

基于最小匹配误差方向预测的快速半像素运动估计^{*})

董海燕 张其善

(北京航空航天大学电子信息工程学院 北京 100083)

摘要 为了减小半像素搜索的计算量,本文提出了一个基于最小匹配误差方向预测的快速半像素运动估计算法。本文提出的算法利用亚像素搜索窗内的匹配误差单峰曲面的特性来预测半像素搜索区域中最小匹配误差方向,从而避免了大量不必要的匹配运算量。实验结果表明,对于各种不同运动程度和空间细节的视频序列,本文提出的算法在保证和半像素全搜索法相同图像质量的同时,平均节省 73% 的计算量,很适合实时应用。

关键词 视频压缩,运动估计,块匹配,半像素搜索

A Fast Half-Pixel Motion Vector Search Algorithm

DONG Hai-Yan ZHANG Qi-Shan

(School of Electronic and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083)

Abstract A minimal matching error direction prediction-based fast half-pixel motion vector search algorithm is proposed in this paper in order to reduce computation load of half-pixel search. Based on the single peak characteristic of half-pixel error matching function inside search grid, the algorithm predicts minimal matching error direction with the help of SAD values comparison results of four integer-pixel points around integer-pixel motion vector, so that a great number of unnecessary computations for matching search can be avoided. The experimental results reveal that, to all kinds of video sequences, the proposed algorithm can obtain almost the same video quality as that of the half-pixel full search algorithm with reduced average 73% computation cost, which is very useful for real-time applications.

Keywords Video compression, Motion estimation, Block matching, Half-pixel search

1 引言

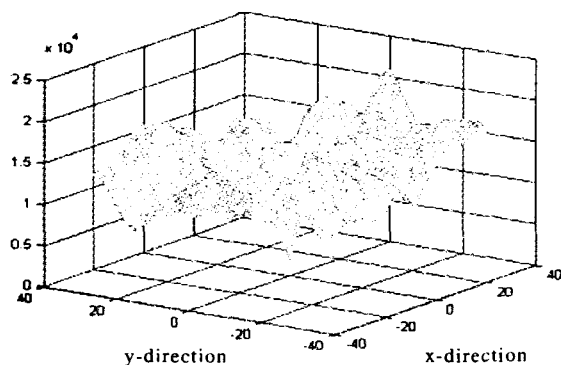
运动估计是视频编码系统的关键部分,同时也是整个视频编码器中计算量最大、耗费时间最多的部分^[1]。运动估计性能的优劣直接影响到整个视频编码器的运行效率和整个视频序列的重构质量。

由于自然物体运动的连续性,相邻两帧之间物体的运动矢量不是以整像素为基本单位,典型的运动估计主要包括两个部分:整像素运动估计和半像素运动估计。目前的快速搜索算法主要集中在整像素运动矢量搜索上,而半像素运动矢量搜索一般采用全搜索方法,即在整像素周围的 8 个半像素插值点以及整像素运动矢量点中找出其中最小块匹配误差点作为半像素精度的运动矢量。由于采用了快速整像素搜索算法,每个块的整像素搜索可以减小到 10 个点以下^[2,3]。而每

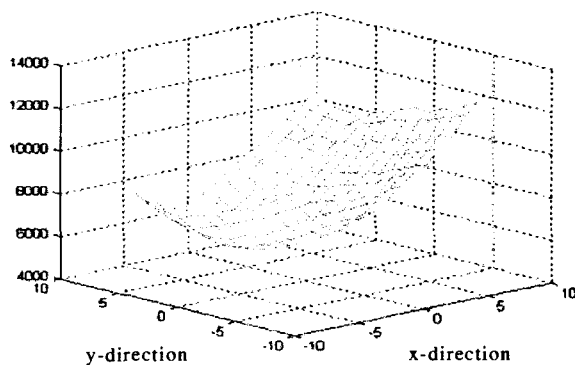
个块的半像素全搜索算法包括:8 个半像素点的插值运算和 9 个点的块失真匹配运算。在整个运动估计中,随着快速整像素运动搜索算法的使用,半像素搜索运算量所占的比例不断突显。文[4]中提出的方法需要进行大量额外的加法除法运算,导致总的计算量很高。为了减小半像素搜索的计算量,研究快速半像素搜索算法有着非常重要的实用价值。

2 半像素搜索窗误差匹配曲面特性

由于搜索窗内细节特性的差异,物体运动的剧烈程度和不规则性,整像素搜索窗内的误差匹配曲面呈现出多峰曲面的特点,如图 1(a)所示。因此,基于误差匹配单峰曲面假设的算法常会落入局部最小点而造成率失真性能的急剧下降,最终导致不适于整像素搜索策略^[5]。



