

[文章编号] 1003-4684(2019)02-0014-04

采用运动矢量的视频快速运动估计算法

涂玲英, 王塬鹏, 秦杰杰, 李儒彬, 王 锐

(湖北工业大学电气与电子工程学院, 湖北 武汉 430068)

[摘 要] 提出了利用运动矢量来解决视频编码时间过长的观点。结合 H.264 标准的 UMHexagons 算法,在起始点预测和搜索模板部分做出了优化,能够有效地降低搜索过程中的搜索点,提高视屏编码效率。通过对 5 种运动程度不同的视频序列进行测试,实验证明,对比原算法,优化后的算法视频的信噪比只下降了 0.01~0.02 dB,运动估计时间却减少了 10%~30%。

[关键词] 运动估计算法; UMHexagons 算法; 起始点预测; 信噪比

[中图分类号] TN919.8

[文献标识码] A

运动估计算法是视频编码的核心部分,占整个编码时间 70% 以上,因此运动估计算法的好坏直接影响视频编码的性能。基于块匹配的运动估计算法因其简单且易在硬件中实现,成为了当前视频压缩中的主流算法,被广泛应用于各种视频压缩标准中,其核心技术是找到相对应的最佳匹配块。经过国内外学者的不懈努力,各种优秀的算法被不断提出和完善。全搜索法是通过搜索全局搜索点,得出最佳匹配块,但是由于搜索时间长、运算复杂度高,不适合应用于实时视频中,只能作为其他算法的对照标准。为了减少运动搜索的时间而提出的三步法^[1-4]、新三步法^[1-4]、六边形搜索法^[1-4]、钻石搜索法^[3],虽然通过减少搜索点,降低了运动搜索时间,但是在适用范围内都具有一定的局限性,容易陷入局部最优化的误区。在此基础上,人们又提出了一种新型的混合搜索算法——UMHexagons 算法,该算法在保证视频质量前提下,大幅度减少了视频运动估计的时间,其运动估计时间最低只为全局搜索法的 10% 左右,但是 UMHexagons 算法在实际应用中还存在不足和待改进部分。基于 UMHexagons 算法,本文在起始点预测部分只采用空间域预测方法并对中值预测结果单独增加了提前终止预测;十字形模板部分根据预测运动矢量值进行了运动类型判定并将搜索区域进行四部分划分;5×5 矩形模板部分分析了运动最佳匹配点的出现概率,采用 3×3 的搜索模

板,为了避免陷入局部最优,对当前最佳匹配点的运动矢量值进行判断来选择搜索不同的边界点;多层次六边形模板部分利用运动预测矢量值来对搜索模板行划分四个区域,同时根据运动对象的中心偏置特性减少搜索单层模板的搜索点。

1 UMHexagons 算法介绍

UMHexagons 算法作为当前 H.264 标准下的主流算法(图 1),其步骤下:

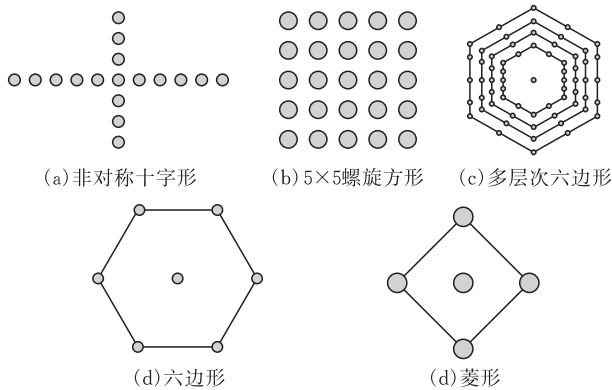


图 1 UMHexagons 算法搜索模板

1)起始点预测:根据视频在空间域和时间域上的相关性,依次采用中值预测、上层预测、起始点预测、前帧预测和前参考帧预测对当前运动矢量 MV 进行预测,确定起始搜索点的位置。

