月正式获得通过的新一代视频编码国际标准。它吸收了现有视频编码标准的优点,同时又提出了许多新的思想和算法,大幅度提高了编码效率和图像质量。统计表明在相同图像质量下, H. 264 比 H. 263+ 或 MPEG-4 简单配置要节省近 50% 的码率^[2]。H. 264 标准和先前的视频标准相比无论在视频编码效率还是图像质量方面都有了很大的提高,它必将取得日益广泛的应用。

多宏块分割模式的运动估计是 H. 264 中采用的一项重要技术,通过对基本宏块进行更细粒度的分割来很好地切合实际运动物体的形状,提高了运动估计的精确程度。但是在对每一种可能的宏块模式进行运动估计和率失真优化使得计算量急剧增加。在运动估计中如果能很快地确定最优宏块模式,减少没必要的宏块模式搜索和代价函数运算,那么将极大地减小运动估计时间和运算复杂度,从而加快运动估计速度。由于采用多宏块分割模式运动估计消耗了大量运算时间,因此占用大量运算时间的细粒度分割是不是对所有情况都值得就成为我们关注的焦点。为此本文通过实际测试分析了宏块分割粒度与编码性能及编码帧类型的关系,并据此得到一个

根据具体码率要求及帧类型进行分割粒度选择的优化算法,之后对这种算法进行了测试,测试结果表明采用这种算法编码与 JM8.2 相比加速比可以达到 4.5 以上,峰值信噪比只是略有下降,实际中不会影响视频效果,从而大大有利于提高编码的实时性能。

2 多宏块分割模式运动估计简介^[3~5]

H. 264 中规定了 7 种用于帧间预测的块模式(16×16, 16×8, 8×16, 8×8, 8×4, 4×8, 4×4)、9 种 4×4 的帧内预测模式和 4 种 16×16 的帧内预测模式。在对图像进行帧间(P 帧和 B 帧)编码时,编码器需对每一宏块使用 7 种帧间预测块模式、13 种帧内预测模式和 1 种 SKIP/DIRECT 模式进行编码,然后运用率失真优化选择最优模式。帧间模式编码中对于每一种分割区域都有一个对应的运动矢量,每个运动矢量也必须进行编码传输,同时分割区域的选择方式也必须编码到压缩码率之中。选用较大的预测块尺寸(如 16×16, 16×8, 8×16)意味着用于标明运动矢量和分割区域类型的比特会比较少,适合于图像细节比较少、比较平稳的区域的匹配。这种多模式的灵活细致的分割,更切合图像中实际运动物体的形状,大大提高了运动估计的精确程度。采用较小的预测块尺寸(如 8×4, 4×4 等)会使运动补偿之后的残差信号能量降低,但用于标明运动矢量和分割区域类型的比特会有所增加,文[6]中就提到要在这两者之间进行折衷。实验表明, H. 264 编码效率的提高有很大部分来源于不同块形状的

^{*})基金项目:国家自然科学基金资助项目(60173040);湖南省教育厅科研项目(项目编号:04B071)。郭建军 硕士生,主要研究方向:视频编码算法、微处理器设计;戴 葵 副教授,博士,主要研究方向:高级计算机体系结构、微处理器设计技术;成 运 副教授,博士生,主要研究方向:视频编码算法及其高效实现、多媒体处理器体系结构等;王志英 教授,博士生导师,主要研究方向:高级计算机体系结构、高性能微处理器设计。

运动估计与率失真优化,然而这一模式选择基本上是一遍历过程,增加了计算复杂度。而且该标准中为了提高编码性能还采用了多参考帧编码,使得运动估计的范围更大,因而,在通常情况下运动估计部分在整个编解码中所占的运算量是最高的,提高运动估计的效率对实现 H. 264 实时编解码具有重要意义。如果能够事先对待测模式进行筛选,获得一个相对较小的待测模式集,便可有效提高编码速度。

3 块切分模式测试分析与研究

H. 264 的显著特点就是可以在中低码率条件下提供较好的图像质量,同时它很注重实时编码效果,因此必须在编码处理时间和编码质量上做更深入的研究。对于帧间预测技术,可以在编码器中加入率失真优化的运动估计和模式选择算法,均衡考虑编码质量和复杂度,预先估计当前编码块最可能采用的块形状和预测位置;也可以利用已处理的预测块信息,合理调整预测方向和及时跳出预测过程^[7]。采用 4×4、4×8、8×4 的小粒度块分割模式的编码图像块在全部编码块只占一小部分,但却占用了大量编码时间^[8]。同样地,对于多参考帧和扩大的搜索范围必须要采用合理的机制在编码效率和编码速度上做出有效的判断和权衡。这里我们着重研究低码率条件下块分割模式的粒度选择问题。