(a) 整像素搜索窗



(b) 半像素搜索窗

图 1 不同搜索窗内的误差匹配曲面

^{*}) 基金项目:国家自然科学基金(60372018)。董海燕 博士,研究方向为视频编码,多媒体系统和图像处理。

若用于运动估计的块大小为 $W \times H$, 搜索范围为 d , 那么, 由于半像素运动估计是在前面获得的整像素运动矢量周围通过线性内插得到的 8 个半像素候选点之间进行, 整个半像素搜索面积为 1, 而整像素搜索面积为 $(W+2d)(H+2d)$, 是前者的 $(W+2d)(H+2d)$ 倍。同时, 相对于整像素搜索窗内误差匹配曲面不规则的特点, 利用相邻整像素数据内插获得的半像素数据具有很强的相关性。因此, 不同于整像素搜索窗, 半像素搜索窗内的误差匹配曲面普遍地具有单峰曲面的特性, 如图 1(b) 所示。

本文采用绝对误差和(SAD)作为块失真匹配函数, 对多个不同视频内容的测试序列的半像素搜索区域误差匹配函数进行了统计。如表 1 所示, 平均 90% 的半像素搜索区域具有这种单峰曲面特性。该单峰曲面特性为本文快速半像素搜索算法的提出提供了客观依据。

表 1 具有单峰特征的半像素搜索区域

测试序列	总块数	单峰块	比率/%
Akiyo	150246	147917	98.45
Mother&Daughter	147825	143198	96.87
Coastguard	129842	120441	92.76
Foreman	136546	119272	87.35
Stefan	136858	111799	81.69
Mobile Calendar	145613	115645	79.42

3 快速半像素运动矢量搜索算法

从上面的分析可知, 半像素匹配误差曲面具有单峰特性。即在半像素搜索区域, 如果某个点越靠近最小匹配失真点, 则这个点的误差匹配就越小。利用半像素搜索的这个特性, 如果能容易找到匹配失真的下降方向, 就能快速找到可能的匹配失真最小的候选点, 从而避免不必要的匹配搜索和运算, 加速半像素运动估计的速度。

如图 2 所示, 标记点 0 为参考帧中整像素运动搜索到的运动矢量指向的位置, 标记点 H_1, H_2, V_1, V_2 分别为以标记点 0 为中心水平向和垂直向的四个整像素点, 标记点 1, 2, 3, 4, 5, 6, 7, 8 分别为整像素运动矢量周围的 8 个半像素点。从运算复杂度和便于硬件实现考虑, 本文采用绝对误差和(SAD)作为块失真匹配函数。

在整像素搜索过程中, 整像素运动矢量周围的四个整像素点 H_1, H_2, V_1, V_2 的 SAD 值都已经得到, 而且半像素搜索窗内的误差匹配曲面具有单峰曲面的特性, 从而可以方便利用已知的这四个整像素点的 SAD 值来预测匹配失真的下降方向。

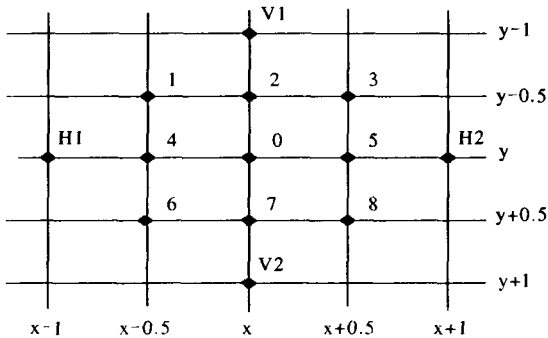


图 2 整像素和半像素的位置

本文提出的快速半像素运动矢量搜索算法具体思想为:

(1) 比较四个整像素点 H_1, H_2, V_1 和 V_2 的 SAD 匹配误差值, 得到一个 SAD 匹配代价最小点 $\min_0$ 和代价亚小点 $\min_1$;

(2) 根据得到的匹配代价最小点 $\min_0$ 和代价亚小点 $\min_1$ 以及它们之间的位置关系, 预测 SAD 匹配代价下降的方向;

