

基于线性预测的半像素运动估计

章伟明 徐元欣 王 匡

(浙江大学信息与通信工程研究所, 杭州 310027)

摘 要 在视频编码系统中,半像素精度的运动估计虽可以明显地改善编码效果,但也因此增加了不少运算量。为了降低运算量 and 提高半像素运动估计的速度,提出了一种全新的半像素搜索算法(half-pixel motion estimation based on linear prediction, 简称 BLPHME),其关键思想是通过分析整像素搜索和半像素搜索结果之间的相关性来建立一个线性模型,通过动态调整判决门限,以便预测并跳过那些不能从半像素块匹配搜索中得到好处的块。实验结果表明,该算法不仅可以明显地降低运动估计的运算量,同时还能得到与传统算法非常接近的图像质量和码率。此外,该算法还可以和基于整像素和半像素的快速运动估计算法一起使用,以进一步降低运算量。

关键词 运动估计 块匹配 线性预测 4-MV 半像素搜索

中图法分类号: TN919.81 TP391.41 **文献标识码:** A **文章编号:** 1006-8961(2007)01-0027-05

Half-Pixel Motion Estimation Based on Linear Prediction

ZHANG Wei-ming, XU Yuan-xin, WANG Kuang

(Institute of Information and Communication Engineering, Zhejiang University, Hangzhou 310027)

Abstract In video coding system, motion estimation at half-pixel accuracy can obviously enhanced coding efficiency compared to the motion estimation at integer-pixel accuracy only. However it requires more computation at the same time. In order to reduce the computation and while enhance the speed of motion estimation at half-pixel, we proposed a novel half-pixel motion estimation algorithm named half-pixel motion estimation based on linear prediction (BLPHME) for video coding in this paper. The key point of the algorithm is setting up a linear model by analyzing the relativity between the results of integer-pixel search and half-pixel search. Then we can modify the threshold dynamically for each frame. By doing so, it can skip over the blocks that can't be benefited from the half-pixel search based on the block size. Experimental results show that significant reduction in computation of motion estimation is achieved together with the increment in bit rate by using the proposed method, but without obvious increment in bit rate and visible loss in video fidelity and increment in bit rate. Moreover, the proposed algorithm can be combined with any of the fast motion estimation algorithm at integer-pixel or half-pixel accuracy to further reduce the complexity.

Keywords motion estimation, block-matching, linear prediction, 4-MV, half-pixel search

1 引言

众所周知,运动估计是视频压缩编码系统的核心部分之一,虽然它可以有效地降低连续视频帧之间的时间冗余度,但同时也占据了整个编码过程的大部分运算量,是影响编码速度的主要因素。

为了提高运动估计的速度,人们相继在整像素

搜索和半像素搜索领域提出了许多快速算法,其中在整像素搜索领域,典型的算法有:三步法(thress-step search, TSS),新三步法(new thress-step search, NTSS)^[1],四步法(four-step search, FSS)^[2],交叉法(cross search, CS)^[3]和钻石算法(diamond search, DS)^[4]等;而在半像素搜索领域虽然还没有一种被广泛认可的快速算法,但也有不少人提出了一些较为实用的快速算法,如 PPHPS^[5] (parabolic

基金项目:国家自然科学基金资助项目(90307002)

收稿日期:2005-04-01; 改回日期:2005-12-08

第一作者简介:章伟明(1980 ~),男。2006年获浙江大学硕士学位。主要研究方向为视频压缩和图像处理。E-mail: glacier1980@hotmail.com

prediction-based fast half-pixel search)算法。

这些算法基本上都遵守了 MPEG4 校验模型 (verification model, VM)^[6] 中关于运动估计的算法流程:

(1) 整像素运动估计 首先计算当前宏块 (macro block, 大小为 16×16) 和参考宏块的绝对误差和 (sum of absolute difference, SAD), 并将搜索范围之内具有最小 SAD_{16} 的 (x, y) 定义为该宏块的整像素运动矢量; 然后分别计算宏块内的 4 个块 (block, 大小为 8×8) 的 SAD_8 , 并定义 $SAD_{4 \times 8} = \sum_1^4 SAD_8(x, y)$ 。

