

## 一种快速分类搜索运动估计新方法

初秀琴 李玉山 徐善锋

(西安电子科技大学电路CAD研究所, 西安 710071)

**摘 要** 在视频编码中, 基于块的运动估计算法被广泛应用。在保证估计质量的前提下, 为了降低运动估计算法的搜索次数, 提出了一种对于不同类型的块采用不同的搜索范围和搜索步骤的分类快速搜索(CFS)运动估计新算法。该算法首先对块进行分类, 然后确定其搜索范围和搜索步骤, 在应用分类搜索法时, 根据运动矢量的中心偏置特性, 将第1步和第2步的搜索窗采用 $5 \times 5$ 的窗口, 第3步采用 $3 \times 3$ 的窗口。结果表明, 该分类快速搜索新算法在运动矢量的估计质量上, 明显优于传统三步搜索法, 且搜索次数与传统三步搜索法相比, 降低了23%, 与全搜索法相比, 降低了91%。实验结果证明, 该算法尤其适用于快速运动、复杂运动序列的运动估计。与传统的全搜索法和三步搜索法相比, 其更适合于用硬件实现。

**关键词** 运动估计 块匹配算法 分类搜索

**中图法分类号:** TP391.41 **文章标识码:** A **文章编号:** 1006-8961(2002)10-1067-04

## A Fast Classified Search Algorithm for Motion Estimation

CHU Xiu-qin, LI Yu-shan

(Institute of Electronic CAD, Xidian University, Xi'an 710071)

**Abstract** For video coding, the algorithm of motion estimation based on blocks is widely used. Under the prerequisite of assuring estimation accuracy, a fast classified search method applying different search range and different search steps for different blocks is given in this paper. According to the criterion of minimization of mean square error (MSE) between two frames, two threshold values  $T_1$  and  $T_2$  are determined. Each block's MSE is compared with  $T_1$  and  $T_2$ . Then blocks' movement is classified into fast, slow and still moving. Different search range and different search steps are applied for different blocks. For fast moving blocks three steps search is used. For slow moving blocks two steps search is used. For still blocks one step search is used. Based on the center-biased property of the motion vectors, the first and second search step use the  $5 \times 5$  window and the third search step uses a  $3 \times 3$  window. The results show that this method is better in quality than 3SS and reduce 3SS search steps by 23% and FS search steps by 91%. Furthermore this algorithm is more suitable for VLSI implement as well.

**Keywords** Motion estimation, Block matching algorithm, Classified search

## 0 引言

运动估计是视频压缩编码中的关键技术, 通常在一个压缩编码方案中运动估计占总计算量的60%~80%, 因此寻找一种简便、快速、高效的运动估计算法成为视频编码领域的一个研究热点。目前

已出现了很多类运动估计算法, 如块匹配法、子波变换法、神经网络法、基于模型的运动估计法、像素匹配法等等, 其中, 块匹配法<sup>[1~6]</sup>因具有简单、实用的特点而得到广泛应用, 并且被大多数视频编码标准(如 MPEG, H. 261, H. 263)所采用。传统块匹配算法的基本思想是将当前的图象帧划分成 $N \times N$ 的子块, 并假定位于同一子块内的所有像素具有同样

基金项目: 国家自然科学基金(60172004); 北京大学视觉与听觉信息处理国家重点实验室基金(2001-03)

收稿日期: 2001-06-01; 改回日期: 2002-01-31

的位移,根据预定的匹配准则,在参考帧中预先确定一个搜索窗,搜索出与该子块最匹配的块,从而得出每个子块的运动矢量。实际传输的数据不是压缩了的像素数据,而是仅仅传输运动矢量和预测误差,这样可以大大降低实际传输的比特数。

全搜索块匹配算法(FS)虽然对运动矢量的估计比较精确,但由于它对搜索窗内的所有块进行搜索,因此计算量很大,不能满足实时性的要求。因而,一些快速运动估计算法,如三步搜索法(TSS)<sup>[1]</sup>、改进的三步搜索法(ITSS)<sup>[2]</sup>等便应运而生。这些算法与全搜索法相比,虽然性能有所下降,但计算量却大大减小。作者通过研究发现,运动矢量具有中心偏置特性,即运动偏移大多限制在围绕搜索窗中心的一个很小的区域内。根据这一特点,本文提出一种新的分类快速块匹配算法,该算法的最大特点是计算量小,便于硬件实现,且估计精度优于传统的三步搜索法。与全搜索法相比,计算量降低了91%,与传统的三步搜索法相比,计算量降低了23%,且估计的均方误差小于传统的三步搜索法。