2)非对称十字搜索模板:如图 1a,以起始预测

[收稿日期] 2018-12-20

[基金项目] 国家自然科学基金(41601399)

[第一作者] 涂玲英(1963—),女,湖北武汉人,湖北工业大学教授,研究方向为嵌入式系统,视频压缩技术

[通信作者] 王塬鹏(1993—),男,湖北黄冈人,湖北工业大学硕士研究生,研究方向嵌入式系统,视频压缩技术

点为中心,根据现实世界大多数运动方式是水平运动,采用 Y 轴搜索点为 X 轴搜索点两倍的非对称十字搜索模板进行搜索。

3) 5×5 矩形搜索模板:如图 1b,以非对称十字搜索模板得到的最优点为中心进行 5×5 的矩形范围内最优点求解。

4) 多层次六边形搜索模板:如图 1c,以上一步得到的最优点为中心,采用搜索半径等间距扩大的六边形进行最优点求解,通常采用 4 层非均匀六边形。

5) 六边形搜索模板:如图 1d,以上一步得到最优点为中心进行六边形模板搜索,直到最优点在六边形模板的中心,停止搜索。

6) 小钻石搜索模板:如图 1e,以六边形最优点为中心进行菱形模板搜索,求出最优点结束。

UMHexagons 算法的提前终止策略为:取当前最佳匹配点的 SAD 值分别与已设置不同阈值 T_{TH1} 和 T_{TH2} ($T_{TH1} < T_{TH2}$) 进行比较,当 $SAD < T_{TH1}$ 时,认定为非常满意区域则直接进入小钻石搜索模板;当 $T_{TH1} < SAD < T_{TH2}$ 时,认定为满意区域则直接进入六边形模板;当 $SAD > T_{TH2}$ 时,认定为不满意区域则进入非对称十字搜索模板。 T_{TH1} 和 T_{TH2} 的取值如下:

$$T_{TH1} = SAD_{pred} (1 + \beta_{Third})$$

$$T_{TH2} = SAD_{pred} (1 + \beta_{Sec})$$

自适应因数 β_{Third} 和 β_{Sec} 为:

$$\beta_{Sec} = \frac{Bsize[blocktype]}{pred_min \ cost^2} - \alpha_{Sec}[blocktype]$$

$$\beta_{Third} = \frac{Bsize[blocktype]}{pred_min \ cost^2} - \alpha_{Third}[blocktype]$$

其中, Sec 和 Third 是常数组分别为 UMH 算法程序中的 Alpha Sec 和 Alpha Third, Bsize 是编码块尺寸大小, Sec 和 Third 分别为 JM 模型算法程序中的 beta Sec 和 beta Third。程序代码中 EARLY_TEREMINATION 宏对应判定的提前终止策略。

2 UMHexagons 算法优化

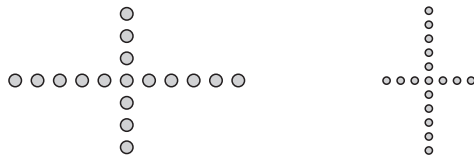
2.1 起始搜索点预测优化

UMHexagons 算法采用五种不同的预测方式进行预测,预测顺序采用中值预测、上层预测、原点预测、前帧预测和前参考帧预测依次执行。根据文献[2]所述,通常情况下,在采用空间域的中值预测、上层预测和原点预测就可以得到最佳起始点。虽然采用时间域的预测方式可以使预测更加精确,但是却明显提高了预测的计算复杂度,故本文只采用空间域的预测方式,不采用时间域的预测方式。根据文献[8]所述,最佳匹配点出现在中值预测的概率比

