

联合分数运动估计优化的多参考帧快速选择算法

陆寄远^{1,2} 孙少晖¹ 朝红阳³

(中山大学信息科学与技术学院 广州 510275)¹ (广东金融学院计算机科学与技术系 广州 510521)²
(中山大学软件学院 广州 510275)³

摘 要 多参考帧和分数运动补偿技术虽然提高了视频编码性能,但同时也使编码器的内存消耗和计算复杂度成倍地增加。其主要原因是多参考帧技术令运动估计重复地应用在每个参考帧上。通过分析这两种技术对编码性能贡献的相关性,提出了一种联合优化的快速算法。首先,根据不同参考帧之间的合成运动向量,选择空间上最接近的候选参考帧。然后,通过二维多项式模型预测最优分数残插值,决定采用分数运动估计的参考帧。实验结果表明,该方法在不损失任何率失真性能的情况下,可以把计算的复杂度减少为原来的 1/3。

关键词 视频编码,多参考帧选择,分数运动估计,资源优化

中图法分类号 TP391 文献标识码 A

Fast Algorithm for Multi-frame Selection with Jointly Optimization of Fractional Pixel Motion Estimation

LU Ji-yuan^{1,2} SUN Shao-hui¹ CHAO Hong-yang³

(School of Information Science and Technology, Sun Yat-Sen University, Guangzhou 510275, China)¹

(Department of Computer Science and Technology, Guangdong University of Finance, Guangzhou 510521, China)²

(School of Software, Sun Yat-Sen University, Guangzhou 510275, China)³

Abstract Multi-frame and fractional pixel motion compensation technology improve the video coding performance at expense of the high memory consumption and computational complexity. We proposed a jointly optimized algorithm for these two technologies by analyzing the connection between them. Firstly, the composed MVs were used to select the candidate reference frame set. Secondly, fractional pixel motion estimation was only applied on the reference frame which has the optimal predicted residual based on a two-dimension polynomial model. According to our experiments, a reduction of 60% execution time is attained while still maintaining the same R-D performance with H. 264 Joint Model (JM).

Keywords Video coding, Multi-frame selection, Fractional pixel motion estimation, Resource optimization

1 引言

作为新的国际视频编码标准, H. 264/AVC 的压缩性能比以往的编码标准都有了很大的提高。这归功于 H. 264/AVC 所引入的多项新技术。其中,多参考帧和 1/4 分数运动补偿技术对于提高视频编码的率失真性能起了很重要的作用。例如,采用这两种技术对一组标清 VGA 格式(640 * 480)视频进行编码,在重建质量和码率上分别比使用单参考帧和 1/2 分数运动补偿进行编码提高了 0.26dB 和减少了 9.71%。但这种编码性能的提高也带来了高的计算消耗。多参考帧和 1/4 分数运动补偿技术扩大了视频编码中运动估计的范围,增加了计算复杂度。以上面为例,5 个参考帧令整数运动估计的搜索范围比单参考帧扩大了 5 倍;1/4 分数运动补偿则令分数运动估计的搜索范围比 1/2 分数运动补偿扩大

了 2 倍。那么总体来说,运动估计的搜索范围扩大了 $5 \times 2 = 10$ 倍。这造成了视频编码运动估计的计算复杂度大幅度地增加。

目前,对于解决多参考帧和 1/4 分数运动补偿技术计算量庞大的问题已经出现了不少的工作^[1-8]。但这些研究工作都是割裂地对以上两种技术构建快速算法,即只优化多参考帧或者只优化分数运动估计,缺少了对这两种技术内在联系的考虑。通过本文第 2 节的分析,我们可以通过这两种技术的内在联系来进一步加快运动估计的速度,同时保证视频编码的性能是至关重要的。下面,简单地回顾一下目前常用的快速多参考帧搜索算法和快速分数运动估计算法。

首先,常用的快速多参考帧搜索算法大致分成以下两类。第一类是相邻参考帧预测法,这类方法是根据一定的策略,只参考几个比较重要的参考帧,跳过其余的参考帧,以此提高多

到稿日期:2009-07-30 返修日期:2009-10-23 本文受国家 863 计划项目——基于动漫数据库的计算机辅助动漫创作关键技术与系统(2007AA01Z340)和广东省自然科学基金项目——新一代视频编码标准中运动估计关键技术的研究(07003728)资助。