## 1 新型分类快速搜索算法

通过大量的实验发现,现实中的图象序列,尤其是在可视电话、视频会议等低比特率的视频编码应用中,其运动矢量主要集中于以原点为中心的一个很小的区域内,即大多数图象序列的运动都是平稳的、缓慢的<sup>[2]</sup>。对于大多数的图象序列,将近80%的块可以看成是静止的或近似静止的。对于很多背景固定不变的序列图象来说,如果相机没有大的运动,那么图象中运动的物体只占很小一部分,比如可视电话中人的头部部分,而传统的三步搜索法对每一帧图象中的每一个子块都进行同样的搜索,这样就会造成计算上的浪费。

新型算法将克服这一缺点,即首先对图象中的块进行分类,即将块分为快运动、慢运动和基本静止3类,然后对不同的块进行不同的搜索;对于快速运动的块,进行三步搜索;而对于慢运动的块,进行两步搜索;对于基本静止的块,进行一步搜索。在该算法的三步搜索中,搜索窗的大小也不同于典型的三步搜索法,它是在第1步与第2步中都采用 $5 \times 5$ 的窗,第3步采用 $3 \times 3$ 的窗,这样由于第1步和第2步中有部分重复的搜索点,因此可以减小计算量。

为便于同其他的算法比较,该算法中图象的子

块也采用 $16 \times 16$ pixels。在块匹配算法中,匹配准则也有多种,如均方误差(MSE)、绝对帧差(MAD)等,其中绝对帧差用得最多,因为它不需要任何乘法运算,而性能与均方误差准则几乎一致。该算法中,考虑到硬件实现的简化性,在进行三步搜索、两步搜索和一步搜索时,均采用绝对帧差作为匹配准则,即

$$MAD_{(m,n)}(x,y) = \frac{1}{256} \sum_{i=0}^{15} \sum_{j=0}^{15} |f_k(m-i, n+j) - f_{k-1}(m+x+i, n+y+j)| \quad (1)$$

其中, $f_k(i,j)$ 和 $f_{k-1}(i,j)$ 分别是当前帧(第 $k$ 帧)和前一帧(第 $k-1$ 帧)中 $(i,j)$ 处的像素值,块 $(m,n)$ 是指该块的左上角位于 $(m,n)$ 。

在对块进行分类时,门限 $T_1$ 和 $T_2$ 是最关键的因素。为了选择合适的门限,引入两个连续帧中块的方差 $P_k$ ,定义为

$$P_{k(m,n)} = \sum_{i=0}^{15} \sum_{j=0}^{15} |f_k(m+i, n+j) - f_{k-1}(m-i, n+j)|^2 \quad (2)$$

图1是两个连续帧中块的方差曲线,其中较平缓的部分可解释为图象中静止部分由于图象捕捉过程所产生的噪声引起的;尖峰部分是由于物体的运动引起的。因此曲线中最小的尖峰值处被确定为门限 $T_1$ ,平缓部分的最大值处被确定为门限 $T_2$ 。

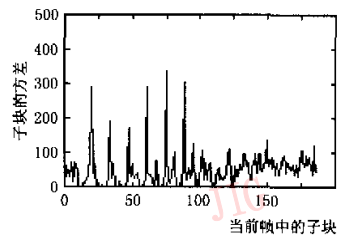


图1 两个连续帧中块的方差曲线

新方法的具体算法如下:

(1) 首先按式(2)计算出当前帧中的子块与前一帧中对应的子块之间的均方误差,根据所有子块的 $P_k$ 值,确定出两个门限 $T_1$ 和 $T_2$ 。

(2) 判断当前帧中的块 $(m,n)$ 与前一帧中对应的块 $(m,n)$ 之间的误差 $P_{k(m,n)}$ ,如果 $P_{k(m,n)}$ 大于或等于门限 $T_1$ ,则认为该块为快速运动的块,转第3步;如果 $P_{k(m,n)}$ 值小于门限 $T_1$ ,但是大于等于门限 $T_2$ ,则认为该块为慢运动的块,转第4步;如果 $P_{k(m,n)}$ 值小于门限 $T_2$ ,则认为该块为基本静止的块,转第5步。

