

H. 264 JM 模型中运动估计算法及改进方案

郑振东 王沛 应骏

(上海师范大学, 上海 200234)

摘要 JM 模型是 JVT(joint video team)发布的 H. 264 标准测试模型,对算法学习和研究有着重要的意义。根据 JM 测试模型的参数设定,其中的运动估计算法有 3 种可选模式。本文结合 JM10.2 的源代码对 UMHExagonS 算法进行了分析,并对该算法进行改进,能够在保证视频序列各分量信噪比的情况下缩短运动估计的耗时。本文利用 UMHExagonS 算法的准确预测以及运动估计代价的相关性来设置阈值达到提前结束搜索的目的。在 JM10.2 的测试模型上进行了算法验证。实验结果表明,利用块与块之间运动估计代价的相关性,在保证编码性能的同时,可以减少运动估计所需时间的 10% 以上。

关键词 H. 264 运动估计 UMHExagonS 算法 JM 测试模型

中图分类号: TP37 **文献标识码**: A **文章编号**: 1006-8961(2007)10-1798-04

H. 264 Motion Estimation Algorithm in JM Model and Its Improvement Program

ZHENG Zhen-dong, WANG Pei, YING Jun

(Shanghai Normal University, Shanghai 200234)

Abstract JM is the standard test model published by joint video team(JVT), which has great significance in algorithm learning and research. In JM, there are three optional modes by setting parameters differently. Based on the study of the principle of the UMHExagonS algorithm, an improved method is proposed to reduce the time of motion estimation with minimal loss in bitrate and reconstructed quality compared with those of the original UMHExagonS algorithm. Correlation between adjacent blocks is used to end the search. In JM10.2, results show that the improved UMHExagonS algorithm can reduce the time of motion estimation with the original performance of the original UMHExagonS algorithm.

Keywords H. 264, motion estimation, UMHExagonS algorithm, JM test mode

1 引言

在 H. 264 的三大开源编码器中, JM (joint model)是 JVT(joint video team)的官方测试源代码^[1], JM 实现了 H. 264 所有的特性,由于是官方的测试源码,所以学术研究的算法都是在 JM 基础上实现并和 JM 进行了比较^[2]。

在 H. 264 的编码过程中,运动估计无疑是复杂度和运算两最高的部分^[3,4]。目前 H. 264 已正式采纳了“非对称十字形多层次六边形格点搜索”

(UMHExagonS 算法),它的运算量相对于快速全搜索算法可节约 90% 以上,同时能保持较好的率失真性能。在 JM 的测试模型中采用了该算法作为整像素运动估计算法,其性能非常接近全搜索算法。

本文对 UMHExagonS 算法的两处进行了改进。一是在预测部分加入一个搜索阈值判断,在运动估计性能达到满意的情况下提前结束搜索;二是根据场景中的运动一般都是某一对象的运动,则运动对象内部的各个块的运动估计的代价具有一定的相关性,则在算法中考虑这一点,在满足条件的情况下提前结束当前的搜索步骤。

收稿日期:2007-07-05;改回日期:2007-07-25

第一作者简介:郑振东(1984 ~),男。上海师范大学通信与信息系统专业硕士研究生。主要研究方向为 H. 264 编解码实现和运动估计。已获专利一项。E-mail: dongdongtoyoyou@126.com

2 UMHexagonS 算法

JM 测试模型中的运动估计主要由两种可选的匹配算法,一种是全搜索法;另一种是 UMHexagonS 算法。理论上,全搜索算法具有最高的精度,但存在着搜索时间长和运算量大的缺点。UMHexagonS 算法是一种已商用的运动估计算法,该算法相对于快速全搜索算法可节约 90% 以上的运算量,同时能保持较好的率失真性能。

UMHexagonS 算法大致分为以下几个步骤^[5,6]:

- (1) 起始搜索点的预测;
- (2) 非对称十字搜索,如图 1 所示;
- (3) 螺旋搜索 5×5 区域,类似于全搜索,只搜索前 25 个点,如图 2 所示;

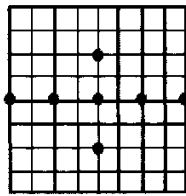


图 1 非对称十字形搜索

Fig. 1 Unsymmetrical cruciform search

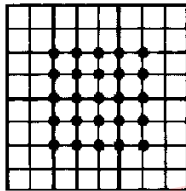
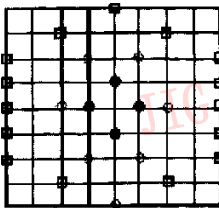