3.1 不同块切分模式下性能分析及结论

如前所述,H. 264 中采用 4×4、4×8、8×4 子块分割模式的编码图像块只占有全部编码块的一小部分,但却占用了很大一部分编码处理时间。采用子块分割技术虽然可以提高信噪比,但是提高的能力是很有限的,并且在实际编码中人们更关心的是编码时间,这直接决定了编码能否达到实时效果。因此采用 8×8 块以下的分割模式是否必要成为我们关注的焦点。为此针对块切分模式对 carphone. qcif 序列在 JM8. 2^[9] 平台下做了测试,测试机器主要配置如下: Pentium4 2. 4G CPU, 1G 内存, 测试结果如表 1 所示,其中 a 表示支持 7 种分割模式, b 表示持前 3 种分割模式, c 表示支持第 1 种分割模式; PSNR 表示峰值信噪比; BR 表示比特率。

表 1 不同量化参数下不同块分割粒度编码比较

QP	模式	a		b		c	
		PSNR	BR	PSNR	BR	PSNR	BR
10		50. 01	1345. 59	49. 83	1375. 99	49. 70	1456. 17
16		45. 68	581. 63	45. 49	603. 53	45. 32	662. 86
22		41. 38	254. 41	41. 15	265. 73	40. 90	297. 20
28		36. 79	95. 82	36. 56	98. 31	36. 19	110. 80
34		32. 60	36. 80	32. 37	36. 56	31. 94	38. 79
40		28. 78	15. 52	28. 77	15. 31	28. 30	15. 32

由表 1 可以看出,当量化参数(QP)较大时较细粒度的块分割对比率率(BR)的减少和峰值信噪比(PSNR)的提高效果并不明显。因此,对于大的 QP 可以不必切分到最小粒度。

3.2 编码帧类型与块切分粒度分析

在同样条件下对其他几个典型图像序列进行测试,发现 P 帧和 B 帧中每种宏块所占比例并不相同,对于 P 帧主要是 8×8 模式以上比例较大,而对 B 帧主要是 SKIP 模式和 16×16 模式,其他的宏块比例相对都较小,表 2 列出了 P 帧和 B 帧各种宏块的比例。

PB 帧编码要在参考帧列表中选择较优的参考帧进行编码,P 帧需在一个参考帧列表中搜索,B 帧需在 2 个参考帧列

表中进行搜索,代价是不同的,如果都按照统一的粒度去进行宏块切分代价相对较大。另外 PB 帧间编码时帧内编码模式(intra)的比例十分小,但是同样要进行搜索,代价和收益不匹配。因此,编码过程中可以针对 PB 帧采用不同的切分粒度,同时去除一些帧内模式。

表 2 几个测试序列各种模式分布比例

序列	carphone. qcif		foreman. qcif		akiyo. qcif	
	P	B	P	B	P	B
SKIP	22. 47	37. 58	15. 31	44. 18	59. 60	63. 64
16×16	26. 65	48. 88	26. 63	45. 45	10. 90	25. 07
16×8	11. 52	4. 49	9. 25	2. 92	3. 28	3. 13
8×16	11. 81	5. 92	21. 64	4. 15	4. 21	4. 93
8×8	26. 20	3. 05	27. 06	3. 30	21. 95	3. 24
intra16×16	0. 33	0. 02	0. 05	0. 00	0. 04	0. 00
intra4×4	1. 01	0. 06	0. 05	0. 00	0. 02	0. 00

3.3 一种新的块分割模式选择算法

综合以上因素,可以首先根据 QP 值来决定所选块分割粒度,再针对编码帧类型进行粒度筛选,从而减少总体运算时间、降低码率。该算法基本思想如下: 设立一个 QP 阈值,首先根据 QP 阈值进行第一次粒度筛选,然后根据编码帧类型进行第二次粒度筛选。

