

一种基于 AVS 的亚像素快速运动估计算法

周 城 谭毅华 田金文 柳 健

(华中科技大学图像所图像处理 and 智能控制教育部重点实验室, 武汉 430074)

摘 要 AVS 是我国自主制定, 拥有自主知识产权的音视频编码技术标准。AVS 采用了可变大小块以及 1/4 像素精度的运动估计技术以提高编码效率。基于 AVS 参考软件提出了一种增强亚像素自适应范围搜索算法。该算法主要使用了小菱形窗的搜索策略和阈值判决方法。与亚像素全搜索算法比较, 在平均性能仅下降 0.009 3dB 的情况下, 减少了 30.25% 的搜索点数。与 AVS 参考软件中的快速亚像素运动估计算法相比较, 在平均性能提高 0.022 1dB 的同时, 节省了 8.78% 的搜索点数。

关键词 AVS 视频编码 亚像素运动估计

中图分类号: TN919.81 文献标识码: A 文章编号: 1006-8961(2007)09-1520-05

A Fast Algorithm of Sub-pixel Motion Estimation Based on AVS

ZHOU Cheng, TAN Yi-hua, TIAN Jin-wen, LIU Jian

(State Key Laboratory for Education Ministry for Image Processing and Intelligent Control, Institute for Pattern Recognition and Artificial Intelligence, HUST, Wuhan 430074)

Abstract AVS is a new audio and video coding standard in China, in which variable block sizes and quarter-pel motion compensation have been applied to improve coding efficiency. In this paper, an adaptive range fractional pel search algorithm is proposed to further speed up the encoding process and reduce the computation complexity. Small diamond search and threshold judgment are used to this algorithm. Compared with the half-pel full search algorithm, experiments show that the proposed algorithm can reduce the sub-pel search points by 30.25% on average with the limited performance lost about 0.009 3dB. And compared with the fast search algorithm in AVS reference software, experiments show that the proposed algorithm can reduce the sub-pel search point by 8.78% on average with the performance achieve a gain up to 0.022 1dB.

Keywords AVS, video coding, sub-pel motion estimation

1 引 言

AVS(audio video coding standard)是由国家信息产业部科学技术司于 2002 年批准成立的数字音视频编解码技术标准工作组所制定的具备自主知识产权的我国第二代信源编码标准^[1]。目前, 标准中涉及视频压缩编码的有两个独立的部分: 第二部分(AVS1—P2)主要针对高清晰度数字电视广播和高密度存储媒体应用, 第七部分(AVS1—P3)主要针对低码率、低复杂度、较低图像分辨率的移动

多媒体应用。AVS 当中独特的核心技术包括: 8 × 8 整数变换、量化、帧内预测、1/4 精度像素插值、特殊的帧内预测运动补偿、2 维熵编码、去块效应环内滤波等。

在 AVS、H.264 等标准中, 帧间运动估计和帧内预测的计算复杂度较高, 特别是运动估计部分, 通常约占全部编码时间的 50% 以上。降低这两部分的计算量可以有效地减少编码的复杂度和加快编码速度。针对 AVS—P2 部分的运动估计算法, 本文提出了一种新的亚像素运动估计算法。该方法利用视频序列在时间和空间的运动矢量所呈现的统计规律,

基金项目: 国家“863”基金项目(2004AA119010-6)

收稿日期: 2006-04-02; 改回日期: 2006-06-02

第一作者简介: 周城(1979 ~), 男, 华中科技大学模式识别与智能系统专业博士研究生。研究方向为视频编码、图像通信。E-mail: zcfriends@gmail.com

通过预测后续运动趋势加速运动估计过程,尽量减少搜索点数,同时力求保持其编码效率。

2 亚像素运动估计算法和运动分布特点

运动估计的基本原理是利用帧间的时空相关性,减小冗余度。视频序列相邻两帧间有很强的相关性,将前后两帧相减(移动物体作相应位移)得到帧间差值,此差值越小,所需码字就越短。运动估计算法一般分为两类:像素递归法和块匹配法。块匹配法是目前最常用的方法。例如 AVS 和 H.264 都采用了块匹配法。块匹配一般采取绝对误差和(sum of absolute difference, SAD)作为匹配准则,其计算过程如下式所示:

$$SAD(i, j) = \sum_{m=1}^M \sum_{n=1}^N |f_i(m, n) - f_{i-1}(m + i, n + j)|$$

其中, M, N 为搜索区域大小, $f_i(m, n)$ 为搜索起点,在待搜索区域的所有候选点中求得最小 SAD 值像素点处,即为块匹配最佳点。

2.1 AVS 中的亚像素运动估计算法

在 AVS 的运动估计算法中,一个 16×16 的宏块 MB(macroblock)可以进一步划分成若干个子块进行补偿。亮度子块的大小包括 $16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8$ 块模式。由于采用 4:2:0 的采样方式,色度子块的大小包括 $8 \times 8, 8 \times 4, 4 \times 8, 4 \times 4$ 块模式。对于每一种块模式都要进行相应的运动估计,在每种模式中都要分别进行整像素、半像素和 1/4 像素搜索。在这样得出的众多 SAD 中,最小 SAD 值所对应的运动矢量即为该次运动估计的结果,也由此导致运算量成倍增加。

UMHexagonS 算法是一种优秀的整像素快速运动估计算法,其全称是非对称十字型多层次六边形格点搜索,该算法能高速完成整像素运动估计,同时保持较高的编码效率^[2]。AVS 与 H.264 标准参考软件均采用该算法进行整像素运动估计。在获得整像素最佳匹配点后,对 $16 \times 16, 16 \times 8, 8 \times 16$ 块分别进行 1/4 像素精度的亚像素全搜索算法(HPFS)来查找最佳匹配点,对 8×8 块采用中心偏倚的亚像素搜索算法(CBFPS)进行 1/4 像素精度运动估计以提高搜索速度,其中 HPFS 算法示意图如图 1 所示。从图 1 可以看出,对每个 $16 \times 16, 16 \times 8$ 与 8×16 块,都需要首先搜索全部 8 个 1/2 像素点,然后搜索

8 个 1/4 像素点,因而每次亚像素运动估计都需要搜索 16 个点,降低了整体运动估计速度。显而易见,HPFS 算法虽然性能是最佳的,但速度最慢。

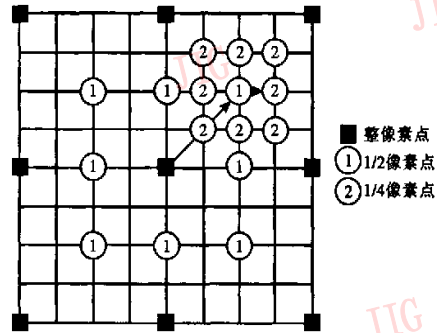


图 1 HPFS 算法

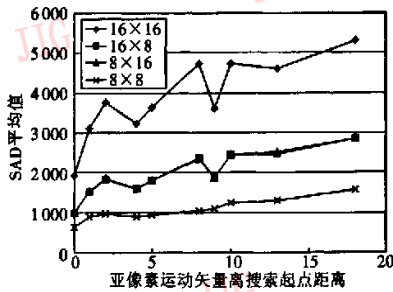
Fig. 1 HPFS algorithm

2.2 亚像素运动分布特点

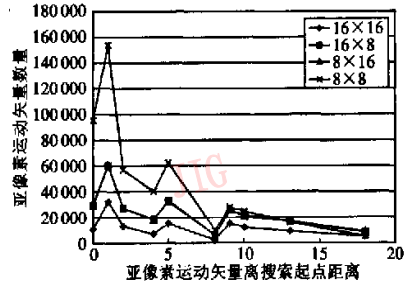
与整像素区域不同,亚像素区域的 1/2 与 1/4 像素点都是由相邻整像素通过线性插值得到,因此亚像素点之间的相关性比较大。亚像素搜索窗内的误差匹配曲面普遍地具有单峰曲面的特性,统计显示平均约 90% 以上的亚像素搜索块具有这种特性^[3]。在亚像素运动估计过程中,可利用这种特性设计合适的亚像素运动估计优化算法,使搜索过程快速定位到该曲面的单峰上。

最重要的问题便是确定亚像素误差曲面的单峰一般出现在哪个位置?是否和初始搜索点有一定的联系?在对大量视频序列的亚像素运动的分布进行统计分析后,得到了一定的统计规律。本文仅选取运动剧烈的 football CIF 序列与细节丰富的 mobile CIF 序列这两个典型视频序列来对此规律进行说明。统计实验包括统计每次亚像素运动估计的结果,包括搜索块大小、每次亚像素运动估计得到的最小 SAD 值以及对应的亚像素最佳运动矢量的数量和距离搜索起点的距离(用距离平方和表示)。统计结果如图 2 所示。