较大,所以本文在采用空间域的预测方式下对中值预测增加一步阈值判定。

2.2 非对称十字搜索模板的优化

UMHexagons 算法采用非对称十字搜索模板的搜索方式,根据现实世界大多数运动方式是水平运动,所以采用 X 轴搜索点为 Y 轴搜索点两倍的非对称十字搜索模板。在一些特殊的领域,有时候垂直方向上的运动多于水平方向上的,所以该算法存在欠缺。本文对非对称十字搜索模板改进有两点:其一根据上一步的预测矢量在 X 轴和 Y 轴上的分量大小,确定该运动是水平运动还是垂直运动。若判断为水平运动,则采用原模板(图 2a)不变,若判定为垂直运动,则采用 Y 轴是 X 轴搜索点两倍的模板(图 2b);其二根据预测运动矢量落入的区域,采用不同的搜索区域,其模板如图 3 所示。



(a) 水平方向的非对称十字形 (b) 垂直方向的非对称十字形

图 2 不同方向的非对称十字形模板

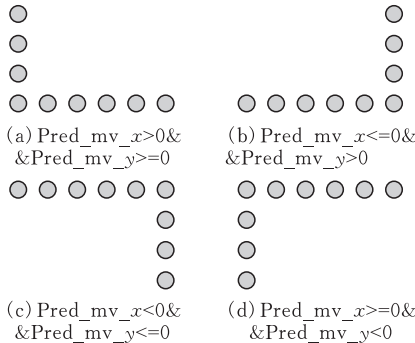


图 3 不同区域的非对称十字形模板

1) 若 $Pred_mv_x > 0 \&\& Pred_mv_y \geq 0$, 判定最佳运动矢量在第一象限, 搜索模板如图 3a;

2) 若 $Pred_mv_x \leq 0 \&\& Pred_mv_y > 0$, 判定最佳运动矢量在第二象限, 搜索模板如图 3b;

3) 若 $Pred_mv_x < 0 \&\& Pred_mv_y \leq 0$, 判定最佳运动矢量在第三象限, 搜索模板如图 3c;

4) 若 $Pred_mv_x \geq 0 \&\& Pred_mv_y < 0$, 判定最佳运动矢量在第四象限, 搜索模板如图 3d。

2.3 5×5 矩形搜索模板的优化

5×5 矩形搜索模板类似于小型全搜索法, 在 (5,5) 的范围内对 25 个搜索点进行螺旋搜索, 因此计算量大, 花费时间长。根据文献[9-10], 通过对常见的视频序列测试得出最佳匹配点, 最佳匹配点出现在 (5,5) 范围内并不是等概率的, 其统计概率如图 3c 所示。根据统计图可知: 离中心越近出现的概率

越高,称之为中心偏置法则。分析最佳匹配点出现的概率情况,可以发现最佳匹配点出现在 3×3 的概率高达 71.792%,相对而言,最佳匹配点出现在 5×5 的概率 81.791%,只少 10%。在此提出了两点改进:1)采用 3×3 矩形模板进行最佳点搜索(图 4)。2)为了避免陷入局部最优化,对 3×3 中得到的最佳匹配运动矢量落入的区域进行分析,根据运动方向的相关性来采用不同的模板,其模板类型见图 5。

