

中图分类号: TN919.81 文献标志码: A 文章编号: 1006-8961(2011)06-0933-06

论文索引信息: 裴琴, 刘国繁, 徐美清. 基于块特征分类的运动估计算法 [J]. 中国图象图形学报, 2011, 16(6): 933-938

基于块特征分类的运动估计算法

裴琴¹⁾, 刘国繁²⁾, 徐美清¹⁾

¹⁾(湘潭大学信息工程学院, 湘潭 411105) ²⁾(湖南工程学院, 湘潭 411104)

摘要: 运动估计作为 H. 264 中最重要最耗时的部分, 大量的算法研究都是通过减少搜索点数来降低计算量。视频序列中, 人们感兴趣的往往是运动的物体, 背景部分一般拥有大量的静止块, 帧间背景间的运动可以认为完全是由摄像机的运动引起的。利用背景块与前景块的图像特征结合 MV 预测提出一种新算法。该算法包括以下策略: 静止块判断、背景块、前景块和边缘块的检测、起点预测和搜索窗口的改进, 同时引进了边缘检测技术。实验结果表明, 该算法与全搜索算法 (FS)、UMHex、SUMHex 相比, 搜索时间减少了, 而搜索精度接近 FS, 码率只有很小的增加。

关键词: 运动估计; 背景块; 前景块; 起点预测

Motion estimation based on block feature classification

Pei Qin¹⁾, Liu Guofan²⁾, Xu Meiqing¹⁾

¹⁾(college of Information Engineering, Xiangtan University, Xiangtan 411105 China)

²⁾(Hunan Institute of Engineering, Xiangtan 411104 China)

Abstract: Motion estimation is the most important and time-consuming part in H. 264. A large number of algorithms are based on the reduction of search points to reduce the computational complexity. In video sequence, people are often interested in moving objects. There are many static blocks in background, and the movement between the two different background of inter-frames is often considered as a result of camera moving. According to image feature of background block and foreground block and using MV prediction, we propose a new algorithm. It is based on these strategies, including static block determination, background-block, foreground-block and edge-block detection, start point prediction and the improvement of search window. Comparing with fast-full search algorithm (FFS), UMHex and SUMHex, experimental results reveal that our algorithm can reduce motion estimation time, while the PSNR value is kept approximate to FFS, with a slight increase in bit rate.

Keywords: motion estimation; background-block; foreground-block; start point prediction

0 引言

运动矢量对于减少时间冗余起到很重要的作用, 因此, 运动估计被广泛运用到视频编码和处理系统中^[1]。同时, 因为要多次计算匹配误差, 所以基于块匹配的运动估计是 H. 264 视频编码中最耗时

的部分。高质量的运动估计算法是高效视频编码的前提和基础。

为了减少计算的复杂性, 快速搜索算法被提出。

1) 通过优化匹配方程减少计算量^[2]; 2) 减少搜索点数。减少搜索点数因为最直观而被算法研究者作为最佳途径。在这类快速搜索算法中, 固定模板搜索法前人已经做出了很多的成效。为了解决固定搜索

收稿日期: 2009-10-29; 修回日期: 2010-05-23

基金项目: 湖南省自然科学基金项目(07JJ6112); 湖南省教育厅资助科研项目(05C588)。

第一作者简介: 裴琴(1985—), 女。湘潭大学信号与信息处理专业硕士研究生, 主要研究方向为视频信号处理。

E-mail: zimuhoniao@163.com。

法没有很好地利用图像本身特征的问题, Chung 等人在文献[3]利用空间相关性提出了一种预测搜索窗口方法,在文献[4]中, Nikitidis 等人提出一种随机向量场模型,文献[5]中 Ka ho Ng 等人利用不同块尺寸的最优运动矢量的特征来决定下一步搜索,文献[6]中 Hou 等人采用自底向上的运动矢量合并来提高运动估计的速度。研究者开始利用序列图像的时空相关性进行预测搜索起点方法^[7]、背景图像快速检测法以及多预测点搜索法等和搜索模板相结合的混合算法研究。

