

判断运动复杂度的快速运动估计算法

黄华¹, 程徽¹, 刘玎玎²

(1. 西安交通大学电子与信息工程学院, 710049, 西安; 2. 华盛顿大学电子工程系, 98195, 美国西雅图)

摘要: 为了提高视频压缩编码中运动估计算法的效率, 提出了一种判断运动复杂度的快速运动估计算法(MC-FME 算法). 首先利用单帧图像的纹理信息与宏块匹配残差自适应地判断帧级运动复杂度, 再利用当前宏块运动向量预测集中预测值的差异性自适应地判断宏块级运动复杂度, 最后根据运动复杂度进行相应的运动估计. 实验结果表明, 与菱形搜索算法和分级运动估计算法相比, MC-FME 算法可以在保证图像质量的前提下, 匹配点数降低 5%~21%, 运动估计时间缩短 3%~5%, 从而达到加快运动估计的目的.

关键词: 运动估计; 运动复杂度; 自适应

中图分类号: TN919.81 **文献标志码:** A **文章编号:** 0253-987X(2009)10-0081-04

A Fast Motion Estimation Algorithm for Motion Complexity Decision

HUANG Hua¹, CHENG Hui¹, LIU Dingding²

(1. School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China;

2. Department of Electronical Engineering, University of Washington, Seattle, Washington 98195, USA)

Abstract: To improve the performance of motion estimation (ME), a new fast motion estimation algorithm is proposed for motion complexity decision (MC-FME). The algorithm adaptively estimates the frame-level motion complexity by using the frame texture information and the macro-block residual value at first, and then adaptively estimates the macro-block-level motion complexity by calculating the variations of predictions in the prediction set of the macro-block motion vectors. The complexity is finally used to estimate the motion. Experimental results and comparisons with the diamond search algorithm and the hierarchical motion estimation algorithm show that the proposed algorithm can get 5% to 21% reduction in the number of matched points and save 3% to 5% motion estimation time while the image quality is guaranteed.

Keywords: motion estimation; motion complexity; adaptive

运动估计作为视频编码的一个重要环节, 不仅决定着整个编码过程的速度, 还直接影响图像质量. 目前, 块匹配算法(BMA)在精确度和复杂性之间提供了一个很好的折中, 人们也提出了许多相关算法^[1-5], 如基于搜索模式的快速块匹配算法和多分辨率下的块匹配算法等. 作为这两类块匹配算法的代表, 菱形搜索算法(DS)^[6-8]和分级运动估计算法(HME)^[9-11]虽然能够达到较好的搜索效果, 但由于没有考虑视频本身的运动信息, 因此搜索效率仍有提升的空间.

本文提出了一种判断运动复杂度的快速运动估计算法(MC-FME 算法), 通过判断视频内容的运动复杂度, 并结合自适应阈值, 选用合适的运动估计算法进行搜索. 实验结果表明, 本文提出的新算法能够在保证图像质量的前提下, 提高搜索效率.

1 菱形搜索算法和分级运动估计算法

DS 算法^[6-8]采用了一种近似于圆形的搜索策略, 具有很好的全方向性. 它从起始点开始按照 LD-SP(Large Diamond-Shaped Pattern)模板以菱形的

形状逐步向外扩展,如果计算的最佳点在中心,则以该点为中心继续按照 SDSP(Small Diamond-Shaped Pattern)模板进行细搜索。

HME^[9-11]算法采用均值法构造一个多分辨率图像金字塔,自底部向上,图像分辨率逐渐减小,例如,一个2级金字塔中,上层的1个像素对应下层4×4像素的图像区域,且上层的像素值为下层对应区域的均值。HME的搜索过程属于多级搜索模式,以低分辨率图像得到的运动向量作为高分辨率图像运动估计的初始值,从整体上缩短运动估计的时间。