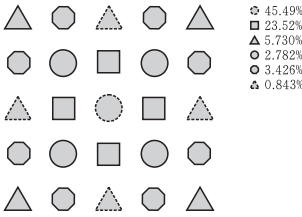


图 4 最佳匹配点在 5×5 模板中的概率分布

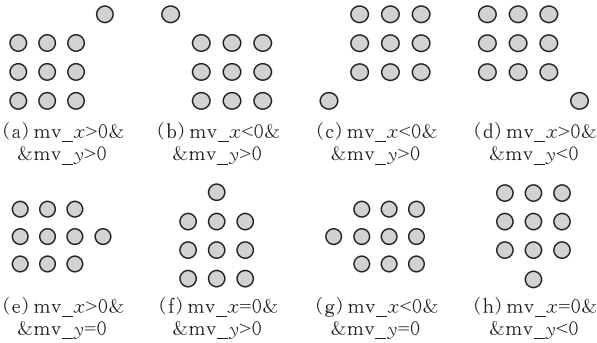


图 5 不同区域搜索模板示意图

- 1)若 $mv_x>0 \& \& mv_y>0$,判定最佳运动矢量在第二象限,搜索模板如图 5a;
- 2)若 $mv_x<0 \& \& mv_y>0$,判定最佳运动矢量在第二限,搜索模板如图 5b;
- 3)若 $mv_x<0 \& \& mv_y<0$,判定最佳运动矢量在第三象限,搜索模板如图 5c;
- 4)若 $mv_x>0 \& \& mv_y<0$,判定最佳运动矢量在第四象限,搜索模板如图 5d;
- 5)若 $mv_x>0 \& \& Pred_mv_y=0$,判定最佳运动矢量在 X 轴正半轴上,搜索模板如图 5e;
- 6)若 $mv_x=0 \& \& mv_y>0$,判定最佳运动矢量在 Y 轴正半轴上,搜索模板如图 5f;
- 7)若 $mv_x<0 \& \& mv_y=0$,判定最佳运动矢量在 X 轴负半轴上,搜索模板如图 5g;
- 8)若 $mv_x=0 \& \& Pred_mv_y<0$,判定最佳运动矢量在 Y 轴负半轴上,搜索模板见图 5h。

2.4 多层次大六边形搜索模板的优化

多层次大六边形搜索模板作为 UMHexagons 算法中粗略搜索的最后一个搜索模板,采用了由内到外多个点数相同大小不等的六边形搜索模板。通常情况下搜索模板为四层,共计搜索 64 个点,虽然

采用较多的搜索点,以牺牲较大的计算复杂度来保证粗略搜索的精确性,但是在其基础上分析最佳匹配的特性,可以减少不必要的搜索点。本文主要在两个方面减少搜索点点数:1)利用运动图像的中心偏置特性,对多层次大六边形搜索模板进行了优化,使其在原来 64 个点的基础上减少到 32 个点,其模板如图 6 所示;2)根据文献[5-7]所述,同一对象的预测运动矢量和最佳匹配运动矢量大概率在同一区域内,所以通过对预测运动矢量的方向进行判断来选择不同的搜索区域,能够减少 3/4 的搜索点,其模板如图 7 所示。

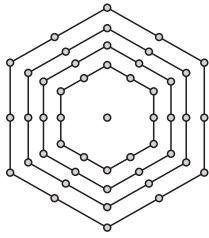


图 6 改进的多层次大六边形模板

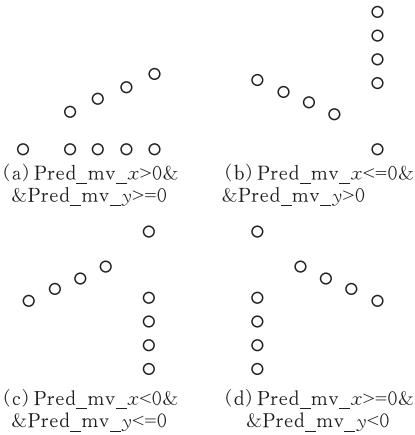


图 7 不同区域的多层次大六边形模板

- 1)若 $Pred_mv_x>0 \& \& Pred_mv_y\geq 0$,判定最佳运动矢量在第一象限,搜索模板如图 7a;
- 2)若 $Pred_mv_x\leq 0 \& \& Pred_mv_y>0$,判定最佳运动矢量在第二象限,搜索模板如图 7b;
- 3)若 $Pred_mv_x<0 \& \& Pred_mv_y\leq 0$,判定最佳运动矢量在第三象限,搜索模板如图 7c;
- 4)若 $Pred_mv_x\geq 0 \& \& Pred_mv_y<0$,判定最佳运动矢量在第四象限,搜索模板如图 7d。

3 仿真实验结果与分析

3.1 仿真实验平台与配置

为了检测优化后的算法性能,笔者将 H.264 标准测试 JM10.1 集成在 VS2008 平台上进行算法测试。实验使用的 PC 配置和编码器配置见表 1。

表 1 实验平台与参数配置

实验平台与配置	
系统	Win7 旗舰版
CPU	智能英特尔 i5-3210M 双核处理器
内存	4GB
测试视频系列	Coastguard、mobile、foreman、news、akiyo
视频格式	176×144 的 QCIF Y:U:V=4:2:0
编码器参数	FramesToBeEncoded=100,FrameRate=30.0,UseHadamard=1,
	SearchRange=16,NumberReferenceFrames=5, QP=28 其他参数采用默认值