提出一种根据编码块的不同属性选择搜索方法的运动估计算法。在算法中,首先进行静止块的判决;其次利用贝叶斯模型进行背景块、前景块的判断,同时利用 sobel 算子判断边缘块;最后对不同的块类型用不同的搜索起点预测法并结合改进的搜索窗口进行模板搜索。实验结果证明,该算法与全搜索算法(FS)、UMHexagons 以及 SUMHexagons 相比,减少了运动估计时间,同时保证了搜索精度。

1 算法理论基础

大多数快速搜索法的搜索策略都基于这一假设前提:匹配误差随着搜索点接近全局最低点而单调减少。然而在实际的视频信号中,在一个大搜索范围内,这种单调性并不是唯一的,因此,这一假设常使匹配点搜索陷入局部最优。如图 1 所示,横向坐标为检测点到全局最优点的搜索距离。

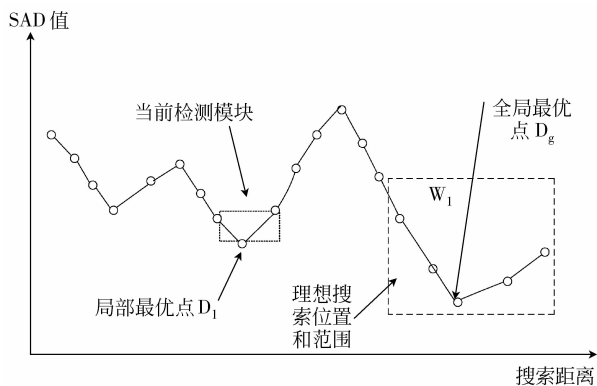


图 1 SAD 值随搜索距离变化的 2 维曲线图

Fig. 1 SAD error 2D curve with the search distance change

点 D_1 是位于模板搜索中心的局部最优点,那么下一步是以这个点为中心进行搜索得到全局最优点。但是,可以看到,全局最优点 D_g 离点 D_1 较远,除

非这个搜索窗口足够大,而且最后一步的搜索模板计算的点足够的多以至于能覆盖到全局最优点 D_g , 否则,这种方式的搜索极有可能会陷入局部最优。从图 1 中可以得出结论 1: 在包含全局最优点的一个小范围内 (W_1), 匹配误差是随着搜索点接近全局最优点而单调递减的。

根据结论 1, 采取窗口策略如下: 1) 起始搜索点尽可能地接近全局最优点; 2) 搜索窗口尽可能得小。

视频运动对象的运动信息对人眼视觉最敏感,也是导致视频压缩编码、传输过程中数据大的主要因素,而同一背景的各部分往往具有一致属性,因此区分出运动目标和背景,然后分别采用复杂度和搜索精度不同的搜索策略可以减少整个搜索时间。物体具有完整性,运动物体的轮廓包含有对象形状的重要信息。含有运动物体轮廓的边缘块,常因匹配的低精确性,在运动补偿时丢失大量的信息而使边缘模糊。因此高匹配性的边缘块运动估计对视频压缩质量也很重要。

2 算法介绍

1) 静止块判断 在许多实时视频和低复杂度运动序列中,有很多块可以认为是静止块。全搜索算法的统计结果表明,在宏块划分中, 16×16 块模式占了很大比例;在亚宏块划分中, 8×8 模式有相当大的比例^[8]。

研究分析指出,静止块的绝对误差和(SAD)远低于非静止块在(0,0)位置的 SAD 值; 16×16 和 8×8 静止块的平均 SAD 值分别为 $450 \sim 1000$ 和 $150 \sim 350$ ^[8]。设定阈值 $T_1 = 460$, $T_2 = 160$ 来分别判断 16×16 和 8×8 静止块。若当前块被判断为静止块,结束运动估计,否则进入步骤 2)。

2) 背景块、前景块和边缘块的判断 区分出前景块和背景块,可以对每个像素点属于前景和背景的概率分别做一个估计,即需估计出前景像素和背景像素的前验概率及后验概率,由于像素点数据分布无法事先确定,因此采用非参数核密度法进行估计^[9]。