上述2种算法具有互补性,DS算法简单,适合于运动平缓的视频序列,但是对于运动变化较大的视频序列则比较耗时。相反,HME算法由于采用了由粗到细的定位过程,可以对大的运动向量进行有效估计,但是需要额外构建塔形结构,这对于背景块和运动向量较小的块来说是没有必要的。

2 判断运动复杂度的快速运动估计算法

为了尽可能正确地反映图像内容的运动剧烈程度,本文将采用2级运动复杂度判断——帧级判断和宏块级判断。

帧级运动复杂度判断(FLD)是对一帧图像的运动复杂度进行整体判断:若运动复杂度较小,则认为该帧运动平缓,适合采用小范围搜索的运动估计算法;否则,认为该帧部分宏块的运动程度较为剧烈,需要进行宏块级判断。

作为帧级判断的补充,宏块级判断可以结合自适应阈值对宏块级运动复杂度进行判断。对于处于静止或者做规则运动的宏块采用小范围搜索的运动估计算法,对于剧烈运动的宏块采用多级分辨率的运动估计算法。

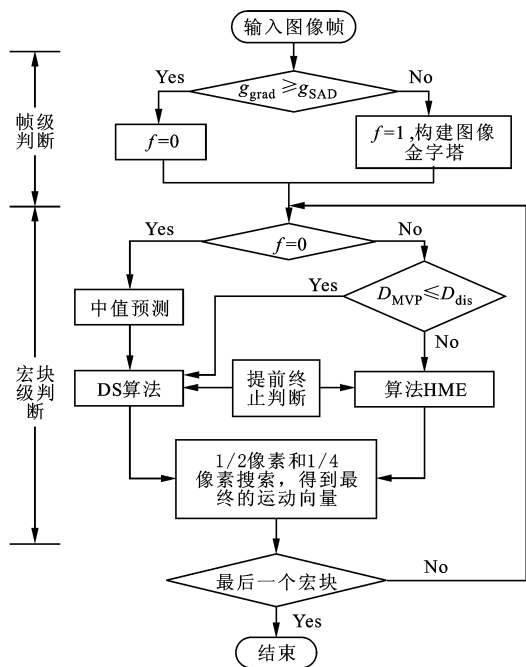
整个算法的判断过程可以用图1的流程图来表示,以下将结合具体实现过程进行说明。

2.1 帧级运动复杂度判断与自适应选择

根据视频序列的时间连续性,可以将前一帧图像的运动复杂度作为当前帧的度量。

运动估计中,宏块的匹配残差主要来自2个方面:图像本身的纹理信息以及运动信息。图像纹理越复杂或者运动越复杂,都会导致宏块匹配残差的增大^[5]。因此,可以根据宏块匹配残差的主要来源进行帧级运动复杂度的判断。

定义单帧宏块匹配残差为该帧图像中宏块与最



g_{grad} : 图像纹理复杂度的梯度信息; g_{SAD} : 1帧图像中宏块与最佳匹配宏块的绝对误差之和; f : 复杂度; D_{MVP} : 预测差异度; D_{dis} : D_{MVP} 的均值

图1 MC-FME算法的流程图

佳匹配宏块的绝对误差之和,即

$$g_{\text{SAD}} = \sum_{i=1}^{N_{\text{mb}}} g_i \quad (1)$$

式中: g_i 是第 i 个宏块与其最佳匹配宏块的绝对误差和; N_{mb} 是该帧图像中的宏块总数。

图像的纹理复杂度可以用图像的梯度信息来描述。本文使用尺寸为 3×3 像素的 Sobel 算子来计算图像像素 $I(i, j)$ 的水平梯度 $G_H(i, j)$ 和垂直梯度 $G_V(i, j)$, 进而得到与该像素对应的梯度信息值

$$g_{\text{grad}}(i, j) = |G_H(i, j)| + |G_V(i, j)| \quad (2)$$

描述整帧图像纹理复杂度的梯度信息 g_{grad} 可由下式得到

$$g_{\text{grad}} = \sum_{(i, j) \in I} g_{\text{grad}}(i, j) \quad (3)$$

式中: I 代表一帧图像像素的集合。

将单帧宏块匹配残差与该帧图像纹理复杂度进行比较,如果纹理复杂度大于或等于单帧宏块匹配残差,则 $f=0$,认为宏块的匹配残差主要来源于图像本身的纹理信息,图像运动较为简单,采用 DS 算法;反之 $f=1$,认为图像运动较为复杂,需要建立参考图像的图像金字塔,并进行后续宏块级判断。