(2) Intra/Inter 编码方式的选择

(3) 半像素运动估计 首先对参考帧进行双线性插值, 然后在整像素搜索的基础上通过进行以半像素为单位的搜索来得到半像素的 SAD_{16} 和 $SAD_{4 \times 8}$ 。如果 $SAD_{4 \times 8} < SAD_{16} - (N_B/2 + 1)$ (式中 $N_B = N \times 2^{(b-8)}$, N 为 VOP 内部所有的点数, 在 VM 中, 由于每像素比特数 $b=8$, 因此, $N_B=256$, 也就是说, 判断条件是差值为 129), 则采用四矢量模式, 否则就用单矢量来描述整个宏块的运动情况。

本文提出的基于线性预测的半像素运动估计算法 (half-pixel motion estimation based on linear prediction, BLPHME) 将开辟一个有关半像素搜索

的全新研究领域, 该算法的关键思想是跳过那些不能从半像素块匹配搜索得到好处的块, 并且该算法与整像素和半像素搜索的具体算法无关, 也就是说, 其能够与其他任何有关整像素及半像素的快速算法融合在一起使用, 以快速提高编码速率。

本文接下来将首先介绍 BLPHME 算法, 然后给出 BLPHME 算法的测试及分析结果, 最后给出结论。

2 基于线性预测的半像素搜索算法

BLPHME 算法的验证过程是在 MPEG4-SP (simple profile) 下完成的。运动估计采用的是块匹配算法, 即首先将图像分割成一个个大小为 16×16 的宏块, 然后再将每个宏块分割成 4 个大小为 8×8 的块, 并以 SAD 最小作为选择最匹配块的依据。

2.1 观察和动机

BLPHME 算法的提出最初是基于这样一个事实: 通过对一些经典序列的测试验证 (测试环境为: MPEG4 校验模型 SP 框架, 整像素运动估计采用全搜索算法 (full search, FS), 采用固定量化系数, $QP=10$, CIF 格式) 发现, 最终采用 4-MV (motion vector) 模式的宏块比率其实并不高。表 1 列出了经典视频序列中实际采用 4-MV 模式的宏块占全部宏块数 (396) 的比率, 记作 ESR (effective half-pixel search ratio)。

表 1 经典视频序列中 4-MV 模式的宏块占全部宏块数的比率

Tab. 1 ESR of the classic video sequence								
序列名	News	Silent	Paris	Tempete	Foreman	Stefan	Mobile	Football
4-MV 概率 (%)	5.06	7.19	11.61	14.24	16.35	19.74	26.26	29.78

在传统的算法流程中为了判断一个宏块是否有必要使用 4-MV 模式, 首先需要对每个宏块进行半像素的块匹配搜索后才能决定。然而从表 1 中可以看出, 由于不管测试序列的运动是缓慢的还是剧烈的, 最后采用 4-MV 模式的宏块比率都不超过 30%, 因此, 针对半像素的块匹配搜索有绝大部分都是无效搜索。如果能事先知道哪些宏块最后不会采用 4-MV 模式, 那么就不必对这些宏块进行半像素的块匹配搜索。基于判断一个宏块是否会采用 4-MV 模式的条件是不等式 $SAD_{4 \times 8} < SAD_{16} - (N_B/2 + 1)$ (这里的值为半像素搜索结果) 是否成立, 考虑是否可以利用整像素搜索的结果呢? 为此, 采用以整像素 $SAD_{16} - SAD_{4 \times 8}$ 的值作为横坐标, 以半像素 $SAD_{16} - SAD_{4 \times 8}$ 作为纵坐标的坐标系, 并以直线来拟合这些