陆寄远(1976—),男,博士生,讲师,CCF 会员,主要研究方向为视频编码、图像传输等,E-mail:dtc005001@163.com;孙少晖(1983—),男,硕士,主要研究方向为视频编码;朝红阳(1957—),女,教授,博士生导师,主要研究方向为视频编码、图像处理等,E-mail:isschhy@mail.sysu.edu.cn (通信作者)。

参考帧运动估计的速度^[1,2]。因为不同的参考帧的重要性是不同的,所以不需要采用相同的处理方式。这类算法利用了相邻宏块参考帧信息的相关性,尽早地决定合适的参考帧。但如果相邻的信息不足以决定正确的参考帧,那么这个错误的参考帧信息就会被不断传播,从而导致编码性能的大幅下降。另一类是合成运动向量法,如, MVCMMME 算法等^[3,4]。这类算法并不是选择性地搜索参考帧,而是根据不同参考帧在时间上的相关性,通过已经被存储的运动向量合成各个参考帧的运动向量,减少了花费在每个参考帧上的搜索时间。其不足之处在于合成运动向量的精度较难把握,在某些情况下合成运动向量造成预测偏差过大,影响编码性能质量。

其次,快速分数运动估计算法的研究主要集中在基于最优分数运动向量预测的方法上。这类方法先通过对最优的分数运动向量的预测,然后使用一个细化模板校正该预测值。CBFPS^[5], SUMH FPS^[6] 以及 EPZS FPS^[7] 方法都是其中具有代表性的算法。这类方法都是以减少运动估计搜索点数为目标,而且都被 H. 264 的联合模型 JM 所采纳。本文在这些算法的基础上面向 H. 265 的测试模型 KTA 构建了一种能进一步减少分数搜索点数的快速算法^[8]。

现有的这些快速算法虽然在一定程度上优化了多参考帧和分数运动搜索的时间,但却忽略了两者的内在联系。如果考虑了各种编码技术内在联系,那么就可以更好地加快各种编码技术的运算。实际上,多参考帧技术所获得的编码收益并非依赖于参考帧数目的多少,而是依赖于视频序列内容的本质特征。对于不同的情况,采用多参考帧技术所带来的效果是不同的。但由于采用了多参考帧而增加的计算复杂度却与参考帧的数目成正比,这造就了多参考帧技术仍有不少可以继续优化的空间。

本文通过分析多参考帧技术与分数运动补偿技术之间的联系提出了一种联合算法,它在不损失编码效率的情况下,可加快多参考和分数运动估计的速度。本文第 2 节介绍了该联合算法;第 3 节是实验的结果;最后是结论。

2 联合分数运动估计的多参考帧快速选择算法

下面将通过对多参考帧和分数运动补偿技术的分析,说明能够提供更好编码性能的参考帧一般是有着分数像素运动的参考帧。然后,给出联合分数运动估计的多参考帧快速选择算法。

2.1 多参考帧运动补偿技术和分数运动补偿技术分析

当视频信号存在周期运动,采用多参考帧运动补偿技术可以达到比单参考帧更好的预测效果,同时明显地提高视频编码效率^[3]。假设我们要对一个小球反复跳动的视频进行编码。由于视频中存在重复的动作,选择与当前帧相隔恰好为一个运动周期的视频帧作为参考帧,小球变化的差别会比直接选择前一帧图像作为参考帧更小。图 1 表示使用 1~5 个参考帧在固定码率情况下对不同测试序列编码的信噪比变化。从图中可以看出,当参考帧数目增加时,信噪比并不随之而线性增长。信噪比的增长是跳跃式的, Mobile, Foreman, Bigship 和 Silent 序列采用 2 个参考帧就可以达到与采用 5 个参考帧十分接近的信噪比。而其余的序列,最多采用 3 个参考帧也能达到与采用 5 个参考帧相当接近的信噪比。由此可知,多参考帧对率失真性能贡献多数时候只集中在少数的

