

一种用于 AVS 视频编码的块运动估计算法^①

郑金华^{②*} 杨 平^{***} 杨志伟^{***}

(* 湘潭大学信息工程学院 湘潭 411105)

(** 上海中科计算技术研究所 上海 201203)

摘 要 通过对现有的视频图像编码中的帧间预测技术所用的运动估计算法的分析,提出了一种适用于 AVS 视频编码标准的快速而有效的块运动估计算法,该算法采用 5 个检测点的大菱形和 9 个检测点的小田字形双模板结构,通过调用大菱形模板和小田字形模板进行块运动匹配搜索,随着搜索的进行,大菱形模板可以被循环调用,但约定小田字形模板仅被调用一次。通过与全搜索算法和菱形搜索算法比较,这种新的块运动估计算法在保证编码性能几乎不变的情况下,编码复杂性能得到有效的降低。

关键词 帧间预测,音视频编解码技术标准(AVS),运动估计,大菱形模板,小田字形模板

0 引 言

音视频编解码技术标准(the audio video coding standard, AVS)是我国自主研发的下一代音视频压缩标准,与 MPEG 系列和 H.26x 系列标准一样,AVS 标准第二部分视频(AVS-P2)中的关键技术帧间预测,采用了基于块运动估计的算法。随着实时监控、视频会议等实时性较高的需求的增长,对这种算法的要求也日益提高。好的运动估计算法不仅能够保证图像清晰,而且还要求算法本身运行效率更高,计算复杂性更低,以能得到真正的应用。

在块运动匹配中,最简单且最可靠的搜索算法是全搜索法(full search, FS),但是其计算的复杂度也最高,因此,为了降低复杂度,人们相继提出了多种快速的算法,如三步搜索法(three-step search, TSS)^[1], 四步搜索法(four-step search, FSS)^[2], 菱形搜索法(diamond search, DS)^[3], 正方形-菱形搜索法(square-diamond search, SDS)^[4]等,希望在图像性能与计算复杂性之间找到一条可协调的途径。本文通过对现有运动估计算法的分析,提出了一种快速有效的搜索算法,该算法在不影响编码性能的前提下能有效地提高编码效率。

1 块运动估计搜索

AVS 采用的是混合视频编码框架,包括了变换、量化、熵编码、帧内预测、帧间预测、环路滤波等技术

模块,这也是当前主流的技术路线,其中帧间预测较为复杂也最为重要,也是多年来众多专家学者和工程技术人员研究的重点。

为了消除视频图像之间的时间冗余性,帧间预测采用基于块运动匹配的搜索技术对当前帧中的某个宏块(在其参考帧中的对应搜索框内)进行匹配搜索,找到最相似块(即预测块),通过保留原始块和预测块的残差信息来压缩码流,而要想快速而又准确地找到最相似块,则要求用好的搜索算法。

最可靠的搜索当选全搜索(FS),即用当前宏块的左上角像素点在其参考帧对应的搜索框内进行穷尽搜索,也就是与所有像素点进行匹配,计算它的搜索代价(search cost, SC),找出具有最小搜索代价(minimum search cost, MSC)的像素点,也就找到具有 MSC 的匹配块。

据统计,在视频图像的运动估计搜索时,最优像素点通常是在以当前点为圆心,两像素点的距离为半径的圆内^[3],如图 1 所示。基于此点,很多算法都考虑在此范围内进行第一层搜索,菱形搜索法(DS)是块匹配搜索算法中性能最优异的算法之一。考虑到粗定位和精确定位,它分别调用大小两种模板进行搜索,即大菱形搜索模板和小菱形搜索模板。先考虑大模板,当最小搜索代价(MSC)出现在大模板的中心时,则换成小模板进行搜索,具有最小搜索代价的搜索点即为最优点;否则改变中心位置,重复考虑大模板进行搜索。

^① 国家自然科学基金(60773047)和湖南省教育厅重点科研(06A074)资助项目。

^② 男,1963 年生,教授,博士生导师;研究方向:进化计算,并行处理,图像处理;联系人, E-mail: jhzheng@xtu.edu.cn (收稿日期:2007-11-14)