(3) 进行三步搜索：

①在搜索区的中心按图 2(a)所示的模式，搜索  $5 \times 5$  窗口的中心点和四周的 8 个顶点，如果这一步得到的最小误差位置在中心点，则转第③步；否则执行下一步；

②搜索窗保持  $5 \times 5$ ，根据上一步的搜索结果进行下面的搜索：如果上一步搜索的最小点位于前一个搜索窗的顶点，则按图 2(b)的模式进行 5 次搜索（图中黑色的点为新增加的点，灰色的点为上一步已经搜索过的点）；如果上一步搜索的最小点位于前一个搜索窗的水平或垂直轴上，则按图 2(c)的模式进行 3 次搜索（图中黑色的点为新增加的点，灰色的点为上一步已经搜索过的点）；

③搜索窗以前面搜索到的最小误差点为中心缩小为  $3 \times 3$ ，按图 2(d)模式在 9 个点中找出最小误差点，并作为估计的运动矢量。

(4) 进行两步搜索：

①在搜索区的中心，按图 2(a)所示的模式搜索  $5 \times 5$  窗口的中心点和四周的 8 个顶点，得到最小误差点；

②搜索窗以前面搜索到的最小误差点为中心，缩小为  $3 \times 3$ ，按图 2(d)模式在 9 个点中找出最小误差点，并作为估计的运动矢量。

(5) 进行一步搜索：在搜索区的中心，按图 2(d)所示的模式搜索  $3 \times 3$  窗口的中心点和四周的 8 个顶点，得到最小误差点，并作为估计的运动矢量。

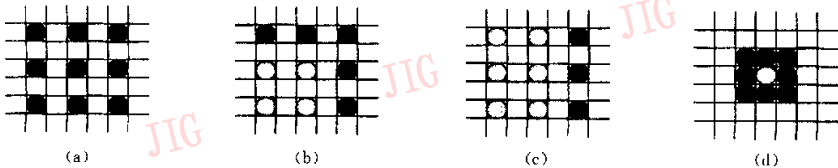


图 2 搜索模式示意图

2 仿真结果

用分类快速搜索新算法(CFS)和经典全搜索法、三步搜索法对 4 种视频序列进行对比测试。

4 种测试视频序列分别为 Flower、Tennis、Mother-daughter 和 Chemical 序列，测试序列的图象格式为 TIFF 格式，每帧图象的像素数目为  $352 \times 240$ 。四种测试序列包含静止、慢运动、快速运动和相机移动以及放大缩小等多种运动方式，其中 Flower、Mother-daughter 和 Chemical 序列主要是慢运动，而快速运动在其中占的比例很小；Tennis 主要是快速运动和一些复杂的运动，包括旋转和缩放。在新算法中，匹配误差采用平均绝对误差 MAD，块的大小采用  $16 \times 16$  pixels。对每一种测试序列，取了前 80 帧进行实验。

表 1、表 2 分别列出了 4 种测试序列，采用不同算法所用的搜索次数和得到的均方误差。通过比较可以看出，对于 4 种序列，新算法在搜索次数上比全搜索法降低了 91%，比传统的三步搜索法降低了 23%；在估计质量上，也比传统搜索法有了一定的提高，对 4 种序列，新算法的均方误差比传统的三步搜索法分别降低了 9.1%，3.4%，2.1%和 2.6%。

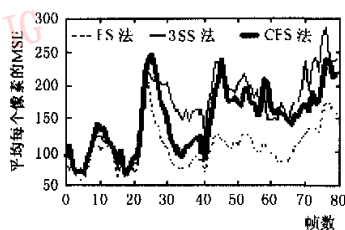
表 1 平均搜索次数比较

算法	测试序列			
	Tennis	Flower	Mother-daughter	Chemical
FS	225	225	225	225
3SS	25	25	25	25
CFS	19.06	19.24	18.6	19.2