(1) 计算背景和前景的核密度概率 $P(f/\omega_1)$ 和 $P(f/\omega_2)$ 。

假设 ω_1 表示背景像素, ω_2 表示前景像素,则背景和前景像素的先验概率分别为 $P(\omega_1)$ 和 $P(\omega_2)$, 背景和前景像素的核密度概率分别为 $P(f/\omega_1)$ 和

$P(f/\omega_2)$, 选择高斯函数作为核密度函数, 则有:

$$P(f/\omega_1) = \frac{1}{n} \sum_{j=1}^n \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(f-g_j)^2}{2\sigma_1^2}} \quad (1)$$

$$P(f/\omega_2) = \frac{1}{m} \sum_{j=1}^m \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(f-k_j)^2}{2\sigma_2^2}} \quad (2)$$

式中, σ_1 和 σ_2 分别为背景和前景的核函数带宽; f 为待检测的像素; g_j 和 k_j 分别为背景和前景样本集中的像素点; n 和 m 分别为背景样本集和前景样本集中的样本总数。

用相距 2 帧的图像之间对应像素点的差值来获取样本点:

$$|f_i(x, y) - f_{i+2}(x, y)| < T_g, g_j = f_{i+2}(x, y) \quad (3)$$

$$|f_i(x, y) - f_{i+2}(x, y)| \geq T_g, k_j = f_{i+2}(x, y)$$

式中, $f_i(x, y)$ 为当前帧中的像素灰度值, $f_{i+2}(x, y)$ 为与当前帧相隔 2 帧的图像灰度值, T_g 为事先设置的判决阈值。

核函数的带宽 σ_i 由公式(4)决定。

$$\sigma_i = \frac{m_i}{0.68\sqrt{2}} \quad (i = 1, 2) \quad (4)$$

式中, m_i 为对应样本集与当前帧对应像素点差值绝对值的中间值, 即

$$m_1 = \text{med}_i |f_i(x, y) - g_j(x, y)| \quad (5)$$

$$m_2 = \text{med}_i |f_i(x, y) - k_j(x, y)|$$

(2) 估计背景和前景的先验概率 $P(\omega_1)$ 和 $P(\omega_2)$

$$P(\omega_1) = \frac{M \times N - (S_2 - S_1)/255}{M \times N} \quad (6)$$

$$P(\omega_2) = \frac{S_2 - S_1}{255 \times M \times N} \quad (7)$$

式中, $M \times N$ 为待检测块的尺寸, S_1 和 S_2 计算公式分别为

$$S_1 = \sum_{k=1}^{M \times N} y_{bk}, \quad S_2 = \sum_{k=0}^{M \times N} y_{fk} \quad (8)$$

式中, y_{bk} 和 y_{fk} 分别表示二值化背景图像和二值化前景图像。

(3) 计算阈值 T

$$T = \frac{P(\omega_1/f)}{P(\omega_2/f)} =$$

$$\left(\frac{P(f/\omega_1)P(\omega_1)}{\sum_{j=1}^2 P(f/\omega_j)P(\omega_j)} \right) / \left(\frac{P(f/\omega_2)P(\omega_2)}{\sum_{j=1}^2 P(f/\omega_j)P(\omega_j)} \right) =$$

$$\frac{P(f/\omega_1)P(\omega_1)}{P(f/\omega_2)P(\omega_2)}$$

令

$$t_1 = \frac{P(f/\omega_1)}{P(f/\omega_2)}, \quad t_2 = \frac{P(\omega_1)}{P(\omega_2)}, \quad T = t_1 \times t_2 \quad (9)$$

(4) 根据所得阈值的大小进行判断, $T > 1$ 时为背景像素, 其余为前景像素。

(5) 背景块、前景块以及边缘块判断。

在进行边缘检测时, 必须先去除原始图像中的噪声。在运动估计时, 并不需要精确地检测到运动物体的边缘, 因此, 在 5×5 的窗口中进行加权中值滤波就可得到平滑后的图像。为简便起见, 采用 Sobel 梯度卷积算子进行边缘检测:

$$g_x = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}, \quad g_y = \begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad (10)$$