2.2 宏块级运动复杂度判断与自适应选择

帧级判断为运动剧烈的图像帧中,许多宏块仍

有可能处于静止状态或者做规则运动,因此需要进行宏块级运动复杂度判断.

根据视频序列的空间相关性,当前宏块的运动向量预测(MVP)集可以由空间相邻宏块(上、下、左上、右上)的运动向量、前一帧相同位置对应的宏块的运动向量以及运动向量(0,0)构成.

宏块所在区域的运动复杂度可以用 MVP 集中预测值的差异性来判断,该差异性可表示为

$$D_{MVP} = \frac{1}{N_{MVP}} \sum_{i=1}^{N_{MVP}} (|x_{pi} - x_p| + |y_{pi} - y_p|)$$

(4)

式中: N_{MVP} 是 MVP 集中可用的预测值个数; (x_p, y_p) 是 MVP 集中最小匹配误差对应的运动向量预测值; (x_{pi}, y_{pi}) 是 MVP 集中除去 (x_p, y_p) 以外的预测值.

为了有效判断宏块所在区域的运动复杂度,我们使用编码过程中已采用 DS 算法的宏块的 D_{MVP} 均值 D_{dis} 作为阈值. 对一个视频序列进行编码时, D_{dis} 的初始值置为 0, D_{dis} 的值随着编码的帧数增长, 逐渐趋于稳定.

对宏块级运动复杂度进行判断,若 $D_{MVP} \leq D_{dis}$, 说明当前宏块所在区域的运动向量接近,运动方式一致,可以将相邻宏块的运动向量作为预测值,结合文献[12]的方法,提前终止判断,采用 DS 算法在小范围内搜索;反之,认为该宏块所在区域的运动较为复杂,需要采用 HME 算法.

经过以上整像素的运动搜索后,对所得到的运动向量继续进行 1/2 像素和 1/4 像素的搜索,从而得到最终的运动向量. 如果当前宏块不是帧中的最

后一个宏块,则继续进行宏块级判断,否则对下一帧进行帧级判断.

3 实验结果及分析

本算法的实现以 jm86 源代码^[13]为基础,编码为 IPPP 格式,选取的宏块大小为 16×16 像素. 用来进行测试的是标准序列 Football. cif(90 帧)以及 2 段摘自足球比赛的视频 Football 2. cif(100 帧)和 Football 3. cif(100 帧). 这 3 段视频序列中存在准静止与剧烈运动 2 种状态的交替,可以有效地检验新算法的自适应选择性能. 实验中采用平均匹配点数和平均运动估计时间 2 个测度作为算法性能的验证标准. 为了得到运动估计的时间均值,每段视频测试 10 次.

为了验证本文 MC-FME 算法的有效性,在相同条件下对视频序列分别采用 DS、HME、FLD 和 MC-FME 算法进行测试,测试结果如表 1~表 3 所示. 由表 1 中数据可以看出,MC-FME 算法在各种序列中所用的匹配点数和运动估计时间都是最少的. 表 2 说明了 MC-FME 算法对 DS、HME 算法在搜索性能上的提升作用:匹配点数分别减少 5.48% 和 21.43%,运动估计时间缩短 3.20% 和 5.38%,能够达到缩短编码时间的目的. 需要提及的是,在编程实现的时候,MC-FME 算法并没有优化诸如内存、缓冲之类的相关因素,因此搜索点数的减少比运行时间的减少更能说明算法的性能提高. 表 3 为几种运动估计算法的图像质量比较结果,可以看出 MC-FME 算法能够达到与 DS、HME 算法基本相同的图像质量.