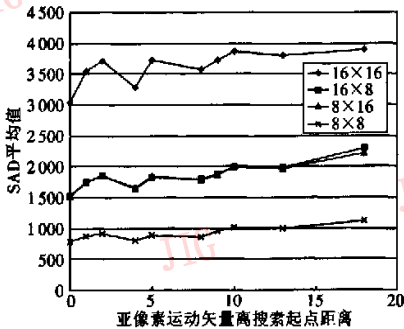
从这种典型序列的统计结果中可以看出以下基本规律:首先,每次运动估计得到的最佳亚像素运动矢量集中在搜索起点附近,离搜索起点越远,运动矢量数量越少,从图 2 可以看出,有 50% 左右的运动矢量为搜索起点附近的 8 个 1/4 像素点和起点;其次,距离搜索起点越近, SAD 值越小,反之越大;另外,由于搜索块的大小不同, SAD 平均值和最佳运动矢量数量之间亦有倍数关系。



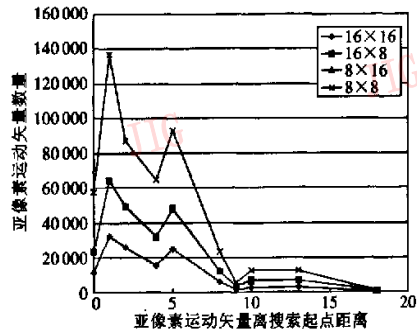
(a) 美式足球序列 120 帧 SAD 平均值分块统计



(b) 美式足球序列 120 帧运动矢量数量分块统计



(c) 汽车序列 120 帧 SAD 平均值分块统计



(d) 汽车序列 120 帧运动矢量数量分块统计

图 2 美式足球、汽车序列统计结果

Fig. 2 Statistical results of football and mobile sequence

3 亚像素运动估计优化算法

针对上述亚像素运动分布特点,同时考虑到 AVS 的特性,本文提出了一种新的混合快速算法:亚像素自适应范围搜索算法 (adaptive range fractional pel search, ARFPS),主要采用下面几种优化准则:

- (1) 根据整像素搜索结果判断亚像素搜索区域大小;
- (2) 采用小菱形搜索窗进行 1/2 和 1/4 像素运动估计,并根据菱形窗搜索结果判断下一步搜索点位置;
- (3) 在亚像素搜索中加入依匹配块大小自适应变化的阈值判决,以此减少不必要的搜索步骤。

3.1 中心偏倚的搜索策略

根据统计规律,运动矢量集中在搜索起点附近,因此在整像素运动估计后,可设置阈值 T_1 ,若当前 $SAD < T_1$,则可直接进入搜索起点附近的 8 个 1/4 像素点进行搜索,此时至少可节省一半亚像素搜索点数。其原理是整像素运动估计的匹配程度已经很高,同时也使得亚像素运动估计的结果有极大概率落在搜索起点附近的 1/4 像素点之中。

3.2 小菱形搜索策略

统计实验表明,视频序列的运动集中在水平和垂直方向,因此各种菱形搜索窗更加符合视频序列的运动规律。Yin 等人^[4]提出了一种根据菱形搜索窗的搜索结果来判断下一步搜索位置的算法,但该算法忽视了一般视频序列在进行亚像素运动估计时运动矢量集中在初始搜索点附近的规律。基于此,本文提出了新的基于小菱形窗的搜索策略。

当 1/2 像素或 1/4 像素小菱形窗搜索完成后,算法记录当前的最佳匹配点和次佳匹配点,如图 3 所示,此时共有如下 4 种情况:

(a) 最佳匹配点位于搜索起点时,说明搜索结果是中心偏倚的。此时若 SAD 小于阈值 T_2 ,则有极大可能最佳匹配点即为搜索起点,可直接终止搜索过程。若不满足条件,且为 1/2 像素,则可直接进入 1/4 像素进一步搜索。