(3) 根据预测的 SAD 匹配代价函数下降方向, 选择位于这个方向的两个半像素点进行匹配计算。如果 $\min_0$ 和 $\min_1$ 在同一水平或垂直线上, 则 SAD 匹配代价函数下降的方向位于这条线上, 就选择这条线上的两个半像素点(例如, $\min_0 = H_1$ 和 $\min_1 = H_2$, 则选择 4, 5 两点)。如果 $\min_0$ 和 $\min_1$ 不在同一水平或垂直线上, 则 SAD 匹配代价函数下降的方向位于 $\min_0$ 和 $\min_1$ 所在的象限, 而且更接近 $\min_0$ 点, 就选择位于这个象限的接近 $\min_0$ 点的两个半像素点(例如, $\min_0 = V_2$ 和 $\min_1 = H_1$, 则选择 6, 7 两点)。

利用这个算法, 只需搜索位于预测方向的两个候选点, 避免了不必要的半像素搜索匹配计算量, 从而在很大程度上减小了半像素运动估计的计算复杂度。表 2 为经过预测保留下来的半像素搜索点。

表 2 经过预测保留下来的半像素搜索点

条件	需要搜索的半像素点
$(\min_0 = H_1) \& \& (\min_1 = H_2)$	4, 5
$(\min_0 = H_1) \& \& (\min_1 = V_1)$	1, 4
$(\min_0 = H_1) \& \& (\min_1 = V_2)$	4, 6
$(\min_0 = H_2) \& \& (\min_1 = H_1)$	5, 4
$(\min_0 = H_2) \& \& (\min_1 = V_1)$	3, 5
$(\min_0 = H_2) \& \& (\min_1 = V_2)$	5, 8
$(\min_0 = V_1) \& \& (\min_1 = H_1)$	1, 2
$(\min_0 = V_1) \& \& (\min_1 = H_2)$	2, 3
$(\min_0 = V_1) \& \& (\min_1 = V_2)$	2, 7
$(\min_0 = V_2) \& \& (\min_1 = H_1)$	6, 7
$(\min_0 = V_2) \& \& (\min_1 = H_2)$	7, 8
$(\min_0 = V_2) \& \& (\min_1 = V_1)$	2, 7

从表 2 可以看出: 需要搜索的半像素点数从原来的 8 个点减少到 2 个。同时, 本文算法无需进行整个图像的半像素插值运算, 只进行这两个半像素点的插值运算。若运动搜索的块大小为 $W \times H$, 水平、垂直方向的半像素搜索点和对角线方向的半像素搜索点的运算次数分别为 $5 \times W \times H - 1$, $7 \times W \times H - 1$ 。如果采用半像素全搜索, 那么每个块需要进行 $48 \times W \times H - 8$ 次运算, 而若采用本文块速算法, 每个块只需 $10 \times W \times H - 2$ 或 $12 \times W \times H - 2$ 次运算。因此, 每个块可节省 6 个半像素搜索点的失真函数运算量和插值运算量, 从而大幅度提高半像素运动估计速度。

4 实验结果和分析

为测试本文提出算法的性能, 本文采用了六个 MPEG-4 标准测试序列。其中, Akiyo 和 Mother&Daughter 属于空间细节不丰富、运动变化小的视频序列; Coastguard 和 Foreman 属于空间细节比较丰富、运动变化小或空间细节不丰富、运动量中等的视频序列; Stefan 和 Mobile Calendar 属于空间细节很丰富、运动变化中等或空间细节比较丰富、运动变化大的视频序列。

为了评价算法性能, 本文采用的评价指标为峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)、编码字节数、搜索点数

和节省的运算量。表3为本文提出的快速算法和半像素全搜索算法的实验数据比较。

表3 采用不同半像素搜索法的实验数据比较

测试序列	算法	运算次数	节省运 算量/%	PSNR /dB	编码字 节数/ Bytes
Akiyo	全搜索	276471099	0.000	40.392	84571
	本算法	72423676	73.804	40.105	84161
Mother&Daughter	全搜索	265636016	0.000	36.200	83138
	本算法	67191033	74.706	35.789	83789
Coastguard	全搜索	264708767	0.000	29.331	85855
	本算法	72849272	72.479	29.098	85461
Foreman	全搜索	265850604	0.000	33.528	82237
	本算法	72532227	72.717	33.118	81976
Stefan	全搜索	254453213	0.000	26.832	79450
	本算法	67462969	73.487	26.291	79560
Mobile Calendar	全搜索	223831318	0.00	23.765	84682
	本算法	58199433	73.999	23.494	84985

分析 PSNR 可以得出:与半像素全搜索法相比,本算法可以提供几乎相同的视频质量,PSNR 平均下降 0.34dB。分析编码码流长度可以得出:码流字节数平均增长仅为 0.78%。从运算次数和节省运算量可以看出:与半像素全搜索算法相比,本文算法平均节省 73% 的计算量。比较结果表明:本文算法对重建视频质量和编码长度的影响很小,可以显著提高搜索速度。根据表 3 中六个 MPEG-4 标准测试序列的测试结果还可以看出:本算法具有很好的鲁棒性,对不同运动程度和空间细节的视频内容均可以获得很好的效果。