表 1 几种运动估计算法的性能比较

视频序列	平均匹配点数				平均运动估计时间/s			
	DS	HME	MC-FME	FLD	DS	HME	MC-FME	FLD
Football. cif	4 613	5 549	4 360	4 428	13.774 5	13.882 9	13.156 1	13.224 7
Football 2. cif	5 278	5 376	4 478	4 495	14.940 2	14.960 6	14.155 8	14.316 0
Football 3. cif	5 442	5 975	4 870	4 958	15.365 5	15.462 2	14.873 1	14.988 5

表 2 MC-FME 算法比 DS 和 HME 算法性能提高的百分比

视频序列	平均匹配点数减少率/%		平均运动估计时间减少率/%	
	DS	HME	DS	HME
Ffootball. cif	5.48	21.43	4.49	5.24
Football 2. cif	15.16	16.70	5.25	5.38
Football 3. cif	10.51	18.49	3.20	3.81

表3 几种运动估计算法的平均峰值信噪比比较

视频序列	峰值信噪比/dB		
	DS	HME	MC-FME
Football. cif	36.39	36.36	36.39
Football 2. cif	36.54	36.49	36.53
Football 3. cif	36.86	36.80	36.85

为了验证 FLD 的有效性,对测试序列只进行帧级判断,即对于 $f=1$ 的帧中宏块全部采用 HME 算法进行搜索,测试结果如图 2 所示. 从图中结果可以看出,结合自适应阈值,帧级复杂度判断可以有效

跟随图像运动信息做出正确的判断. 另外,为了验证宏块级运动复杂度判断对帧级判断性能的提升,在表 1 中也对帧级判断的结果进行了统计,通过与 MC-FME 算法作对比,可以看出宏块级判断是对帧级判断的有效补充,可以进一步减少搜索点数,加快搜索过程.

以上实验结果说明,MC-FME 算法可以较为准确地对图像运动信息进行判断,在相同的图像质量下,搜索性能优于 DS 和 HME 算法.

4 总 结

本文提出了一种判断运动复杂度的快速运动估计算法,通过对视频序列进行帧级和宏块级运动复杂度的判断,选择出适合当前宏块的快速运动估计算法. 实验结果证明,本文所提算法可以在保证图像质量的前提下,有效地缩短编码时间. 值得一提的是,本文算法中的运动复杂度判断阈值是自适应选取的,无需人工交互. 由于宏块划分的复杂性,本文仅做到宏块级判断,对实际编码过程中遇到的不同尺寸的块的判断将在后续工作中进一步展开.

参考文献:

[1] LI Renxiang, ZENG Bing, LIOU M L. A new three-step search algorithm for block motion estimation [J]. IEEE Trans on Circuits Syst Video Technol, 1994, 4 (4): 438-442.

[2] ZHU Ce, LIN Xiao, CHAU L P. Hexagon-based search pattern for fast block motion estimation [J]. IEEE Trans on Circuits Syst Video Technol, 2002, 12 (5): 349-355.

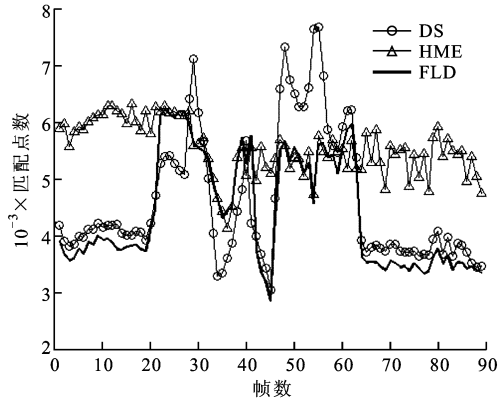
[3] 丁贵广,郭宝龙. 基于线性搜索的快速运动估计算法 [J]. 西安交通大学学报, 2004, 38(2): 136-139.