定义 $f'_i(x, y)$ 为与原始块同位置的平滑块中的像素点, 以该点为中心, 计算其邻域

$$S_i(x, y) = |f'_i(x, y) * g_x| + |f'_i(x, y) * g_y| \quad (11)$$

记 $B_i(x, y)$ 为

$$B_i(x, y) = \begin{cases} 1 & S_i(x, y) > T_e \\ 0 & \text{其他} \end{cases} \quad (12)$$

若当前块满足 $\sum_{x=0}^M \sum_{y=0}^N B_i(x, y) \geq T_{\text{count}}$, 记判断参数 $E = 1$, 认为当前块为边缘块; 否则 $E = 0$ 。其中, T_e 为像素梯度阈值, T_{count} 为边缘点数阈值。

计算当前块中 4 个顶角像素点的 T 值, 如图 2 所示。

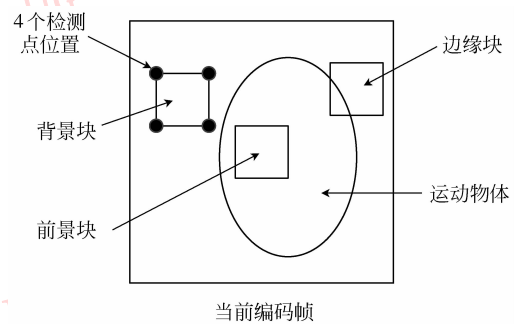


图 2 判决像素点的选取以及块的分类

Fig.2 Detection pixels selection and block classification

若 4 个点都满足 $T > 1$, 进行边缘检测得 $E = 0$, 则当前块为背景块, 进入步骤 3);

若 4 个点的 T 均为 $T \leq 1$, 进行边缘检测得 $E = 0$, 则当前块为前景块, 进入步骤 4);

其余情况,当前块为含有边缘的边缘块,进入步骤 5)。

3) 背景块运动估计 由于摄像机的高采样频率,被拍摄背景本身可以认为是不动的,帧间背景的运动可以认为完全是由于摄像机的运动引起的,而目标能独立运动^[4]。对背景图像的分析统计发现,背景图像的相邻块的运动矢量之间有很强的空间相关性,相邻块具有相同运动矢量的比例为 85.7%。因此可用与当前块相邻的块的运动矢量来预测其搜索起点 D_K 。 $D_K = P_1 + d$ 。其中, P_1 为当前块的左边块或上边块的匹配点坐标, d 的确定方法为:若 P_1 为左边块,则 $d = M$;若 P_1 为上边块,则 $d = -N$ 。 M 、 N 为块 P_1 的尺寸 $M \times N$,如图 3 所示。搜索在 5×5 搜索窗口内用小菱形搜索模板搜索。

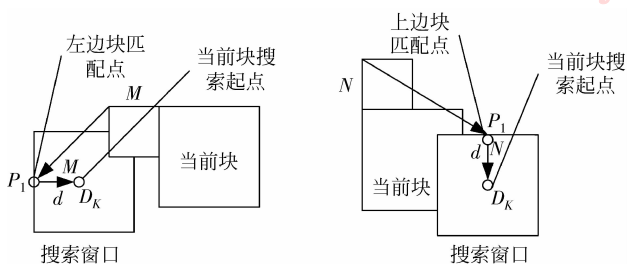


图 3 当前背景块搜索起点与搜索窗口

Fig. 3 Starting point and search window for current background block

4) 前景块运动估计 由于充分运用了时间域与空间域的相关性,预测运动矢量不会偏离真实运动矢量太远。

为了减少对起始搜索点搜索的计算量,直接用预测运动矢量的方向点作为起始搜索点。

对于当前编码块 B_c ,利用以下 4 种预测模式求当前块的预测运动矢量 MV_p :

帧内中值预测、上层预测、前帧对应块运动矢量预测、时间域的邻近参考帧运动矢量预测,将以上 4 个 MV_p 分别记为 MV_{p1} 、 MV_{p2} 、 MV_{p3} 、 MV_{p4} ,其指向的点分别为 d_1 、 d_2 、 d_3 、 d_4 ,改进的搜索区域 S 定义为

$$S = S_1 \cup S_2 \cup S_3 \cup S_4$$

其中, S_i ($1 \leq i \leq 4$) 为分别以 d_i ($1 \leq i \leq 4$) 为中心的 $(2W+1) \times (2W+1)$ 搜索区域,如图 4(a) 所示。