点, 然后计算当 $Y=129$ 时的 X 值 (取整), 并将其作为判断是否需要进行搜索的半像素块搜索的门限, 即可得到如表 2 所示结果。

表 2 整像素 ($SAD_{16} - SAD_{4 \times 8}$) 与半像素 ($SAD_{16} - SAD_{4 \times 8}$) 的关系

Tab. 2 The relationship between integer-pixel and half-pixel about $SAD_{16} - SAD_{4 \times 8}$		
序列名	拟合直线	判决门限
News	$Y=0.93232X+19.016$	118
Silent	$Y=0.82965X+21.193$	130
Paris	$Y=0.98537X+36.778$	94
Tempete	$Y=0.67431X+50.864$	116
Foreman	$Y=0.9803X+28.788$	102
Stefan	$Y=1.0684X+41.531$	82
Mobile	$Y=0.54189X+58.554$	130
Football	$Y=1.0949X+231.619$	98

由于各个序列的运动特征以及运动的剧烈程度不同,也导致门限各不相同,因此,在面对输入序列未知的情况下,需要寻找一种能自动调节门限的自适应算法。在考虑了各种变量以后,发现没有一种方法可以直接决定某个序列的门限。尽管如此,还是可以通过检验实时的压缩结果来判断当前的门限

是否适合这个序列,并可作相应的调整。

2.2 一个线性模型

本文将表 2 中设定了判决门限以后需要进行半像素块搜索的比率定义为最佳半像素搜索率(optimal half-pixel search ratio, OSR)。表 3 分别列出了 8 个测试序列的 ESR 与 OSR。

表 3 ESR 和 OSR
Tab.3 ESR and OSR

指标	序列名							
	News	Silent	Paris	Tempete	Foreman	Stefan	Mobile	Football
ESR(%)	3.571	5.220	8.629	7.025	13.434	12.057	11.177	24.582
OSR(%)	4.518	6.003	11.75	12.631	17.518	19.58	15.04	26.947

以 ESR 为横坐标,以 OSR 为纵坐标将这些点画到直角坐标系上(如图 1 所示),从图 1 可见,其具有比较明显的线性特征。通过拟合这些点,即可得到如下的线性关系式:

$$OSR = 1.1 \times ESR + 0.026 \quad (1)$$

这为实时的信号压缩提供了一种逼近 OSR 的方法:即先预设一个门限,同时计算当前帧的实际半像素块搜索率(actual half-pixel search ratio, ASR)和有效半像素块搜索率;然后利用等式(1)求出当前情况的最佳半像素块搜索率;最后通过比较当前 ASR 与 OSR 之间的关系,就可以判断当前门限是否适合这个序列,以便调整这个值作为下一帧的门限。

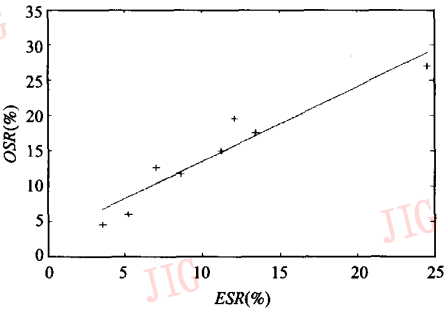


图 1 ESR 与 OSR 之间的线性关系

Fig.1 The linear relationship between ESR and OSR

2.3 动态调整门限

BLPHME 算法的具体实现步骤如下:

- (1) 设 $m4v_searchMB = 0, m4v_usefulMB = 0, del_threshold = 111$ (该初始值是上述 8 个序列的平均值);
- (2) 对于每个宏块,在整像素搜索结束以后就

可以得到 SAD_{16} 和 $SAD_{4 \times 8}$, 当 $SAD_{16} - SAD_{4 \times 8} > del_threshold$ 时,则令 $flag = 1$, 否则 $flag = 0$;

(3) 若 $flag = 1$,则该宏块进行半像素的块匹配搜索, $m4v_searchMB++$, 如果该宏块最终采用了 4-MV 模式,则 $m4v_usefulMB++$, 若 $flag = 0$,则跳过该宏块的半像素块匹配搜索;