图 2 螺旋搜索 5×5 区域

Fig. 2 Spiral search

- (4) 超六边形模板搜索,以搜索范围的 $1/4$ 为间隔的大六边形模板搜索,如图 3 所示;



□表示超六边形模板 ○表示中六边形模板 ●表示小菱形模板

图 3 超六边形模板、中六边形模板以及小菱形模板

Fig. 3 Hexagon template and rhombus template

- (5) 多圈的中六边形模板搜索,在搜索范围之内进行多圈的中六边形模板的搜索,如图 3 所示;

- (6) 小菱形模板反复搜索,得到最终的运动矢量,如图 3 所示。

UMHexagonS 算法流程图如图 4 所示。

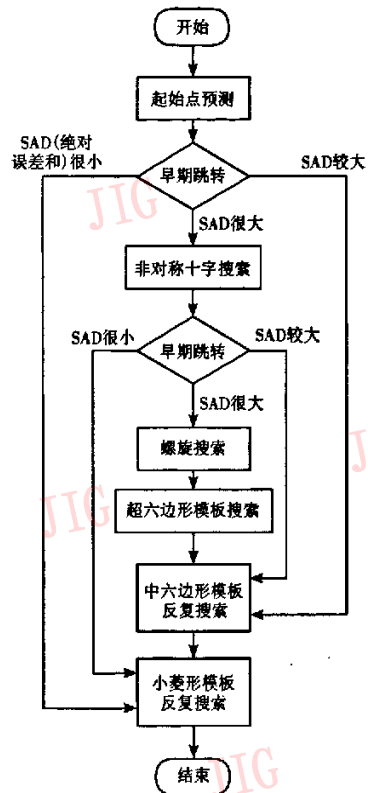


图 4 UMHExagonS 算法流程图

Fig. 4 UMHExagonS algorithm search flowchart

3 基于 UMHExagonS 的改进方法

在 UMHExagonS 算法的步骤 4、5 的搜索中使用的都是多圈的六边形搜索,而步骤 6 采用的小菱形模板反复搜索也是以“最佳点在模板中心”为准则来判断是否最优、是否可结束搜索^[7],这 3 处都是非常消耗搜索点数的步骤。也就是说即使已经搜索到最佳点,但是仍然需要把剩余的点搜索完毕才可以继续下一个搜索点,而这些剩余点数的计算量不可小视。以下改进都是基于提前中止搜索的思想,希望能够在保证 Y 、 U 、 V 各分量信噪比的前提下缩短搜索时间。

场景中一般都是以某一对象为单位的运动,那

么在该对象内部的代价应该是有很高的相关性的,如图 5 所示为一残差图像,可以从图中看到,除了运动对象边缘比较突出外,对象内部的残差值是同一灰度,即非常相似的残差值,因此可以断定,在对象内部各块的代价也具有相似性。

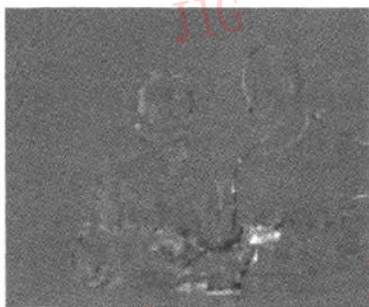


图 5 第 1 帧与第 2 帧经过块预测后的残差图像

Fig. 5 Residual image between two frames

另设一变量 $pred_cost$ 来存储上一个点运动估计后的代价 min_mcost , 在下一个点运动估计的步骤 4、5 中, 如果满足

$$min_cost < pred_cost \times percent \quad (1)$$

则直接结束搜索, 其中 $percent$ 是一个小于 1 的数。具体的改进步骤如下所述。

在步骤 4 超六边形搜索中, 首先以 $search_range/4$ 为半径的六边形进行预测点的搜索, 然后增加半径为 $search_range/2$ 进行预测点的搜索, 以此类推直到搜索半径为 $search_range$, 则搜索结束。

改进方法为在每一个六边形搜索结束后加入如式(1)的判断, 如果满足式(1)的条件则结束搜索, 改进后超六边形搜索的流程图如图 6 所示。

然后在每次整个超六边形搜索算法接受前将当前块的最小代价记录下来

$$pred_cost = min_mcost \quad (2)$$

对于步骤 5 的中六边形搜索也做类似改进, 经过测试也可以取得一定的效果。

4 实验结果与分析

4.1 测试环境

使用 JM10.2 作为测试模型, $QP = 28$, 分别选取了以下 3 个 QCIF 测试序列:

- (1) akiyo.yuv 整个场景运动比较低的类型代表。
- (2) foreman.yuv 整个场景运动一般的类型代表。

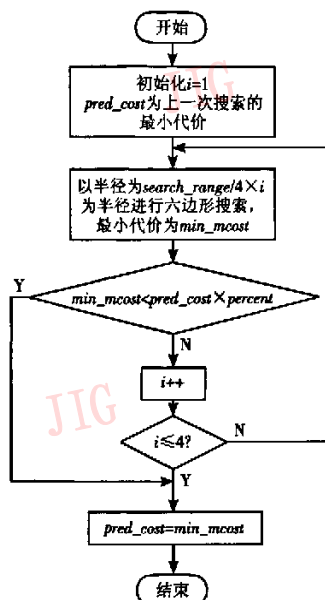


图 6 改进后的超六边形搜索流程图

Fig. 6 Flowchart of improved hexagon search

(3) mobile.yuv 场景变换比较剧烈的类型代表。

4.2 实验数据

性能比较主要取了 Y 、 U 、 V 各分量的平均信噪比, MET 为运动估计的时间。

从表 1 可以看出, 对于场景运动较低的 akiyo.yuv, 改进后的算法性能与原算法保持一致, 原因是场景

表 1 测试序列的实验数据

Tab. 1 Experimental data of test sequence

序列	percent 取值	平均信噪比 (dB)			MET (s)
		Y	U	V	
foreman	原算法	36.83	40.84	42.37	12.981
	0.9	36.84	40.88	42.34	12.981
	0.8	36.84	40.85	42.35	12.367
	0.7	36.83	40.86	42.43	10.987
	0.6	36.84	40.87	42.41	11.150
mobile	原算法	33.63	36.18	35.71	12.697
	0.9	33.63	36.17	35.72	11.017
	0.8	33.62	36.16	35.73	11.039
	0.7	33.63	36.17	35.72	11.237
	0.6	33.63	36.18	35.71	11.224
Akiyo	原算法	38.82	41.14	42.09	6.372
	0.9	38.82	41.14	42.09	6.499
	0.8	38.82	41.14	42.09	6.210
	0.7	38.82	41.14	42.09	6.350
	0.6	38.82	41.14	42.09	6.246

运动较低的视频序列进行起始点预测的精度较高,在早期判断时,因为绝对误差和(SAD)很小,所以直接跳过了超六边形于中六边形搜索,直接进行小菱形的搜索,因此没有用到提前判断退出。而对于场景运动一般以及场景运动较激烈的 foreman 和 mobile,改进后的算法无论是在平均信噪比和 ME 所花的时间上都有一定的下降。另外关于利用代价的相关性判断的位置有待于进一步的实验测试,表 1 的实验数据是基于把该判断放置于步骤 4 与步骤 5 中产生的效果。

5 结 论

本文在结合 JM 模型的基础上,首先介绍了 H. 264 中运动估计算法 UMHExagonS 的算法,然后针对该算法的特点及不足之处提出了两处改进,一是利用 UMHExagonS 的准确的起始点预测,二是利用了相邻块代价的相关性。在保证视频序列各分量的信噪比,甚至有所提高的情况下,使搜索时间平均降

低了 10% 以上,降低了编码器端运动估计的耗时。

参考文献 (References)

- 1 JVT-Q042, Revised H. 264/MPEG-4 AVC Reference Software Manual[S].
- 2 JVT-0079, Text Description of Joint Model Reference Encoding Methods and Decoding Concealment Methods JVT[S].
- 3 ISO/IEC JTC1/SC29/WG11 and ITU-T SG16/Q. 6 Draft ITU-T, Recommendation H. 264 and Final Draft International Standard[S].
- 4 Thomas Wiegand, Overview of the H. 264 _ AVC video coding standard[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 560 ~ 576.
- 5 Choudhury A Rahman. UMHExagonS algorithm based motion estimation architecture for H. 264/AVC[A]. In: Proceedings of the Fifth International Workshop on System-on-Chip for Real-Time Applications[C], Washington, DC, USA, 2005: 207 ~ 210.
- 6 Yang Yu-hong. Explore and improvement based on UMHExagonS algorithm [J]. Computer Engineering and Application, 2006, 11(3): 52 ~ 54. [杨育红. 快速运动估计 UMHExagonS 算法的探讨与改进[J]. 计算机工程与应用, 2006, 11(3): 52 ~ 54.]
- 7 JVT-F017, Fast Integer Pel and Fractional Pel Motion Estimation for JVT[S].