对于不同码率,阈值的设置应该有所区别,因此对于视频序列 foreman. qcif 进行了如下测试,编码 100 帧,帧模式采用 IBPBP 格式,编码方法采用基于上下文的二进制算术编码(CABAC)。测试中发现在一定码率范围内,该阈值与码率成近似线性关系,如图 1。根据该图,得到线性回归公式如公式(1)所示。

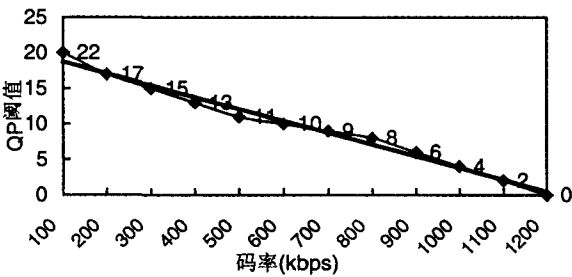


图 1 QP 阈值与码率关系

$$threshold_qp = \lceil 20.3409 - 0.017448 \times bitrate \rceil \quad (1)$$

该算法的步骤如下:

步骤 1: 获取当前实时码率值,根据该码率利用得到的经验公式计算出 QP 的阈值。

步骤 2: 根据该阈值选择合适的块切分模式进行帧间预测。如果当前 QP 值大于该阈值则采用粗分割粒度方式,转步骤 4,否则转步骤 3。

步骤 3: 判断当前编码帧类型。若为 P 帧,粒度切分到 8×8,若为 B 帧只进行 16×16 切分。

步骤 4: 进行 PB 帧编码中 intra 模式 16×16 或 4×4 粒度选择。如果码率改变,转步骤 1。

4 性能评测

根据该块分割模式选择算法再对典型视频图像序列 claire 及 highway 分别进行编码测试。所得的部分结果如图

(下转第 45 页)

• 网卡资源:主要考察网卡带宽。在网卡队列处,网卡资源使用率 $\rho_{net} = \lambda/\mu$, 如果发现在指定的条件下 ρ_{net} 的值接近 1 时,可知 Domino 邮件服务器在网卡处存在瓶颈。可以考虑更换速度更快的网卡。

• 磁盘资源:主要考察磁盘带宽。计算方法与网卡相同。在磁盘队列处,磁盘资源使用率 $\rho_{disk} = \lambda/\mu$, 如果发现在指定的条件下 ρ_{disk} 的值接近 1 时,可知 Domino 邮件服务器在磁盘处存在瓶颈。可以考虑更换读写速度更快的磁盘。同时还需考察磁盘容量。根据实验可知,每个用户邮箱初始化时占用磁盘空间为 5.5M, 如果为每个用户分配的邮箱最大容量为 50M, 在计算所需的磁盘空间时,我们取均值 25M。给定用户数 n , 则除了 Domino 应用程序所占空间外,磁盘还需留出至少 $(25 * n)M$ 空间。

• CPU 资源:主要考察 router 进程, dbserver 线程, transfer 线程, deliver 线程的占用 CPU 时间的情况。根据到达率和服务率,可得进(线)程的使用率 ρ , 若各使用率均接近 1, 则说明 CPU 的处理能力是瓶颈,需要更换处理速度更快的 CPU, 若只有个别使用率接近 1, 则可以通过调整该进(线)程的优先级来解决。

• 内存资源:主要考察内存的使用量。我们在模型中没有将内存带宽作为瓶颈考虑,因此没有建立内存资源的队列,而且在现实中内存容量的需求可以很容易地被满足。所以我们只给出 Domino 邮件服务器内存需求的经验公式。设给定并发用户数 n , 内存使用的公式如下:

所需的内存容量的最小值 $= (n * 300 * 3)k$

小结 本文中我们利用排队论理论对基于 Domino 的邮件服务器进行性能建模和分析,通过建立该模型我们实现了对系统支持的在线用户数和响应时间等指标进行预测,对 Domino 邮件服务器的资源瓶颈进行考察。

我们的模型只是基于一个 Domino 域建立的模型,在下一步的工作中我们将进一步考虑多个 Domino 域的情况。

参考文献

- 1 随机过程. 中国统计出版社, 1997
- 2 管理 DOMINO 系统(第一卷). LOTUS
- 3 运筹学. 清华大学出版社, 1982
- 4 排队论. 北京邮电学院出版社, 1994
- 5 Slothouber L P. A Model of Web Server Performance. Starline Com, 1996
- 6 乐捷 DOMINO 邮件系统用户手册. 乐捷公司, 2001