(b) 最佳匹配点在外,而次佳匹配点为搜索起点。此时说明运动有向外趋势,下一步搜索最佳匹配点旁两个同精度像素点进行验证即可。

(c) 最佳匹配点与次佳匹配点都不为搜索起点

且相邻。此时运动有斜方向趋势,下一步搜索两点之间的同精度像素点进行验证即可。

(d) 最佳匹配点与次佳匹配点都不为搜索起点且相对。此时无法判断运动趋势,则需将剩余同精度像素点全部搜索完成后才能确认最佳匹配点。

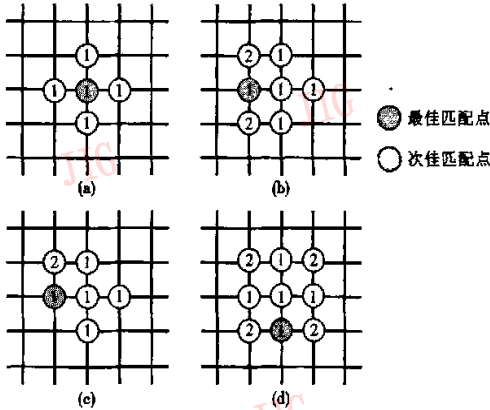


图3 小菱形窗搜索策略

Fig.3 Search strategy of small diamond search pattern

3.3 阈值判决

每搜索一个点后,若当前 SAD 值已经相当小,说明匹配程度已经很高,像素匹配精度已经达到要求,即可提前终止搜索或由 1/2 像素精度搜索进入 1/4 像素精度搜索。

采用上述优化准则后,与图 1 所示的参考模型中运动估计算法相比,本文提出的算法——ARFPS 在各级运动搜索的中途终止了大量的不必要的运动

搜索计算,加速了运动估计的计算,同时充分体现了视频序列的运动规律。

4 仿真结果和分析

本文算法的仿真平台为 AVS-P2 参考软件 rm52d,采用了 14 个 AVS 标准测试序列。这些序列基本涵盖了各种类型的视频序列特征,因此可认为测试结果具有普遍意义。测试条件包括帧率为 30Hz,编码 120 帧,帧间预测最多使用两个参考帧,采用 IBBP 序列,其中 B 帧为双向预测,使用 Hadamard 变换,CAVLC 熵编码,不使用 RD 优化和码率控制。性能评价参数包括综合峰值信噪比 PSNR,码率以及亚像素运动估计搜索点数。其中 PSNR 是按视频编码性能通用评价方法计算每个序列的 4 个 QP(27,30,35,40) 仿真结果的综合平均值所得^[5]。仿真的微机配置为 P4 2.0G 512M RAM。

为评价本文算法性能,对于每个序列都完成了 3 组仿真。这 3 组仿真的整像素运动估计部分统一使用 UMHexagonS 算法。亚像素运动估计部分,第 1 组采用亚像素全搜索算法 (HPFS),第 2 组采用 AVS 参考软件的默认方法,即 HPFS 算法和中心偏倚的亚像素搜索算法 (CBFPS) 的混合,第 3 组则采用本文提出的亚像素自适应范围搜索算法。表 1 中的结果为 ARFPS 算法分别对比 HPFS 算法以及 HPFS、CBFPS 混合算法在不同性能指标上的增益。

