

基于选择预测的自适应运动估计算法

魏星 蒋建国 齐美彬

(合肥工业大学计算机与信息学院,合肥 230009)

摘要 依据图像序列的运动矢量的时空相关性和中心偏移特性,首先对宏块进行类型划分,并利用相关块而非相邻块进行初始点预测的选择预测方法,提出了一种对起点进行选择预测的自适应运动估计算法,而在搜索过程中,则根据不同的图像内容对不同的块灵活地采用不同的搜索模式。实验证明,该算法在大大提高搜索速度的同时,还能保证图像性能,还兼顾了运动估计复杂度和精确度的要求。

关键词 运动估计 运动矢量 时空相关性 起点预测

中图分类号:TN919.81 文献标识码:A 文章编号:1006-8961(2005)07-0873-05

A New Adaptive Motion Estimation Algorithm Based on Selecting Predictive Initial Search Point

WEI Xing, JIANG Jian-guo, QI Mei-bin

(Department of Computer and Information, Hefei University of Technology, Hefei 230009)

Abstract By studying the spatio-temporal correlation and the center-biased characteristic of motion vectors in image sequences, this paper classifies the macro-blocks as different types and takes advantage of the relative blocks instead of neighbour blocks to predict the initial point. A new adaptive motion estimation algorithm based on selecting predictive initial search point is presented, according to which different modes are agilely selected in accordance with image contents. It has been proved with the experiments that this algorithm greatly improves the searching speed and at the same time guarantees image performances with consideration to complexity and accuracy of motion estimation.

Keywords motion estimation, motion vector, spatio-temporal correlation, initial search point prediction

1 引言

运动估计已广泛应用于视频信号处理,是视频压缩中用于减少运动图像时间冗余度的有效方法之一。在ITU-H.26X和ISO MPEG^[1]系列视频压缩标准中,由于运动估计在整个系统中的计算复杂度最大,占运算量的50%以上,因此如何提高运动估计的效率以及速度和精确度已成为目前的研究热点。

目前运动估计的搜索算法有很多种,其中全搜索运动估计算法(full search method, FS)^[2]由于精度高而应用比较广泛,但由于它的运动搜索复杂度

很高,硬件实现难度很大,因此不适合实时应用。为了减少FS算法的运动搜索复杂度和数据读取复杂度,现已出现了许多改进的运动估计算法,如早期的三步搜索算法(three step search, TSS)^[3]、2维对数搜索算法(two-dimensional logarithmic, TDL)^[4]、交叉搜索算法(cross search, CS)^[5]均主要是通过限制搜索位置的数目来减少计算量,但是由于这些算法初始步长较大,易使搜索陷入局部极小,因此在此基础上又出现了利用运动矢量中心偏移特性的新四步法(novel thress step search, NFSS)^[6]和基于块的梯度下降法(block-based gradient descent search, BBGDS)^[7],但它们均没有利用到相邻块的运动相

基金项目:安徽省“十五”科技攻关计划项目(010506E2)

收稿日期:2004-09-07;改回日期:2004-12-08

第一作者简介:魏星(1980~),女。2002年获合肥工业大学计算机与信息学院学士学位,现在该校攻读信号与信息处理专业硕士学位。主要研究方向为视频信号处理、DSP技术应用。E-mail: starwxing@sohu.com

关性。1999 年 10 月被 MPEG4 国际标准采纳并收入验证模型 (verify model, VM)^[8] 的菱形算法 (diamond search, DS)^[9] 具有简单、鲁棒、高效的特点,是现有性能最优的快速搜索算法之一,但由于其始终采用固定的搜索模板,不仅不区分运动类型,而且没有结合图像的内容进行适当的调整,因而导致一定程度上的计算资源浪费,且估计精确度仍与全搜索算法差别较大。