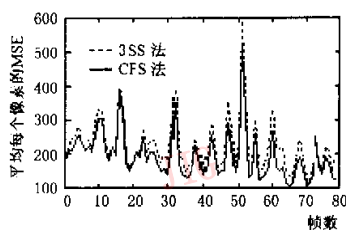
表 2 均方误差比较

算法	测试序列			
	Tennis	Flower	Mother-daughter	Chemical
FS	108.02	279.21	20.30	89.64
3SS	166.16	326.34	28.22	98.82
CFS	150.96	315.24	27.64	96.30

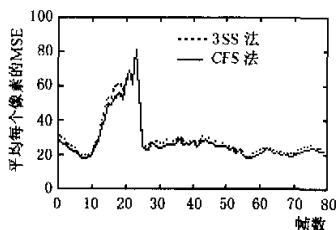
图 3(a)表示了全搜索法、传统的三步搜索法和新算法对 Tennis 序列进行运动估计所得到的每帧中平均每个像素点的均方误差，从图中曲线看出，用新算法所得到的均方误差明显小于传统三步法，且对前 25 帧，用新算法估计得到的均方误差曲线与全搜索法估计得到的均方误差曲线基本重合；图 3(b)、(c)、(d)分别表示了传统的三步搜索法和新算法对 Flower 序列、Mother-daughter 序列和 Chemical 序列进行运动估计所得到的每帧中平均每个像素点的均方误差，这三种序列主要是慢运动，从图中曲线可以看出，实验的 80 帧中，用新算法估计比用传统三步法估计得到的均方误差均有下降。



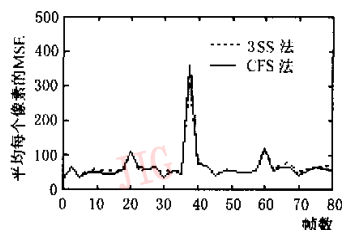
(a) Tennis 序列



(b) Flower 序列



(c) Mother-daughter 序列



(d) Chemical 序列

图3 不同搜索算法平均每个像素的MSE比较

### 3 结束语

由上面的实验结果可以看出,本文新算法由于充分利用了视频序列运动矢量的中心偏置特性,在搜索步骤中采用了分类搜索的方法,且对于快速运动的块在第1步与第2步搜索中采用了相同大小的窗,因此其计算量比全搜索法降低了91%,比传统的三步搜索法降低了23%;在估计质量上,本文新算法优于传统的三步搜索法,对4种实验序列的估计结果进行对比看出,对Tennis序列,本文新算法比传统三步搜索法估计的均方误差更低,这一点说明,新算法尤其适用于快速运动、复杂运动序列的运动估计。与传统的全搜索法和三步搜索法相比,该算法更适合于用硬件实现。

### 参考文献

- 1 Koga T, Iinuma K, Hirano A *et al.* Motion compensated interframe coding for video conferencing [A]. In: Proc. Nat. Telecommun. Conf[C]. New Orleans, LA, 1981:531~535.
- 2 Po Lai-Man, Ma Wing-Chung. A novel four-step search algorithm for fast block motion estimation[J]. IEEE Trans. Circuits Syst. Video Technol., 1996, 6(3):313~317.
- 3 Li R, Zeng B, Liou M L. A new three-step search algorithm for block motion estimation[J]. IEEE Trans. Circuits Syst. Video Technol., 1994, 4(4):438~442.

- 4 Liu B, Zaccarin A. New fast algorithms for the estimation of block motion vectors[J]. IEEE Trans. Circuits Syst. Video Technol., 1993, 3(2):148~157.
- 5 Chan Y L, Siu W C. New adaptive pixel decimation for block motion vector estimation[J]. IEEE Trans. Circuits Syst. Video Technol., 1996, 6(1):113~118.
- 6 Chan Yui-Lan, Siu Wan-Chi. Block motion vector estimation using edge matching: An approach with better frame quality as compared to full search algorithm [J]. IEEE International Symposium on Circuits and System, 1997, 2(B):1145~1148.



初秀琴 1972年生,西安电子科技大学讲师,目前在攻读电路与系统专业博士学位。主要研究方向为计算机视觉、图象处理。



李玉山 1946年生,西安电子科技大学教授,博士生导师。主要研究方向为计算机视觉、图象处理、图象处理专用集成电路设计。

徐善锋 1977年生,西安电子科技大学电路与系统专业硕士研究生。主要研究方向为计算机视觉、图象处理。