表 2 改进算法与原算法的实验测试结果

测试序列	UMH 原算法			UMH 改进算法		
	PSNR/dB	BR/(kb·s ⁻¹)	MET/s	PSNR/dB	BR/(kb·s ⁻¹)	MET/s
coastguard	34.01	245.07	42.340	34.02	245.94	30.299
mobile	33.33	423.46	30.336	33.35	422.83	26.369
foreman	36.42	131.69	33.329	36.42	132.54	26.306
news	36.70	74.91	22.725	36.69	75.12	20.063
akiyo	38.28	29.61	21.377	38.26	29.78	19.388

表 3 改进算法变化测试结果

测试序列	变化差值			MET 变化/%
	Δ PSNR/dB	Δ BR/(kb·s ⁻¹)	Δ MET/s	
coastguard	0.01	0.87	-12.041	-28.4
mobile	0.02	-0.63	-3.967	-13.1
foreman	0.00	0.85	-7.023	-21.1
news	-0.01	0.21	-2.662	-11.7
akiyo	-0.02	0.17	-1.989	-9.3

3.2 实验结果与分析

算法改进前后的结果如表 2 和表 3 所示:从表 3 实验数据可以发现,改进后的算法与原算法相比,信噪比 PSNR 变化范围在-0.01~0.03 dB 之间,比特率 BR 变化范围相对比较小,运动估计的时间减少了 10%~30%,其中对运动激烈的视频序列 coastguar 和 foreman,运动估计时间减少了 28.4%和 21.1%;对运动程度一般的视频序列 news 和 akiyo,运动估计时间减少了 11.7%和 9.3%;对于运动缓慢的视频序列 mobile 而言,运动估计时间减少了 13.1%。对比几种不同运动程度的视频序列可以得出,随着视频序列运动程度的上升,其运动估计时间减少的比例增大。

4 结束语

通过对原 UHM 算法的介绍,就其中的不足进行了分析,分别对起始点预测和部分搜索模板做出了优化。实验证明在保证一定的信噪比的基础上,该算法降低了运动估计时间,提高了算法的整体效率,对比原始算法,具有明显优势。

电出版社,2007.

[2] 林永.基于 H.264 中视频压缩编码算法的研究与改进[D].南昌:江西师范大学,2012.

[3] 胡一凡. 基于 H.264 运动估计算法的视频监控研究[D].武汉:湖北工业大学,2017.

[4] 张斌. 基于 H.264/AVC 的快速运动估计算法研究[D].太原:中北大学,2016.

[5] 叶文龙,袁友伟,汪世瑜,周成武.H.264 块匹配运动估计中 UMHexagonS 算法的优化[J].计算机工程与应用,2011,47(25):133-136.

[6] 郑振东,王沛,应骏.H.264 JM 模型中运动估计算法及改进方案[J].中国图象图形学报,2007(10):1798-1801.

[7] 李浩炜,韩文波,田苗.H.264JM 模型中 UMHexagonS 算法优化[J].长春理工大学学报(自然科学版),2014,37(01):116-119.

[8] 王艳营.基于运动方向预测 UMHexagonS 算法的改进[J].计算机系统应用,2011,20(12):177-180,157.

[9] Zhao W , Xu S . Research and Optimization of UM-Hexagons Algorithm Based on H.264[C]// Multimedia Information Networking and Security (MINES), 2012 Fourth International Conference on. IEEE, 2012.

[10] Wang S , Xue Y . Optimization of UMHexagonS algorithm for AVS+[C]// IEEE International Symposium on Broadband Multimedia Systems & Broadcasting. IEEE, 2015.