(上接第 214 页)

标、多目标;室内、室外;简单和复杂背景的视频序列都可以得到较好的检测效果,而且本算法不需要考虑光源的照射方向,场景的几何特征以及考虑如何对运动目标建模等问题。在执行算法时,需要调整的参数少,抗干扰能力强,可以较精确地提取出运动目标。此外,该算法复杂度较低,算法的运行速度不与目标数成正比,易于实现实时运动图像处理。

参考文献

- Prati A, Mikic I, Trivedi M M, Cucchiara R. Detecting Moving Shadows: Algorithms and Evaluation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2003, 25(7): 918~923
- Kilger M. A shadow handler in a video-based real-time traffic monitoring system. In: *Proc. Applications of Computer Vision*, Nov. - Dec. 1992. 11~18
- Stauder J, Mech R, Ostermann J. Detection of Moving Cast Shadows for Object Segmentation. *IEEE Trans. Multimedia*, 1999, 1(1): 65~76
- Haritaoglu I, Harwood D, Davis L S. W4: Real-Time Surveillance of People and Their Activities. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2000, 22(8): 809~830
- Cucchiara R, Grana C, Piccardi M, Prati A. Detecting Moving Objects, Ghosts, and Shadows in Video Streams. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2003, 25(10): 1377~

结论 快速半像素运动矢量搜索算法利用半像素搜索窗内的误差匹配曲面单峰曲面的特性,通过比较整像素运动矢量周围四个整像素点的 SAD 值来预测最小匹配误差的下降方向,从而达到大大提高半像素搜索速度的目的。实验结果表明,对于各种不同运动程度的视频序列,本算法都可以获得和半像素全搜索法几乎相同的图像质量,同时平均节省 73% 的计算量。总之,本文提出的算法在搜索速度和搜索效果两方面更具优势,尤其适用于实时性要求比较高的场合。

参考文献

- Fukunaga S, Nakaya Y, Nagumo S S T. MPEG-4 video verification model version 14. 0 [S]. ISO/IEC JTC1/SC29/WG11 MPEG99/N2932, Victoria, Australia, Oct. 1999
- Tourapis A M, Cl. Au O, Liou M L. Predictive motion vector field adaptive search technique (PMVFAST): enhancing block-based motion estimation [A]. In: *Proc. SPIE Vol. 4310 [C]. Visual Communications and Image Processing*, 2001. 883~892
- LI Bo, LI Wei, TU Yaming. A fast block-matching algorithm using smooth motion vector field adaptation search technique [J]. *Journal Comput. Sci. & Technol*, 2003, 18(1): 14~21
- Du Cheng, He Yun, Zheng Junli. PPHPS: A Parabolic Prediction-Based, Fast Half-Pixel Search Algorithm for Very Low Bit-Rate Moving-Picture Coding [J]. *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY*, 2003, 13(6): 514~518
- Zhou Bo, Chen Jian. A fast two-step search algorithm half-pixel motion estimation [J]. *ICECS IEEE*, 2003. 611~614
- 1342
- Mikic I, Cosman P, Kogut G, Trivedi M M. Moving Shadow and Object Detection in Traffic Scenes. In: *Proc. Pattern Recognition*, Sept. 2000. 321~324
- Nadimi S, Bhanu B. Physical Models for Moving Shadow and Object Detection in Video. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2004, 26(8): 1079~1087
- Kato J, Watanabe T, Joga S, Rittscher J, Blake A. An HMM-Based Segmentation Method for Traffic Monitoring Movies. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2002, 24(9): 1291~1296
- Wren C R, Azarbayejani A, Darrell T, Pentland A P. Pfunder; Real-Time Tracking of the Human Body. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1997, 19(7): 780~785
- Kumar P, Sengupta K, Lee A. A comparative study of different color spaces for foreground and shadow detection for traffic monitoring system. In: *The IEEE 5th Intl. Conf. on Intelligent Transportation Systems*, 2002. 100~105
- Branca A, Attolico G, Distanto A. Cast shadow removing in foreground segmentation. *The 16th International Conference on Pattern Recognition*, Aug. 2002. 1: 214~217
- Grest D, Frahm J-M, Koch R. A Color Similarity Measure for Robust Shadow Removal in Real-Time. *VMV. Munich, Germany*, Nov. 2003