几帧中。不考虑其对编码贡献的轻重,统一的搜索策略只会把计算资源浪费在无用的计算上。

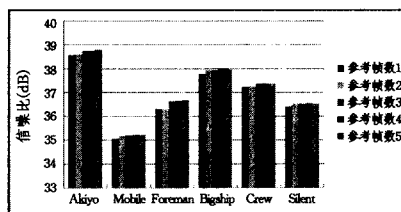


图1 固定码率下使用5个参考帧编码的信噪比对比

多参考帧运动补偿技术可以有效地消除两种类型的冗余。一种是周期运动,该类型的时间冗余仅仅通过单参考帧预测是很难消除的,前面已经作过介绍了;另一种就是分数像素精度运动^[1]。亚像素精度运动源于数码摄像头的采样。一个物体运动的位移不一定是整数个像素点。在不同时刻的视频帧中,对物体的采样位置各有不同,即物体有可能刚好落在采样点上,也有可能没有落在采样点上。正因如此,不同时刻的图像可以在时间间隔较远的参考帧中找更好的匹配。另一方面,要消除这种分数像素运动的时间冗余,可以采用的一个手段是分数点运动补偿技术^[5]。分数运动补偿技术解决了视频信号中频率混淆的问题,因为实际图像的纹理信号和视频帧中实际运动的位移精度有可能高于采样设备的精度,所以在压缩标准中一般会引入分数运动向量。现实中的视频信号是连续的,而摄像头在捕捉影像时必须离散化视频信号,也就是采样和量化。一帧图像由许多不同颜色的点组成,这些点被称作采样点。如果实际的运动向量不落在采样点上,就必须通过分数点插值计算对应的分数点。采用分数运动补偿技术在一定程度上与采用多参考帧运动补偿有着异曲同工的作用。

现在流行的视频编码标准如: H. 264, AVS 和 VC-1 等,都支持 5 个以上的参考帧和 1/4 像素的运动补偿技术。图 2 是使用 H. 264 编码器对 Crew 序列编码后的率失真曲线,其中基准是采用 1/2 分数点运动补偿技术以及一个参考帧。从图 2 可以看到,在加入多参考帧或 1/4 分数点后,编码性能都得到了显著的提升。但同时加入 1/4 分数点和多参考帧后所得到的编码收益,却远低于单独加入这两者的收益之和。由此可知,多参考帧技术和分数运动补偿技术两者之间存在着非常多的重叠部分。本文正是考虑到了这个重叠性,针对多参考帧运动补偿技术提出了一种联合分数运动估计的快速算法。

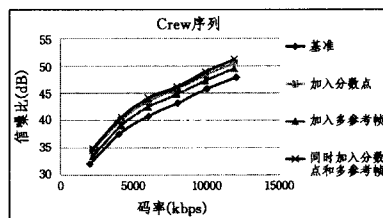


图2 多参考帧技术和1/4像素运动补偿技术对编码性能的影响

2.2 联合分数运动估计的多参考帧快速选择算法

由于现有的视频编码标准多数采用 1/4 像素精度的运动向量,因此在前后连续两帧里,物体的运动可以是整数像素步长的、半像素步长的、或者是 1/4 像素步长的。如图 3 所示,如果物体运动正好落在数字摄像镜头的采样点上,那么称这

个物体为“整数像素定位”(图3中左方和右方小图);如果物体运动落在半像素或1/4像素的插值点上,那么称这个物体为“半像素定位”或“1/4像素定位”(图3中间的小图)。