当前块与其相邻块之间的运动矢量存在差异,定义其为 D ,选择 $D=2$ 作为改进的搜索窗口范围,

即窗口大小为 5×5 ^[2]。改进后的搜索窗口如图 4(b)。在以上 4 种预测中,上层块预测性能最优。选取上层预测运动矢量指向的点 d_2 为起始搜索点。搜索用六边形搜索模板 (HEXBS)。

5) 边缘块运动估计 人类的视觉系统很敏感,极容易察觉到运动物体边缘的运动补偿是否精确。物体边缘具有闭合性,而且在大多数情况下,边缘形状不会发生变化。对边缘块的运动估计可以采用简单搜索方法和高准确性的匹配函数^[10]。

定义边缘块的搜索范围为当前块的左边块、上边块、右上块以及上层块的最大运动矢量值:

$$search_range = \max(MV_{左}, MV_{上}, MV_{右上}, MV_{Uplayer})$$

在参考帧中,选取上层预测运动矢量指向的点为中心,以 $sear_chrage$ 为半径,提取范围内所有与当前边缘块尺寸相同的边缘块,计算其匹配性,匹配性最强的即为最佳匹配块。

匹配性判断方法为:提取候选边缘块中边缘的最大梯度值点,与当前边缘块中的最大梯度值点进行匹配,拥有最多灰度相关性最强的最大梯度值点的边缘块即为当前块的最佳匹配块。

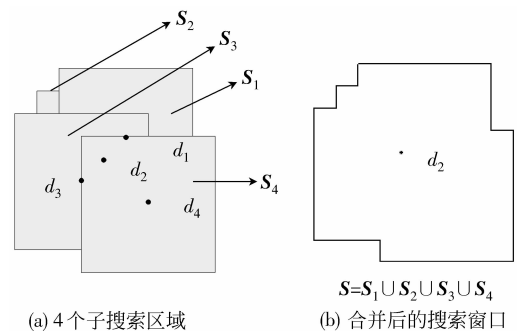


图 4 改进的搜索窗口

Fig. 4 Improved search window

3 实验结果与分析

测试序列选取 3 个 QCIF 格式视频,分别为 Foreman、Mobile 和 Coastguard。图像序列以 30 帧/s 的速度进行编码,原始的搜索窗口范围为 16,采用量化参数为 18、24、28 和 32 进行测试,设置 T_g 为 15, T_{count} 为 8, T_e 为 40,参考帧数为 5。将本文算法与 FS、UMHex、SUMHex3 种算法比较,结果如表 1 所示。

表 1 本文算法与 FS、UMHex、SUMHex 的比较

Tab.1 Comparison of experimental results among proposed algorithm, FS, UMHex, and SUMHex

算法	测试序列	量化参数	PSNR 变化	码率 变化	ME 时间下降 度/%
FS	Foreman	18	-0.002	+3.78	83.71
		24	-0.003	+3.64	84.67
		28	-0.002	+3.57	84.73
		32	-0.001	+1.96	84.25
	Mobile	18	-0.001	+4.51	85.32
		24	-0.001	+1.31	84.86
		28	-0.002	+1.11	84.81
		32	-0.002	+0.91	83.96
	Coastguard	18	-0.005	+2.37	88.64
		24	-0.006	+1.77	88.56
		28	-0.011	+1.33	88.94
		32	-0.008	+1.45	89.54
Foreman	18	+0.006	+1.45	62.27	
	24	-0.001	+0.79	63.15	
	28	+0.062	-0.10	60.56	
	32	+0.006	+0.02	60.97	
UMH	Mobile	18	+0.002	-0.09	63.10
		24	+0.010	+0.87	61.92
		28	+0.490	-1.38	61.97
		32	+0.020	+0.34	59.50
	Coastguard	18	-0.002	+1.58	73.26
		24	+0.025	+0.37	72.87
		28	+0.005	-0.37	71.79
		32	+0.027	+0.09	71.61
	Foreman	18	+0.034	-1.73	48.50
		24	+0.026	-0.24	50.54
		28	+0.047	-0.07	48.45
		32	+0.102	-0.17	49.48
SUMH	Mobile	18	+0.000	-2.36	55.69
		24	+0.068	-1.06	54.86
		28	+0.513	-0.85	55.22
		32	+0.020	-0.63	55.13
	Coastguard	18	-0.012	+1.27	60.33
		24	+0.004	+0.77	61.65
		28	+0.015	+0.03	60.96
		32	+0.036	+0.44	62.72