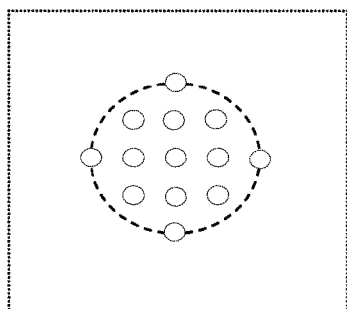


图 1 最优点分布

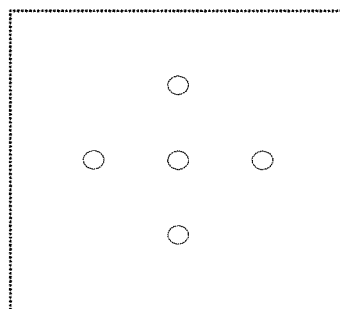


图 2 大菱形模板

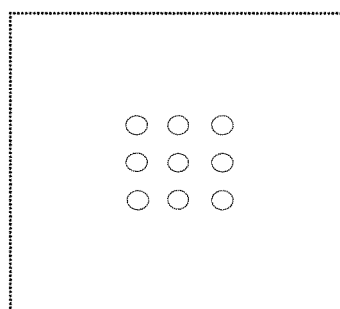


图 3 小田字形模板

对搜索代价(SC)的确定有多种,主要有绝对误差和(sum of absolute difference, SAD)^[5],平均绝对差值(mean absolute error, MAE),差值的平方和(sum of squared difference, SSD)等方式,为了提高效率,同时不失判断的准确度,在这里,SC 则考虑两个方面:一是计算原始块与预测块的 SAD,即两个宏块的匹配程度,如式

$$SAD = \sum_{i=0}^{15} \sum_{j=0}^{15} |f_o(x+i, y+j) - f_p(x+i, y+j)| \quad (1)$$

所示,其中 f_o 和 f_p 分别表示原始像素值和预测像素值, x 和 y 表示当前宏块左上角元素坐标;二是计算预测块的运动矢量代价(motion vector cost, MVC),即预测块距离原始块的位移代价,如式

$$MVC = \alpha \cdot \sum_{i=0}^1 f(MV_{o_i} - MV_{p_i}) \gg \beta \quad (2)$$

所示,其中 α 和 β 为常量, i 用来区分当前的坐标轴, MV_o 表示原始坐标, MV_p 表示预测坐标。函数 f 是一个恒定的查表函数, $\gg$ 表示右移。然后根据式

$$SC = SAD + \lambda \cdot MVC \quad (3)$$

计算原始宏块的搜索代价(SC)。其中 λ 是一可调参数。

菱形搜索(DS)算法的缺陷是当视频图像有剧烈运动时,大模板调用次数较多,而大模板的检测点有 9 个,因此搜索耗费的时间则较多,针对此问题,下节中提出的新算法,对此能有很好的改善。

2 大菱形小田字型搜索算法

大菱形小田字形搜索(diamond-farmland search, DFS)算法考虑到搜索的并行性,也使用了两个模板,如图 2 和图 3 所示,为了快速地锁定最优区域,大模板由 5 个检测点组成,而为了进行精确定位,小

模板则由 9 个检测点(即中心点与其周围的 8 个像素点)组成,算法具体步骤如下:

步骤 1 以当前搜索点为中心,以大菱形为模板,分别对 5 个检测点进行匹配计算,找出最小搜索代价(MSC)点,若 MSC 点在大菱形的 4 个顶点处,则将当前搜索点指向该 MSC 点,进行步骤 3;否则进行步骤 2。