本文利用运动矢量的时间和空间的相关性进行了选择性预测,并结合运动矢量的分布特点和根据图像的内容,在原有菱形算法的基础上提出了一种自适应的快速块匹配运动估计算法。

2 算法原理

2.1 运动矢量的时空相关性和中心偏移特性

众所周知,实际上绝大多数图像序列的运动都很小,如视频电话、视频会议的人肩图像序列,其背景几乎是不变,而且前景的运动也比较缓慢,变化不大。由于它们的运动矢量通常集中分布在搜索窗的中心位置附近,所以进行运动估计时,一般是从块的几何中心处向外进行逐步搜索。常见的基于中心偏移特性的搜索算法中,较为成功的有菱形搜索法。然而如果块内包含有整齐的边缘,或者在运动速度较大的情况下,采用中心偏移的搜索准则,则会产生搜索方向的错误和陷入局部最小的问题。

运动矢量的另一个特性是时空相关性,即在视频序列图像中,物体的运动在时间和空间上具有很强的相关性。在基于块的运动估计中,由于物体运动是连续的,即相邻帧对应位置块的运动矢量具有时间上的相关性,同时由于图像中对象和背景都覆盖相当多的块,且属于同一对象或背景的块的运动场也具有很大的相似性,因此可以利用时间空间上的相邻块来预测当前块的运动矢量,并将其作为当前块的初始搜索位置^[10]。由于这样估计出的运动矢量场较为精确,而且避免了中心偏移搜索法带来的错误方向估计,这样就有效地减小了搜索范围,从而减少了计算量,提高了运算速度。

2.2 基于内容的运动类型划分预处理和初始搜索点的预测

一般而言,由于图像序列中静止和低速活动的区域相对于高速运动的区域要占大部分,如在空间细节不丰富,运动变化小的运动序列(如 Miss

American、Mother and daughter)中,分布在中心点的 3×3 区域窗内,即分布在 $(\pm 1, \pm 1)$ 范围内的运动矢量分别有 93% 和 90%;对于一般的运动物体序列 Foreman,有 87% 的运动矢量分布在 3×3 中心区域窗内;而对于空间细节很丰富,并且运动剧烈的视频序列 Football,分布在 3×3 中心区域窗内的运动矢量仍然有 72%;而高速运动的块,也即运动位移达到搜索区域边界的块,只占全部运动块的大约 10% 到 20%,因此可以将块进行运动类型划分,将那些静止缓动块和高速运动块划分开来,区别对待,以便采用不同的搜索窗和搜索模板,这样对于占大多数的静止缓动块采用小的搜索窗和简单的搜索方法,就可以减少计算量,即可以提高系统编码效率,而对于少部分高速运动块则需扩大搜索范围,并采用更加精确的搜索方法,以提高图像的匹配精确度。

另外,结合前面分析的运动矢量的时空相关性,尚可以依据运动矢量的时空相关性来进行块类型划分,同时可依据运动矢量的空间相关性来进行预测,但在进行块类型划分时,必须先确定一个划分标准。一般而言,与前一帧的像素值的绝对差和 SAD 在一定程度上反映出图像的变化剧烈程度,其定义如下:

$$SAD(\Delta x, \Delta y) = \sum_{x=0}^M \sum_{y=0}^N |f_k(x, y) - f_{k-1}(x + \Delta x, y + \Delta y)|$$

其中, $(\Delta x, \Delta y)$ 为位移矢量, $M \times N$ 是宏块大小, $f_k(x, y)$ 和 $f_{k-1}(x, y)$ 分别为当前帧和上一帧的像素值。当 $\Delta x, \Delta y$ 取 0 时, SAD 即为当前块与前一块相同位置块的绝对差和,其值大,反映图像变化大,运动速度快;其值小,反映图像改变小,图像变化不明显,因此可利用与前一帧相同位置块的 SAD 来进行大致的块类型划分。同时当利用运动矢量的空间相关性来进行预测时,一般传统的算法是利用相邻帧和相邻块来进行预测(如图 1 所示),即用 1, 2, 3, 4, 5 块来预测当前块 (current block, CUR),但是当利用时间相关性时,由于要保存前一帧的所有块的运动矢量 V^{motion} ,这就加大了存储量和系统复杂度,所以在实际预测中不采用相邻帧,只用 1, 2, 3, 4 块来预测,而将运动矢量的时间相关性体现在块类型的划分处理这一过程中。