从表 1 中,可以看出,虽然在算法中引入了像素点的性质判断以及边缘检测,但由于只进行少量点的检测计算,以及采用了静止块策略和小范围的搜

索窗口,整体上运动估计时间下降,分别减少了 83%、60% 和 48% 以上。提出的算法的 PSNR 和 FS 的相比,只有极小的降低,并且略高于 UMH 和 SUMH 的。和其他算法相比,总的来说,码率增加很少。

4 结 论

利用视频序列帧的图像特征,进行静止块判决、背景块、前景块和边缘块判断,并引进了不同的起点搜索方法和改进的搜索窗口。该算法对于视频监控和视频会议这种含有大量静止块视频以及镜头移动的全局运动视频有很好运动估计性能。对运动剧烈的视频序列,由于采取了更为精细的窗口策略和边缘块策略,搜索时间大大减少,同时也获得了较好的搜索精度。

该算法还有很多可以改进的地方,比如可充分利用物体边缘的完整性,进行更加简单的背景块与前景块的判断;采用更加精确的边缘匹配算法;进行更好的起点预测等。

参考文献 (References)

- [1] Dufaux F, Moscheni F. Motion estimation techniques for digital TV: a review and a new contribution [J]. Proceedings of the IEEE, 1995, 83(6):858-876.
- [2] Nguyen Viet Anh, Tan Yap peng, Efficient block-matching motion estimation based on integral frame attributes [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(3):375-385.
- [3] Chung Kuoliang, Chang Lungchun. A new predictive search area approach for fast block motion estimation [J]. IEEE Transactions on Image Processing, 2003, 12(6):648-652.
- [4] Nikitidis S, Zafeiriou S, Pitas I. Camera motion estimation using a novel online vector field model in particle filters [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, 18(8):1028-1039.
- [5] Ng Kaho, Po Laiman, Wong Kaman. Multiple block-size search algorithm for fast block motion estimation [C]//7th International Conference on Information, Communications and Signal Processing. Macau, China: ICICS, 2009:1-4.
- [6] Hou Kaichung, Chen Meijuan, Hsu Chingting. Fast motion estimation by motion vector merging procedure for H.264 [C]// IEEE International Conference on Multimedia and Expo. Washington, DC, USA: IEEE, 2005:1444-1447.

- [7] Choudhury A Rahman. UMHexagons algorithm based motion estimation architecture for H. 264/AVC [C] // Fifth International Workshop on System-on-Chip for Real-Time Application, Washington, DC, USA: IEEE 2005: 207-210.
- [8] Chen Chun, Yang Zhi, Bo Jiajun, et al. A fast variable - block-size motion estimation algorithm for H. 264 [J]. Journal of Image and Graphics, 2007, 12 (2): 272-276. [陈纯, 杨智, 卜佳俊, 等. 一种基于 H. 264 的可变块快速运动估计算法 [J]. 中国图象图形学报, 2007, 12 (2): 272-276.]
- [9] Liu Yingxia, He Changwei, Wang Xin. Bayesian modeling for dynamic background detection [J]. Journal of System Simulation, 2007, 19 (21): 5042-5058. [刘英霞, 贺长伟, 王欣. 基于贝叶斯模型的动态背景检测 [J]. 系统仿真学报, 2007, 19 (21): 5042-5058.]
- [10] Li Haichao, Zhang Guangjun. Fast stereo edge-matching algorithm based on corner-guidance [J]. Journal of Beijing University of Aeronautics and Astronautics, 2007, 33 (5): 557-560. [李海超, 张广军. 一种基于角点引导的快速立体边缘匹配方法 [J]. 北京航空航天大学学报, 2007, 33 (5): 557-560.]