(4) 令 $ASR = m4v_searchMB/396, ESR = m4v_usefulMB/396$, 则利用等式(1)即可求出相应的 OSR;

(5) 调整 $del_threshold$ 作为下一帧的门限。门限的调整公式定义如下:

$$del_threshold = del_threshold \times (1 + ASR/OSR)/2 \quad (2)$$

当 $ASR > OSR$ 时,由等式(2)可知,由于门限将被提高,从而可降低下一帧的 ASR,这就起到一个负反馈的作用,反之亦然。

3 实验结果及分析

为了测试 BLPHME 算法的性能,在 PC 机(Pentium IV 1.8G,512M 内存)上,对 BLPHME 算法和传统算法进行比较测试。测试环境是 MPEG4 校验模型 SP 框架,测试样本序列为本文前面已提到的 8 个经典的 CIF 格式的视频序列(这些序列基本包括了各种类型的运动变化),帧频为 25fps,量化系数 $QP = 10$,测试长度除了 Football 和 Stefan 是 90 帧外,其余的都为 100 帧,搜索窗口大小为 $(-16, 15.5)$ 。为了更好地测试本文算法的性能,整像素和半像素运动估计都采用了全搜索算法,其中第 0 帧是 I 帧,其后各帧都是 P 帧。表 4 给出了传统算

表 4 两种算法流程的性能比较

Tab.4 Compare performances between the two algorithms

指标	算法	序列名							
		News	Silent	Paris	Tempete	Foreman	Stefan	Mobile	Football
Y 分量峰值 信噪比 (dB)	传统算法	34.80	32.91	31.55	30.48	33.87	31.56	29.49	34.35
	BLPHME 算法	34.78	32.90	31.53	30.45	33.83	31.54	29.44	34.34
	PSNR 增加值	-0.02	-0.01	-0.02	-0.03	-0.04	-0.02	-0.05	-0.01
平均码率 (kbps)	传统算法	152.2	152.2	366.2	754.4	293.2	806.6	1429.6	698
	BLPHME 算法	153.0	151.4	370.0	755.8	291.8	808.2	1445.6	694.4
	平均码率增加值	+0.8	-0.8	+3.8	+1.4	-1.4	+1.6	+16	-3.6
	ASR (%)	8.525	10.132	12.457	9.035	17.439	13.468	8.982	32.334

法和 BLPHME 算法在信噪比、比特率方面的性能比较以及所有帧的实际半像素块搜索率的平均值。

仿真结果表明, BLPHME 算法可以极大地降低半像素块匹配搜索的数目, 该数目平均只占传统算法的 14%, 即平均每个宏块只需要搜索 4.56 个半像素点(由于半像素搜索包括以宏块为单位和以块为单位的二次搜索, 因而传统的算法共需搜索 $8+8=16$ 个宏块, 但一般说成 8 个, 即整像素运动矢量周围的 8 个点。因此可将 BLPHME 算法的结果也统一折算成以宏块为计算单位, 可得如下结果 $8\times 50\%+8\times 14\%\times 50\%=4.56$), 而图像峰值信噪比平均只下降 0.025dB, 码率平均提高 1.825kbps, 可见本文算法编码获得的图像质量和码率与传统算法相当。同时所选择的测试序列的运动特征和激烈程度虽各不相同, 但都获得了很好的结果, 这说明 BLPHME 算法具有相当广泛的适用性。

由于 BLPHME 算法在跳过大多数无效的块匹配搜索的同时, 也跳过了某些事实上需要采用 4-MV

模式宏块的块匹配搜索, 从而导致匹配效果不如传统算法的好, 即使得图像的峰值信噪比略有下降, 并使残差增大、码率上升, 但一般认为, 峰值信噪比下降在 0.1dB 以内, 其对图像质量影响并不明显。进一步的分析发现, 对于那些运动比较平缓的序列而言, 用新算法编码后的图像, 无论在峰值信噪比, 还是码率上, 和传统算法的差别都不大, 而对运动剧烈、细节较多的序列(如 Mobile), 图像峰值信噪比的下降就相对较大, 这是由于这类序列整像素和半像素的相关性较弱所引起的。