DING Guiguang, GUO Baolong. New fast motion estimation algorithm based on line search [J]. Journal of Xi'an Jiaotong University, 2004, 38(2):136-139.

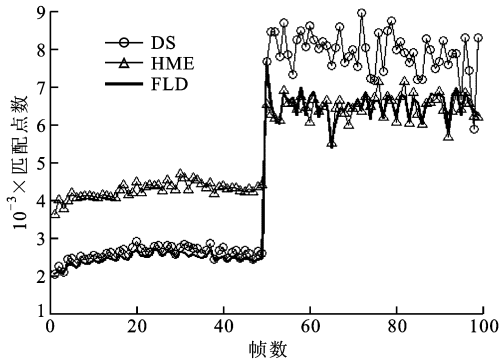
[4] LIU Shaowei, WEI Shou-der, LAI Shanghong. Fast optimal motion estimation based on gradient-based adaptive multilevel successive elimination [J]. IEEE Trans on Circuits Syst Video Technol, 2008, 18(2): 263-267.

[5] DIAZ G, DIAZ M. Adaptive multipattern fast block-matching algorithm based on motion classification techniques [J]. IEEE Trans on Circuits Syst Video Technol, 2008, 18(10): 1369-1382.

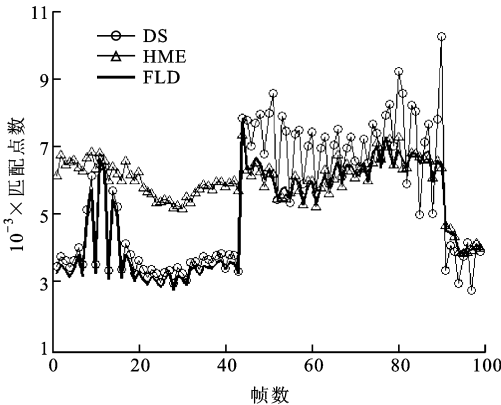
[6] THAM J Y, RANGANATH S, RANGANATH M, et al. A novel unrestricted center-biased diamond



(a)序列 Football. cif(90 帧)



(b)序列 Football 2. cif(100 帧)



(c)序列 Football 3. cif(100 帧)

图2 DS、HME算法和帧级判断基于平均匹配点数的比较

- search algorithm for block motion estimation [J]. IEEE Trans on Circuits Syst Video Technol, 1998, 8(4): 369-377.
- [7] ZHU Shan, MA Kaikuang. A new diamond search algorithm for fast block matching motion estimation [J]. IEEE Trans on Image processing, 2000, 9(2): 287-290.
- [8] WANG Yanni, FAN Yangyu. An adaptive diamond searching algorithm on space-time correlation[C] // Proceedings of IEEE Conf on Image and Graphics. Piscataway, NJ, USA: IEEE, 2007: 220-223.
- [9] NAM K M, KIM J S, PARK R H, et al. A fast hierarchical motion vector estimation algorithm using mean pyramid [J]. IEEE Trans on Circuits Syst Video Technol, 1995, 5(4): 344-351.
- [10] SONG Xudong, CHIANG Tihao, LEE Xiaobing, et al. New fast binary pyramid motion estimation for MPEG2 and HDTV encoding [J]. IEEE Trans on Circuits Syst Video Technol, 2000, 10(7):1015-1028.
- [11] YOON H, KIM H, KIM M, et al. Hierarchical integer pixel and adaptive fractional pixel motion estimation [C]//Proceedings of IEEE Conf on Computer and Information Technology. Piscataway, NJ, USA: IEEE, 2008: 391-395.
- [12] XIE Zhengguang, LIU Yong, LIU Wenjun, et al. A general method for detecting all-zero blocks prior to DCT and quantization [J]. IEEE Trans on Circuits Syst Video Technol, 2007, 17(2): 237-241.
- [13] Joint Video Team. Jm86 开源代码[EB/OL]. (2007-08-01) [2009-01-14]. <http://iphome.hhi.de/suehring/tml/download/old-jm/>.

(编辑 刘杨)