研究发现,由于图 1 所示的 4 个相邻的预测块,并非都与当前块是一个整体,而是具有因果关系,相邻块有可能是属于不同的运动对象或背景,也就是说,属于不同的图像内容^[11]。如果不加以区分地使用这些不同类型的相邻块来进行运动估计,则会产生无效,甚至是错误的估计,这就相应地增加了计算

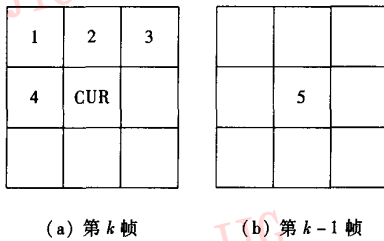


图 1 预测块的位置

Fig. 1 Position of predictive blocks

量,甚至出现误导,如果再采用基于中心偏移的搜索算法,则很容易就陷入了局部最小,因此必须对相邻块进行判别和选择。这里首先定义一个候选预测块,也就是传统的预测方法中所选择的预测块(即图 1 中的 1,2,3,4 块),然后在这些候选预测块中选择出与当前块具有相关性的块参与预测,称为预测块(predictive block, PB),其筛选的原则可以利用前面初步划分的运动类型,对已判断为高速运动的当前块,其缓动的相邻块不参与其初始位置的估计,即缓动块不能作为高速块的预测块。这样就可根据图像的内容来自适应地选择预测块,而不是固定预测块,这不仅可以提高预测的精确度,而且具有基于内容的匹配方法的特点。

2.3 基于选择预测的自适应快速算法

综前分析,在兼顾速度和精度的要求下,本文对不同的块采用不同的搜索范围和不同的搜索方法,如对于占大多数的缓动块采用尽量小的搜索范围和简单的搜索方法,以提高算法速度;对于占少数的高速运动块则将 16×16 的宏块再细分成 4 个 8×8 的小块,分别进行运动估计,以求得更精确的块匹配。同时结合选择预测和菱形搜索算法来对不同的块类型选取不同的搜索模板,就可以取得较好的效果,具体算法描述如下:

事先对块的运动类型做初步划分,其划分的依据是计算当前块与相邻帧对应块的 SAD,并将其定义为 SAD_p (下角 P 代表 predictive),并通过设定两个阈值 T_1 和 T_2 ,将运动类型划分为以下 3 种:

当 $SAD_p < T_1$,则可以认为当前块是缓动块或静止块;

当 $T_1 \leq SAD_p < T_2$,则认为当前块为中速的运动块;

当 $SAD_p \geq T_2$ 时,则判定为高速运动块;

对于第 1 种类型的缓动块,可不作初始位置估计,直接以块中心为搜索中心,搜索窗大小为 3×3 ,只进行一步搜索,且仅和周围 8 个点比较,其中 SAD

最小的就作为该块的运动矢量,据此就可将此类缓动块划分出来,并可一步终止搜索,这既保证了精度,又极大地降低了计算量。

对于中速块,则需先进行初始搜索点预测,然后在预测的起点用具有中心偏移特性的菱形搜索法来找到最佳匹配块。由于阈值选取的缘故,这里不采用候选块筛选的方法,以免破坏图像的整体性,本文选取了如图 1 所示的 4 个点进行运动矢量预测,预测的运动矢量 \hat{V}^{motion} 定义如下:

$$\hat{V}^{motion} = \text{med}(V_1^{motion}, V_2^{motion}, V_3^{motion}, V_4^{motion}) \quad (1)$$