如本文第 1 部分所述, 目前在整像素搜索领域有许多快速算法, 它们大多被一些实时的视频编码系统所采用。在接下来的测试中, 将采用文献[4]中提出的钻石算法作为整像素的搜索算法, 并结合 BLPHME 算法进行性能测试, 测试环境及各参数和前一个实验相同。表 5 给出了当整像素搜索采用钻石算法时, 两种算法流程在编码图像的信噪比、码率和平均编码时间 3 个方面的整体性能比较结果。

表 5 两种算法流程的性能比较

Tab.5 Compare performances between the two algorithms

指标	算法	序列名							
		News	Silent	Paris	Tempete	Foreman	Stefan	Mobile	Football
Y 分量峰值 信噪比 (dB)	传统算法	34.68	32.88	31.52	30.46	33.65	31.47	29.49	34.21
	BLPHME 算法	34.66	32.87	31.50	30.43	33.60	31.44	29.44	34.19
	PSNR 增加值	-0.02	-0.01	-0.02	-0.03	-0.05	-0.03	-0.05	-0.02
平均码率 (kbps)	传统算法	147.2	147.8	365.0	748.2	282.0	842.2	1417.8	799.0
	BLPHME 算法	147.0	146.8	368.0	748.8	280.0	847.0	1435.4	790.6
	平均码率增加值	-0.2	-1.0	+3.0	+0.6	-2.0	+0.8	+17.6	-8.4
平均每帧编码 时间 (ms)	传统算法	52.1	52.5	52.5	54.3	53.8	48.3	54.7	49.6
	BLPHME 算法	45.3	46.4	45.4	47.1	48.4	43.4	48.1	46.5
	编码时间 变化率 (%)	13.052	11.619	13.524	13.260	10.037	10.145	12.066	6.250

仿真结果显示,当整像素采用钻石搜索时,BLPHE 算法相比传统算法,在图像信噪比和码率上依然保持了很好的性能,而平均每帧的编码时间却下降了 11% 以上。这充分说明,在整像素搜索越来越快的情况下,降低半像素搜索的计算量对于提高编码器效率的重要性。当整像素搜索采用比钻石算法更有效的算法时,BLPHE 算法的作用就会显得更加重要。

4 结 论

本文针对整像素与半像素搜索结果之间的相关性,提出了一种全新的半像素搜索算法。该算法通过研究整像素宏块与块匹配搜索结果之间的大小关系,提出了一个线性模型,用来动态地调整判决门限,以便预测并跳过大部分宏块内的半像素块匹配搜索。实验结果表明,BLPHE 算法可以有效地降低半像素搜索的运算量,同时又可保持相当好的图像信噪比和码率。

BLPHE 算法还可以和其他任何整像素及半

像素的快速算法结合起来使用,其不仅可极大地降低运动估计部分的运算量和提高视频压缩的效率,而且算法本身的复杂度很低,易于实现。

参考文献 (References)

- 1 Li Ren-xiang, Zeng Bing, Liou Ming L. A new three-step search algorithm for block motion estimation [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1994, 4(4): 438 ~ 442.
- 2 Po Lai-man, Ma Wing-chung. A novel four-step search algorithm for fast block motion estimation [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1996, 6(3): 313 ~ 317.
- 3 Ghanbari M. The cross-search algorithm for motion estimation [J]. IEEE Transactions on Communications, 1990, 38(7): 950 ~ 953.
- 4 Zhu Shan, Ma Kai-kuang. A new diamond search algorithm for fast block-matching motion estimation [J]. IEEE Transactions on Image Processing, 2000, 9(2): 287 ~ 290.
- 5 Du Cheng, He Yun, Zheng Jun-li. PHPS: 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.
- 6 ISO/IEC. Information Technology-generic Coding of Audio-visual Objects, Part 2: Visual [S]. ISO/IEC 14496-2. Nov. 1998