步骤 2 以当前搜索点为中心,以小田字形为模板,对 9 个检测点进行匹配计算,找出 MSC 点,进行步骤 4。

步骤 3 以当前搜索点为中心,重复步骤 1。

步骤 4 当前 MSC 点即为最优匹配点。

由算法可知,在搜索过程中模板的改变最多只有两种情况,即大菱形模板到小田字形模板与大菱形模板到大菱形模板,分别如图 4 和图 5 所示。

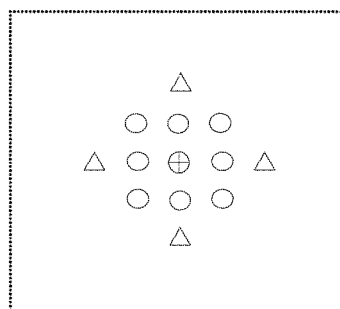


图 4 大菱形模板到小田字形模板

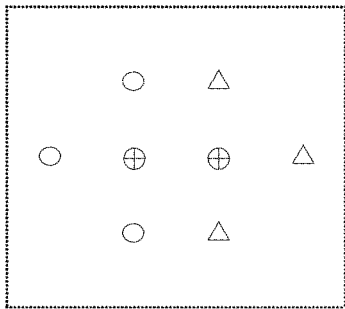


图 5 大菱形模板到大菱形模板

图中符号△代表原始模板的检测点,○代表预测模板的检测点,⊕代表前后两模板重叠的检测点。由图 4 和图 5 可知,大菱形模板到小田字形模板有一个重叠的检测点,大菱形模板到大菱形模板有两个重叠的检测点,这样在计算搜索代价(SC)时只需算一次。

3 实验分析

为了验证大菱形小田字形搜索(DFS)算法的合理性及实用性,从计算复杂性和图像质量两个方面将 DFS 与全搜索法(FS)和菱形搜索法(DS)比较。测试环境为 Pentium(R) 4 CPU 2.40G Hz, 256MB 内存,实验采用了 AVS 标准组提供的参考代码 rm52i 作为实验平台,用 ANSI C 编程实现。编码参数设置如表 1 所示。

表 1 编码参数配置

视频格式	cif
帧率	30fps
率失真优化 RDO	开启
编码格式序列	IPPP……
编码帧数	150
参考帧数	1
Hadamard 变换	关闭
帧间搜索模式	16 × 16
搜索范围	44 × 28

3.1 计算复杂性

因为在新算法中计算搜索代价(SC)是最主要的代价,所以为了更精确地统计复杂度,在这里统计宏块计算 SC 的次数。实验测试了 7 个序列,其中量化参数(quantization parameter, QP)设为 30,测试结果如表 2 所示。

由表 2 可知,对于固定搜索框,所有序列中 FS 计算 SC 的次数不变,并且最高,DFS 和 DS 则相对比

较少,而同时 DFS 又较 DS 少,这在视频序列图像越剧烈时,效果越明显。

表 2 各种序列下宏块计算 SC 的次数

序列	平均每个宏块的计算 SC 次数/总计算次数		
	FS	DS	DFS
hall	1232/72692928	12.392/731187	12.238/722100
foreman	1232/72692928	13.804/814511	12.848/758088
football	1232/72692928	18.248/1076700	15.8/932292
subway	1232/72692928	13.205/779147	12.694/748980
tempete	1232/72692928	12.927/762734	12.52/738711
bus	1232/72692928	13.619/803546	12.766/753252
mobile	1232/72692928	12.219/720995	12.082/712863

3.2 编码性能分析

编码性能考察的是图像的质量,在这里比较的参数是峰值信噪比(peak signal noise rate, PSNR)。我们将 DFS 与 FS 和 DS 在不同量化参数(QP)下进行比较。在此测试了 3 个序列,分别如图 6、图 7 和图 8 所示。