$$\hat{V}^{motion} = (\hat{v}^x, \hat{v}^y)^T$$

$$\hat{v}^x = \text{med}(x_1, x_2, x_3, x_4) \quad (2)$$

$$\hat{v}^y = \text{med}(y_1, y_2, y_3, y_4) \quad (3)$$

式中, \hat{v}^x 、 \hat{v}^y 分别表示预测运动矢量的水平和垂直分量, V_1^{motion} 、 V_2^{motion} 、 V_3^{motion} 、 V_4^{motion} 分别为 1,2,3,4 块的运动矢量。这样在得到预测的初始位置后,即可再采用具有中心偏移特性的菱形法来进行搜索。

对于高速运动的块,这类块只占 10% ~ 20%,是少数块,本文采用传统的最佳 SAD 预测方法,并结合预测候选块的方法来预测运动矢量,并且将 16×16 的宏块细分为 4 个 8×8 的小块分别进行运动估计和块匹配。对宏块的划分见图 2,图中 a, b, c, d 为 4 个 8×8 的小块。

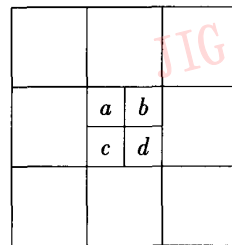


图 2 高速运动宏块的划分

Fig. 2 Partition of fast macroblock

初始运动矢量采用最佳 SAD 结合筛选法来预测,即首先判断相邻块是否为缓动块,如果是,则不参与预测,否则判为预测块 \hat{B}_i ($0 \leq i \leq 4$);然后分别以这些预测块的运动矢量 \hat{V}_i^{motion} ($0 \leq i \leq 4$) 作为预测值,用于计算当前块的 SAD,并选择 SAD 最小者的 \hat{V}^{motion} 作为 4 个 8×8 小块的运动位移的预测值;最后再对这 4 个小块独立的进行菱形搜索。这样一方面可减少计算 SAD 的次数,另一方面也可使预测更加准确,由于灵活地改进了算法,使之具有适应性,因此对高速块能够自动提高匹配的准确性。

3 试验结果

为了验证本文算法的适应性,选取了以下几种不同类型的视频序列:缓慢运动的 Mother and daughter;一般运动的 Foreman;剧烈运动的 Football 进行了运动估计试验。在测试中,分别使用这 3 个序列的前 20 帧,且只采用这些帧的亮度分量进行运动估计。同时还对全搜索法、菱形搜索法和本文自适应选择预测快速算法的效果和速度进行了比较。实验采用的宏块大小为 16×16 ,搜索最大步长为 7,图像格式为 CIF 格式(352×288)。算法使用 Visual C++ 6.0 编程实现。表 1 统计了 3 种算法的平均搜索点数,表 2 比较了用 3 种算法进行运动估计后重建图像的平均峰值信噪比,用以检测算法的性能,并将其作为算法复杂度的测度。图 3 给出了 3 种算法编码后重建图像的主观质量的比较。

表 1 不同视频序列不同搜索算法的平均搜索点数

Tab.1 Average searching points of different algorithms on different sequence

算法	视频序列		
	Mother and daughter	Foreman	Football
全搜索法 FS	225	225	225
菱形搜索法 DS	13.1	15.755	16.812
本文算法	6.2	12.161	14.942

表 2 不同视频序列不同搜索算法的平均峰值信噪比比较

Tab.2 Comparison of average PSNR using several algorithms on different sequence

算法	视频序列		
	Mother and daughter	Foreman	Football
全搜索 FS	43.375	33.923	27.261
菱形搜索法 DS	42.402	31.294	23.916
本文算法	43.315	33.485	25.675



(a)原始图像

(b)菱形搜索算法

(c)本文算法

图 3 Foreman 序列(第 17 帧)编码后重建图像主观质量比较

Fig. 3 Image performances of several algorithms on Foreman sequence(17th frame)