表 1 ARFPS 算法仿真实验结果

Tab.1 Simulation results of ARFPS algorithm

序列	ARFPS		ARFPS vs HPFS			ARFPS vs HPFS + CBFPS		
	平均 PSNR (dB)	平均比特率	平均 Δ PSNR (dB)	平均比特率改变 (%)	亚像素搜索点数减少 (%)	平均 Δ PSNR (dB)	平均比特率改变 (%)	亚像素搜索点数减少 (%)
tempete_football_704 x 480i	34.9275	5269.26	-0.0118	0.2316	28.6258	0.0329	-0.6399	8.3656
football_704 x 480i	36.2275	3650.77	-0.0195	0.3518	28.0031	0.0058	-0.1029	7.1081
tempete_704 x 480i	34.2900	5322.41	0.0051	-0.1111	29.6299	0.0368	-0.8122	9.3410
mobilecalendar_720 x 576i	33.5200	10498.76	-0.0058	0.1294	29.9674	0.0297	-0.5894	9.9065
bus_interview_704 x 576i	35.7900	3784.47	-0.0200	0.3715	31.8703	0.0631	-1.1742	10.5740
husky_704 x 576i	33.0525	15332.25	-0.0064	0.0875	28.7028	0.0768	-1.0735	11.4390
kayak_1920 x 1080i	36.4225	64117.64	-0.0044	0.0613	30.3265	0.0081	-0.1121	9.9255
Fireworks_1920 x 1080i	33.3675	136674.98	-0.0021	0.0251	30.1560	0.0450	-0.5733	9.2734
flamingo_1920 x 1080i	36.3150	53266.78	-0.0147	0.2189	29.2373	0.0214	-0.3196	9.3618
city_1280 x 720p	36.1950	4458.79	-0.0112	0.3298	30.0031	-0.0007	0.0711	6.6306
crew_1280 x 720p	38.4875	3260.18	-0.0171	0.4990	34.0247	-0.0117	0.2872	8.8566
Harbour_1280 x 720p	35.7800	8376.19	-0.0114	0.2828	29.4734	0.0012	-0.0230	8.0898
night_1280 x 720p	36.8800	6838.39	-0.0023	0.0679	33.7038	0.0080	-0.0230	9.8989
Spincalendar_1280 x 720p	35.5400	5795.24	-0.0089	0.3913	29.7563	-0.0076	0.3414	4.1777
TOTAL AVERAGE	35.4854	23331.87	-0.0093	0.210	30.249	0.0221	-0.339	8.782

从表 1 可以看出,同亚像素全搜索算法比较,本文算法在平均峰值性噪比下降仅为 0.009 3dB 的情况下,平均减少了超过 30% 的亚像素搜索点数,同时比特率仅增加了 0.21%。同 AVS 现有亚像素快速运动估计算法比较,平均峰值性噪比增益为 0.022 1dB,码率降低了 0.34%,搜索点数平均减少了近 9%。在最佳情况下,Husky 序列的性能增加了 0.076 8dB。可见本算法充分利用了亚像素运动分布的规律,在减少搜索点数的同时也达到了相当高的精度,和 AVS 现有亚像素快速算法相比较,不但搜索点数有一定下降,而且性能有所增加,有效地减少了编码器的计算量。

5 结 论

本文通过对亚像素邻域内 SAD 误差特性的研究,得出了 SAD 函数随着与搜索起点的距离增加而增大,并且运动矢量集中在搜索起点附近的结论。在此基础上提出了一种亚像素自适应范围搜索算法 ARFPS。仿真结果表明,与亚像素全搜索算法相比,在性能损失极小的情况下,对标准测试序列利用 AVS 编码器编码时可以降低亚像素搜索过程中 30.25% 的运算量。与 AVS 现有亚像素快速运动估计算法相比,在得到 0.022 1dB 的性能增益的同时,

亦减少了近 9% 的运算量,因此本算法能有效提高 AVS 编码器的运行速度和编码效率。本算法也同样适用于其他采用了亚像素运动估计的视频编码标准。

参考文献 (References)

- 1 Beijing: AVS Audio Video Coding Standard Workgroup of China, AVS Video Group. Information Technology—Advanced Audio Video Coding[R]. AVS-N1063, 2003; 6~8. [北京: AVS 数字音视频编解码技术标准工作组, AVS 视频专题组. 信息技术——先进音视频编码[M]. AVS-N1063, 2003; 6~8.]
- 2 Chen Z, Zhou P, He Y. Fast integer and fractional pel motion estimation for JVT[A]. In: JVT-F017r. doc, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG. 6th Meeting[C], Awaji, Island, Japan, 2002, 12: 1~6.
- 3 Dong Hai-yan, Zhang Qi-ahan. A fast half-pixel motion vector search algorithm[J]. Journal of Image and Graphics, 2005, 7(11): 884~887. [董海燕,张其善. 一种快速半像素运动矢量搜索算法[J]. 中国图象图形学报, 2005, 7(11): 884~887.]
- 4 Yin P, Cheong H Y, Tourapis A M, et al. Fast mode decision and motion estimation for JVT/H.264[J]. Proceedings of the 2003 IEEE International Conference on Image Processing, 2003, 9(3): 853~856.
- 5 Bjontegaard G. Calculation of average PSNR differences between RD-curves[R]. ITU-T Q. 6/SG16 (VCEG) Document VCEG-M33, 2001, 3: 1~4.

更 正

2007 年第 7 期 1301 页作者排版有误,应为林谋广,罗笑南。特此更正。

中国图象图形学报编辑部

2007 年 8 月 23 日