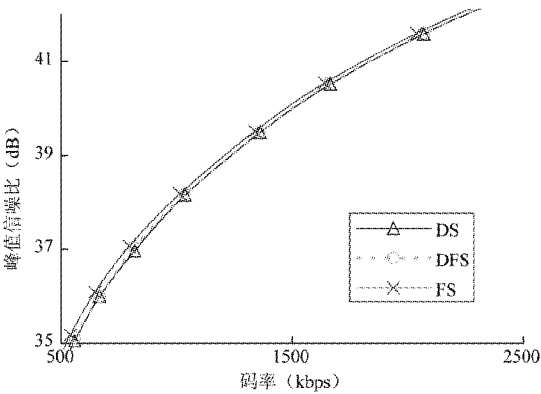


图 6 foreman 序列性能曲线

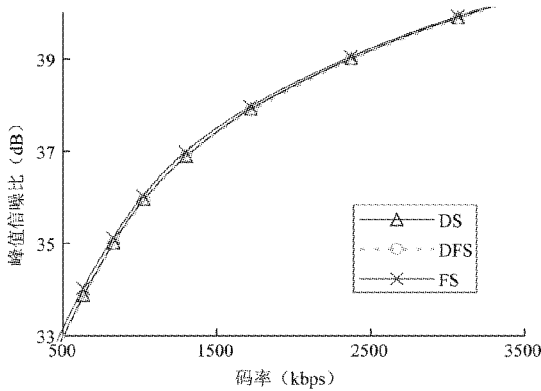


图 7 subway 序列性能曲线

因为搜索越准确,峰值信噪比(PSNR)值则越高,因此全搜索(FS)的 PSNR 被认为是高的,从图上可验证这一点,另外从图 6、图 7 和图 8 也可看出,DS 和 DFS 的两条曲线彼此接近,相差不大,这说明它们的编码性能基本相当。

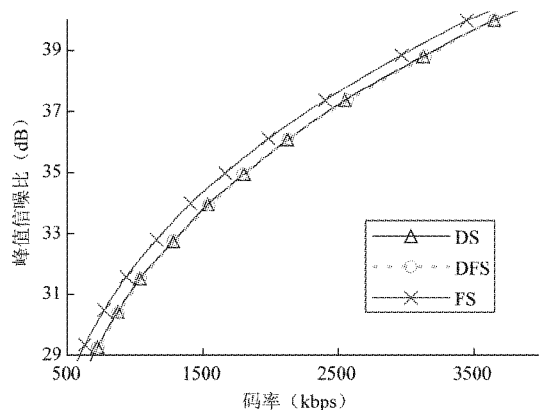


图 8 bus 序列性能曲线

4 结 论

针对 AVS 视频编码,本文提出了一种新的块运动估计搜索算法,通过循环调用 5 个检测点的大菱

形模板和调用 9 个检测点的小田字形模板实现,通过实验对新算法进行计算复杂性和编码性能的分析,与传统的全搜索(FS)法和菱形(DS)法相比,大菱形小田字形搜索(DFS)算法在保证编码性能基本不变的情况下,计算复杂性能得到有效的降低,并且图像越剧烈,效果越明显。

参考文献

- [1] Li R X, Zeng B, Liou M L. A new three-step search algorithm for block motion estimation. *IEEE Transactions on Circuits and Systems for Video Technology*, 1994, 4(4): 438-442
- [2] Po M L, Ma W C. A novel four-step search algorithm for fast block motion estimation. *IEEE Transactions on Circuits and Systems for Video Technology*, 1996, 6(3): 313-317
- [3] Zhu S, Ma K K. A new diamond search algorithm for fast block matching motion estimation. *IEEE Trans Image Processing*, 2000, 9(2): 287-290
- [4] 刘海峰,郭宝龙,冯宗哲.用于块匹配运动估计的正方形-菱形搜索算法. *计算机学报*, 2002, 25(7): 747-752
- [5] 刘东华,杨立志,陈益强.一种用于 AVS-M 视频编码的快速帧内预测模式判断算法. *中国图像图形学报*, 2006, 11(11): 1566-1569

A block motion estimation algorithm for AVS video coding

Zheng Jinhua^{*}, Yang Ping^{***}, Yang Zhiwei^{***}

(^{*} Institute of Information Engineering, Xiangtan University, Xiangtan 411105)

(^{***} Shanghai Division Institute of Computing Technology, Chinese Academy of Sciences, Shanghai 201203)

Abstract

The paper proposes a quick and efficient block motion estimation algorithm for inter prediction, a key technology in the audio video coding standard (AVS). The algorithm uses a double-pattern structure consisting of a large diamond pattern with five detecting dots and a small farmland pattern with nine detecting dots, and conducts the matched search for block motion by using the two patterns. For a matched block search, the former can be called many times and the latter can be called only once. Compared with the full search (FS) method and the diamond search (DS) method, the proposed algorithm can reduce the computational complexity and almost has the same performance as them.

Key words: inter prediction, audio video coding standard (AVS), motion estimation, large diamond pattern, small farmland pattern