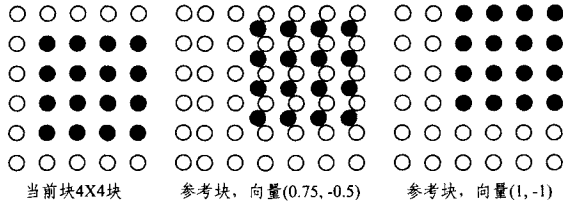


图3 整数像素和亚像素预测的示例

图4是一个多参考帧分数运动估计的例子,其中帧 t 表示当前帧,帧 $(t-1)$ 至帧 $(t-3)$ 是参考帧。从图中可知,帧 t 中的像素点可以在每个参考帧中分别找到其最匹配的像素。由于帧 $(t-3)$ 中像素的位移是1/4,而在帧 $(t-1)$ 和帧 $(t-2)$ 中却是1/2,那么可以知道这是一个周期的运动。在这种情况下,时间相隔较远的参考帧反而匹配程度要比时间上靠近的参考帧要好。在这一情况下,本文给出多参考帧的选择算法,首先对最接近的参考帧 $(t-1)$ 完成运动估计。因为多数的最优预测块都会落在最接近的参考帧上,所以对 $(t-1)$ 帧完成标准的运动估计是必要的。然后,通过运动向量合成方法^[4]找到在各参考帧的预测运动向量,再使用菱形模板搜索整数运动向量。因为整数运动估计不需要用到插值点,只需要完成匹配运算,所以这样不仅可以获得各个参考帧的整数运动向量,而且计算简单。

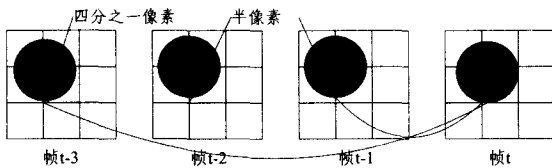


图4 运动估计的参考帧选择

以下将要完成的是多参考帧的分数运动估计。本文提出的方法并不需要在每个参考帧上都完成分数运动估计,只有满足下面两个条件之一的参考帧才需要完成分数运动估计:一是帧距远的参考帧有着比帧距近参考帧更小的运动向量;二是通过二维分数模型估计的最优残差值比帧 $(t-1)$ 要小。第一个条件针对的是视频中出现的周期运动。一般情况下,参考帧与当前帧的距离越远,运动向量应该越大,只有周期运动出现时才会导致时间距离越远的两帧运动向量反而更小。当周期运动出现的时候,除了与当前帧时间距离最短的参考帧 $(t-1)$ 帧有机会成为最优的参考帧外,空间距离最短的参考帧也同样有机会成为最优参考帧。空间距离最短的参考帧指的是与当前要编码宏块的运动位移最短的那个参考帧。当周期运动出现时,时间距离最短的参考帧并不等于就是空间距离最短的参考帧。第二个条件针对的是多参考帧中出现分数运动位移的情况。因为在分数运动估计完成前难以预先知道运动向量是否具有分数位移,所以本文采用抛物线模型来预测分数最优残差^[5],并只在拥有最优预测残插值的参考帧上使用分数运动估计,以减少分数运动估计时间。分数残差的抛物线模型如下:

$$F(x, y) = A \times x^2 + B \times x + C \times y^2 + D \times y + E \quad (1)$$

式中, x 和 y 为分数搜索区域的坐标。 A, B, C, D 和 E 为相邻整数位置的残差值,如图5所示。图5中心的整数点 E 是整

数运动估计中的残差最优位置。因为本文采用菱形模板完成整数运动估计,所以点 E 周围相邻的4个整数点(B, C, D 和 E)的残差值都是已知的。利用这5个整数点的残差值,就可以建立如式(1)的抛物线模型。根据这个抛物线模型可以得到预测最优分数残差为:

$$Cost_{pred} = E - \frac{B^2}{4 \cdot A} - \frac{D^2}{4 \cdot C} \quad (2)$$

式中, A, B, C, D 和 E 的含义与式(1)相同。每个候选参考帧在完成整数运动估计后,都会用式(2)对最优分数残差进行预测。而分数运动估计仅仅应用在拥有最小预测值的那个参考帧中。

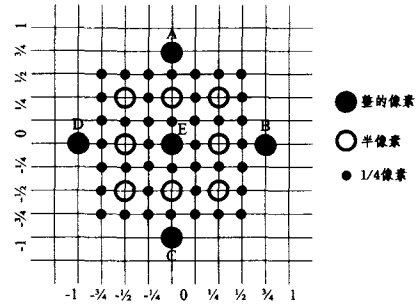


图5 整数像素和亚像素预测的示例

通过上述的多参考帧选择条件减少分数运动估计需要搜索的参考帧数目后,再配合动态分数像素插值方法,不仅可以大幅减少计算的时间,也可减少内存的消耗。所谓动态分数插值方法并不是预先把所有分数点都计算出来,而是只预先计算某部分的分数点,其余大部分的分数点是在使用时才计算。采用这种插值方法可以在减少内存消耗的同时,节省计算资源。不过,有效利用动态插值方法的前提就是优化的多参考帧选择和分数运动估计方法。这恰恰正是本文所提出的优化算法。