由表 1、表 2 可见,对于运动较小的图像序列 Mother and daughter,本文算法搜索点数明显下降,是菱形搜索算法的一半以上,不仅提高了运算速度,同时仍然能使重建图像保持较高的信噪比;而对于高速运动细节较多的图像序列 Football,在保证重建图像峰值信噪比 PSNR 比菱形算法高的情况下,搜索点数仍低于菱形算法,可见本文算法能保证匹配精度和图像质量,而从图 3 的主观质量可见,本文算法的主观质量也比菱形算法好,其重建图像接近原始图像。

4 小结

本文提出了一种基于起点预测的搜索算法,该算法融入了基于内容的特点,在充分利用运动矢量

的中心偏移和时空相关性特点的基础上,提出了选择性预测的自适应运动估计算法,并针对不同块类型采用不同的搜索模板,尤其是在高速运动块上设计了再细分小块的搜索方法。实验结果表明,本文的选择预测自适应快速算法在保持图像性能的基础上,还有效地提高了运算速度。总体而言,本文算法由于抛弃了以往对整幅图像采用相同搜索策略进行运动估计匹配的方法,因而对不同类型的图像具有适应性,相应地也消除了传统算法所带来的大量计算冗余,这样就在运动估计的精确度和复杂度这个矛盾问题上找到了一个平衡点。

参考文献 (References)

- 1 Ji Wen-ping, Guo Bao-long. Overview of the international digital

- video coding standards [J]. *Application Research of Computers*, 2003, 20(4):1~5. [计文平, 郭宝龙. 数字视频压缩编码的国际标准[J]. *计算机应用研究*, 2003, 20(4):1~5.]
- 2 ISO/IEC/JTC. Information technology. coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s. part 2: video[S]. ISO/IEC 11172-2-1993, ISO/IEC/JTC, 1993.
- 3 Ding Gui-guang, Ji Wen-ping, Guo Bao-long. *Visual C++ 6.0 Digital Image Coding* [M]. Beijing: China Machine Press, 2004: 101~103. [丁贵广, 计文平, 郭宝龙等. *Visual C++ 6.0 数字图像编码* [M]. 北京: 机械工业出版社, 2004: 101~103.]
- 4 Jain J, Jain A. Displacement measurement and its application in interframe image coding [J]. *IEEE Transactions on Communication*, 1981, 29(12):1799~1808.
- 5 Ghanbari. The cross-search algorithm for motion estimation [J]. *IEEE Transactions Communication*, 1990, 38(7):950~953.
- 6 PO M, MA W C. A novel four-step algorithm for fast block motion estimation [J]. *IEEE Transactions on Circuits Systems Video Technology*, 1996, 6(3):313~317.
- 7 Liu L K, Feig E. A block-based gradient descent search algorithm for block motion estimation in video coding [J]. *IEEE Transactions on Circuits Systems Video Technology*, 1996, 6(4):419~422.
- 8 ISO/IEC. MPEG4 Video Verification Model Version 13.0 [S]. JTC1/SC29/WG11 N2687, ISO/IEC, 1999.
- 9 Zhu S, Ma K-K. A new diamond search algorithm for fast block matching motion estimation [J]. *IEEE Transactions on Image Processing*, 2000, 9(2):287~290.
- 10 Xiang You-jun, Guo Bao-long. Fast Block Matching Algorithms for Motion Estimation [J]. *Computer Engineering*, 2003, 29(13):62~64. [向友君, 郭宝龙. 运动估计快速块匹配算法 [J]. *计算机工程*, 2003, 29(13):62~64.]
- 11 Yang Cheng-jun, Wei Chun-cheng, Chen Ming, et al. Spatial and temporal motion vector prediction and fast motion estimation [J]. *High Technology Letters*, 2002, (4):27~31. [杨成军, 魏春城, 陈铭等. 视频编码中基于时空信息运动估计的研究 [J]. *高技术通信*, 2002, (4):27~31.]