(上接第 26 页)

2 和图 3 所示。这里 PB 帧编码没有采用 intra 模式。

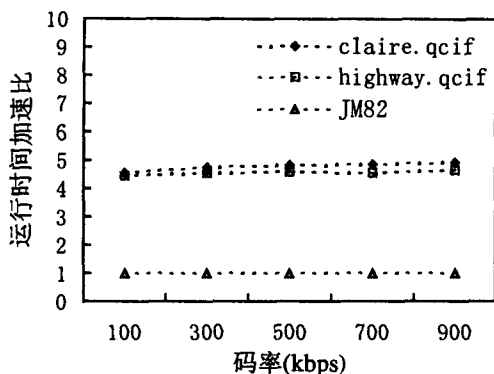
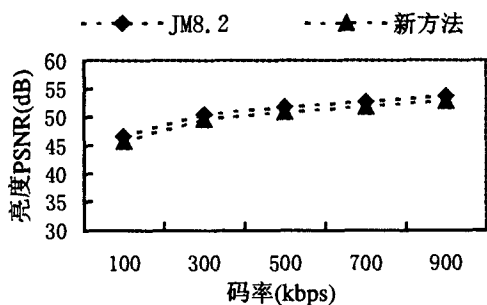
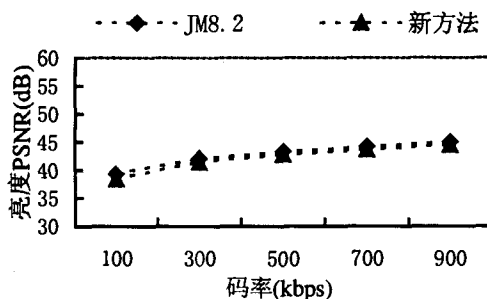


图2 几种典型序列运行时间加速比



(a) claire 序列亮度率失真比较



(b) highway 序列亮度率失真比较

图3

图2、图3表明对于几种代表序列该算法相对 JM8.2 的运行时间加速比可以达到 4.5 以上, PSNR 略有下降, 但实际当中不会对视频质量造成很大影响。而且在该算法步骤 2 中 intra 模式的粒度可以放宽, 如可以只采用 intra16×16 模式, 测试表明这样 PSNR 下降较少, 只是加速比要稍有降低, 实际中可以根据需要来调整 intra 的粒度。

结束语 H. 264 视频编码标准具有编码效率高、图像质量好、适合低码率传输等优点, 并能很好地处理信息包丢失问题, 必将促进视频压缩编码技术和多媒体通信技术的进一步发展。H. 264 主要面向的是低比特率、低延时的应用。因此, 减少运算时间、降低码率是其中一个关键因素。本文在分析目前 H. 264 中采用的关键技术的基础上, 针对其中较为耗时的块分割模式选择模块进行优化, 提出了一种新的块分割模式选择算法, 通过对比实验测试, 验证了该算法和 JM8.2 相比, 在 PSNR 稍有下降的情况下运行时间加速比达到了 4.5 以上, 大大提高了编码的实时效果。

参考文献

- 1 ITU-T Rec. H. 264 | ISO/IEC 14496-10 AVC, Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification, document JVT-G050. March 2003
- 2 Wiegand T, Sullivan G J, Bjøntegaard G, Luthra A. Overview of the H. 264/AVC Video Coding Standard. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7)
- 3 李宾, 高平. H. 264 编码系统的特点及其应用前景. 数字电视与数字视频, 2003 (6)
- 4 王嵩, 薛全, 张颖, 陈建乐. H. 264 视频编码新标准及性能分析. 数字电视与数字视频, 2003 (6)
- 5 Richardson I E G. H. 264 / MPEG-4 Part 10 White Paper: Prediction of Inter Macroblocs in P-slices
- 6 Richardson I E G. H. 264 and MPEG-4 Video Compression. UK: Wiley & Sons, 2003
- 7 朱冬冬, 戴琼海. H. 264 快速帧间编码模式选择算法. 有线电视技术, 2004(9)
- 8 Evaluation of the H. 264 codec Internal Report. Carlos Miguel Tavares Calafate, Manuel Pérez Malumbres, 2003
- 9 Joint Video Team (JVT). Test Model JM8. 2. at <http://bs.hhi.de/~suehring/ttml/download/>, July 2004