3 实验结果

本文的实验环境是 H. 264/AVC 的联合模型 JM 15.0 (Joint Model 15.0)。参考帧的数目为5,运动估计的搜索半径为32,使用高精度的率失真优化选项,所有的测试序列采用 IPPPP 格式。为了充分体现算法的有效性,本文对 Akiyo, Mobile, Foreman, Bigship, Crew 和 Silent 等多种分辨率和不同运动情况的序列进行测试。参与比较的算法有 JM 中的全搜索多参考帧算法、MVCMMME 快速多参考帧选择算法^[4]。

图6和图7是不同序列使用不同多参考帧选择算法后率失真性能的比较。因为这些序列的分辨率和运动情况都有很大的差别,它们编码后的码率相差比较悬殊,所以图7以百分比的方式对码率变化进行比较。从两图中可以看出,采用本文的方法无论在信噪比还是码率,都保持与精度最高的全参考帧搜索方法相同,并不损失任何编码性能。

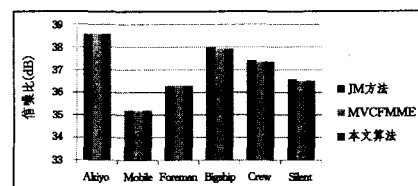


图6 率失真性能比较

(下转第302页)

sign for High Performance Memory Systems[A]//IEEE Proceedings of the Second International Symposium on High-Performance Computer Architecture[C]. 1996;254-263

[3] Lai A, Fide C, Falsafi B. Dead-Block Prediction and Dead-Block Correlating Prefetchers[A]//Proceedings of the 28th International Symposium on Computer Architecture[C]. 2001;144-154

[4] Schumann R C. Design of the 2 1 1 7 4 Memory Controller for Digital Personal Workstations[J]. Digital Technical Journal, 1997, 9(2):57-69

[5] Zhao Z, Zhichun Z, Zhang X. A permutation-Based Page Interleaving Scheme to Reduce Row-Buffer Conflicts and Exploit Data Locality[A]//Proceedings of the 33rd Annual International Symposium on Microarchitecture[C]. 2000;32-41

[6] Burger D C, Austin T M. The SimpleScalar Tool Set [R]. CS-TR-97-1342. 1997

[7] Samsung Semiconductor. 512 Mb B-die DDR400 SDRAM Specification[R]. Revision 1. 0. 2003

(上接第 285 页)

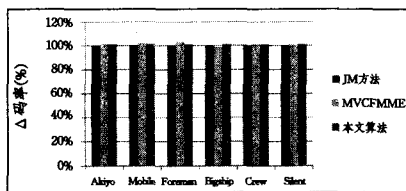


图 7 码率变化比较

速度的对比通过表 1 给出。表 1 中列出了在 16 和 28 两种量化参数下的实验结果,分别对应高码率和低码率两种情况。无论在低码率还是高码率,本文的算法对于全参考帧搜索方法(即 JM 中使用的算法)有至少一倍的提高,而对于快速算法 MVCMMME,也有平均 20% 的减少。在 Akiyo, Silent 和 Crew 这些运动不甚剧烈的序列中,本文算法能提供更快的速度。其原因主要是较静止的序列内周期运动和分数运动出现得较少,本文的算法可以跳过绝大多数对冗余参考帧的搜索。

表 1 本文算法与 H. 264 JM 快速运动估计算法的比较结果

视频序列	量化值	格式	Δ 复杂度(%) 与 JM 方法	Δ 复杂度(%) 与 MVCMMME
Akiyo	16	CIF	-56.31	-27.14
Mobile	16	4CIF	-55.27	-24.72
Foreman	16	4CIF	-45.54	-25.6
Bigship	16	720p	-53.44	-22.48
Crew	16	720p	-58.18	-21.98
Silent	16	CIF	-57.89	-28.15
Akiyo	28	CIF	-51.11	-26.3
Mobile	28	4CIF	-49.2	-25.02
Foreman	28	4CIF	-55.41	-25.03
Bigship	28	720p	-49.89	-23.57
Crew	28	720p	-50.82	-28.8
Silent	28	CIF	-52.9	-27.9

结束语 本文提出了一种联合分数运动估计的多参考帧快速选择算法。首先,通过对多参考帧选择技术和分数运动估计技术的分析发掘了两者之间的相关性;然后,利用合成的运动向量和分数残差曲面模型,确定进行分数运动估计的参考帧。该方法仅在时间相邻、运动相邻和分数预测残差最小

的参考帧上进行分数运动估计,在大幅地加快多参考帧运动估计速度的同时,保持了视频编码的率失真性能。

参考文献

[1] Shen Liquan, Liu Zhi, Zhang Zhaoyang, et al. An Adaptive and Fast H. 264 Multi-Frame Selection Algorithm Based on Information from Previous Searches [C]//2007 IEEE International Conference on Multimedia and Expo. 2007

[2] Sun Qichao, Chen Xin-hao, Wu Xiaoyang, et al. A Content-adaptive Fast Multiple Reference Frames Motion Estimation in H. 264 [C]//2007 IEEE International Symposium on Circuits and Systems (ISCAS). 2007

[3] Liu Zhenyu, Li Lingfeng, Song Yang, et al. Motion Feature and Hadamard Coefficient-Based Fast Multiple Reference Frame Motion Estimation for H. 264 [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, 18(5):620-632

[4] Chen Mei-juan, Li Gwo-long, Chiang Yi-yen, et al. Fast multi-frame motion estimation algorithms by motion vector composition for the MPEG-4/AVC/H. 264 standard [J]. IEEE Transactions on Multimedia, 2006, 8(3):478-487

[5] Chen Zhibo, Xu Jianfeng, He Yun, et al. Fast integer-pel and fractional-pel motion estimation for H. 264/AVC [J]. Journal of Visual Communication and Image Representation, 2006, 17(2):264-290

[6] Yi Xiaoquan, Zhang Jun, Ling Nam, et al. Improved and simplified fast motion estimation for JM, JVT-P021. doc [C]//Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q. 6) 16th Meeting. Poznan, Poland, 2005

[7] Alexis T, Pankaj T. Fast Subpixel Motion Estimation Support for the Enhanced Predictive Zonal Search Scheme, JVT-Q079 [C]//Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q. 6) 17th Meeting. Nice, France, 2005

[8] 陆寄远, 张培钊, 朝红阳. 一种面向 H. 264 KTA 的快速分数运动估计方法 [J]. 中国图象图形学报, 2010, 15(3):367-371

(上接第 288 页)

[4] Chen L F, Liao H Y, Ko M T, et al. A new LDA-based face recognition system which can solve the small sample size problem [J]. Pattern Recognition, 2000, 33(10):1713-1726

[5] Zhuang X S, Dai D Q. Improved discriminant analysis for high dimension data and its application to face recognition [J]. Pattern Recognition, 2007, 40(5):1570-1578

[6] 宋枫溪, 程科, 杨静宇. 最大散度差和大间距线性投影与支持向量 [J]. 自动化学报, 2004, 30(6):890-896

[7] Yang M H. Kernel Eigenfaces vs. Kernel Fisherfaces face recognition using kernel methods [A]//Proceedings of Fifth IEEE International conference on Automatic Face and Gesture Recognition[C]. Washington DC, USA, 2002;215-220

[8] Hu H F. Orthogonal neighborhood preserving discriminant analysis for face recognition [J]. Pattern Recognition, 2008, 41(9):2045-2054

[9] Georgiades A S, Belhumeur P, Kriegman D, et al. From Few to Many: Illumination Cone Models for Face Recognition Under Variable Lighting and Pose [J]. IEEE Trans. Pattern Anal. Mach. Intelligence, 2001, 23(6):643-660