

一种预测质量可控的快速运动估计搜索算法

鲁照华, 李 华, 褚晶辉, 俞斯乐

(天津大学电子信息工程学院, 天津 300072)

摘 要: 本文提出了一种预测质量可控的快速运动估计搜索算法. 该算法利用运动矢量的空间、时间相关性, 通过预测初始搜索中心位置、判定是否为静止宏块、调整搜索窗口、预测质量可控的渐进式块匹配准则搜索等步骤进行运动估计, 并引入调节函数在预测质量和搜索速度之间增加了可控性. 实验结果表明本算法大大减少了计算量, 明显提高了运算速度, 且具有很好的重建视频质量, 适合于在低码率视频编码中应用.

关键词: 预测质量可控; 渐进式块匹配准则; 运动估计; 视频编码

中图分类号: TN919.81; TP391.41 **文献标识码:** A **文章编号:** 0372-2112(2004)06-0980-03

A Fast Motion Estimation Search Algorithm with Adjustable Predictive Quality

LU Zhao hua¹, LI Hua¹, CHU Jing-hui¹, YU Si-le¹

(Electronic & Information Engineering Department, Tianjin University, Tianjin 300072, China)

Abstract: This paper presents a fast motion estimation search algorithm with adjustable predictive quality. Taking advantage of the spatial and temporal correlation of motion vectors, motion estimation of the algorithm is realized by predicting initial search center location, analyzing whether a silent macroblock or not, adjusting windows size, and using progressive block matching measure of adjustable predictive quality, and at the same time, introducing adjustable function to add controllability between predictive quality and searching speed. Experimental results show that the algorithm can reduce computation amount significantly, increase computing speed greatly, provide good reconstructed video quality, and be applied for low bit rate video encoding.

Key words: adjustable predictive quality; progressive block matching measure; motion estimation; video encoding

1 引言

视频压缩国际标准中, 例如 MPEG 和 H. 26X, 采用的均是带有运动补偿的预测编码与 DCT 变换编码相结合的混合压缩编码方法, 其中运动估计的计算量是影响编码速度的关键因素. 对于设计一个高效的视频编码器而言, 用较小的计算量获得准确的运动矢量是至关重要的.

全搜索块匹配算法(FS), 由于其精度高, 硬件易实现, 是目前应用最为广泛的运动估计算法, 但它搜索参考窗中的所有参考位置, 运算量非常大, 将耗费编码器 85% 左右的计算资源. 常采用的改进型块匹配算法, 如二维对数搜索法(2DLOG)、正交搜索法(OSA)、新三步搜索法(NTSS)、四步搜索法(4SS)、环状搜索法(DS)、三步搜索法(TSS)等, 见文献[1~7], 均是利用一定的算术准则通过减少搜索位置数来节省计算资源, 虽然比起全搜索块匹配运算来大大减少了运算量, 但往往只能找到搜索窗内的局部最优值, 使估计出的运动矢量信息不够准确, 影响解码后的图像质量.

近来, 一些快速搜索算法在不减少搜索位置数的情况下, 通过宏块降采样, 如交替降采样搜索算法(ASSA)^[8], 或按照某种渐进累积的失真准则快速判断参考宏块是否为最佳匹配宏块, 如标准化局部失真搜索算法(NPDS)^[9], 来减少实际参与运算的像素点数目, 不仅提高了运算速度, 并能尽可能地找

到全局最优值, 获得接近于 FS 的图像质量. 然而, 这些算法在预测质量和搜索速度之间缺乏可控性, 且其搜索速度与全搜索算法相比提高的理论极限均为 16 倍, 有待改进.

本文提出了一种预测质量可控的快速运动估计搜索算法(AQS), 该算法尽可能多地保留搜索窗中参与匹配运算的位置数目来找到全局最优值. 实验结果表明, 本算法大大减少了计算量, 明显提高了运算速度, 并在预测质量和搜索速度之间增加了可控性.

2 误差可控的快速运动估计搜索算法

2.1 算法流程

本算法流程包括预测初始搜索中心位置、判定是否为静止宏块、调整搜索窗口、渐进式块匹配准则搜索、宏块亚采样块匹配一步搜索等步骤. 对当前宏块做运动估计时, 首先利用图像运动的帧间、帧内相关性, 找到其初始搜索中心位置, 再根据一定的准则快速判定其是否是静止宏块, 若是, 就跳过此宏块直接对当前帧中的下一宏块进行运动估计; 否则计算参考帧中的全局运动矢量来调整搜索窗口的水平宽度和垂直高度(在搜索位置数不变的情况下), 对该窗口每隔一个位置点按照 AQM 准则进行判定, 然后进行一步搜索, 如此反复, 直到找到当前帧中所有宏块的运动矢量.

2.2 初始搜索中心位置的预测

根据运动区域的特征, 特别是对同一物体, 构成它的宏块

具有高度相关的运动矢量, 因此如果利用相邻宏块的运动矢量来预测当前宏块 (i, j) 的初始运动矢量 $iv(i, j)$, 找到初始搜索中心位置 (ic_x, ic_y) , 就可以使最终得到的运动矢量更接近于全局最优值, 甚至优于用 FS 找到的运动矢量. 初始搜索中心位置的计算可由式(1)~(3)得到, 式(1)中 (i, j, t) 、 $(i-1, j-1, t)$ 、 $(i, j, t-1)$ 分别代表当前宏块、当前宏块左上角宏块、参考帧中当前宏块的对应宏块, 依次类推.

$$iv(i, j) = \begin{cases} 0.375v(i, j-1, t) + 0.375v(i-1, j, t) \\ + 0.125v(i-1, j-1, t) + 0.125v(i-1, j+1, t), & \text{当参考帧为 } I \text{ 帧时} \\ 0.5v(i, j, t-1) + 0.15v(i, j-1, t) + 0.15v(i-1, j, t) \\ + 0.1v(i-1, j-1, t) + 0.1v(i-1, j+1, t), & \text{当参考帧为 } P \text{ 帧时} \end{cases} \quad (1)$$

$$ic_x = i + iv(i, j)_x, \quad ic_y = j + iv(i, j)_y \quad (2)$$

2.3 静止宏块的判定

通常情况下, 用固定的阈值来确定静止宏块, 如 SAD=512 常用做 16*16 宏块的阈值. 本算法中提出一种与相邻宏块 SAD 值有关的自适应阈值 Th_0 , Th_0 的实现见式(3), SAD 的定义见式(8). 若当前宏块与其在参考帧中的初始对应宏块所得的 $SAD \leq Th_0$, 则其运动矢量就为 $iv(i, j)$, 跳过此宏块, 对下一宏块进行运动估计.

$$Th_0 = \begin{cases} 0.375SAD(i, j-1, t) + 0.375SAD(i-1, j, t) \\ + 0.125SAD(i-1, j-1, t) + 0.125SAD(i-1, j+1, t), & \text{当参考帧为 } I \text{ 帧时} \\ 0.5SAD(i, j, t-1) + 0.15SAD(i, j-1, t) + 0.15SAD(i-1, j, t) \\ + 0.1SAD(i-1, j-1, t) + 0.1SAD(i-1, j+1, t), & \text{当参考帧为 } P \text{ 帧时} \end{cases} \quad (3)$$

2.4 搜索窗口的调整

为了简便, 参考帧中全局运动矢量 t 为其内部所有宏块 (共有 M 个) 运动矢量 v_i 的平均值, 见式(4).

$$t = \frac{1}{M} \sum_{i=1}^M v_i \quad (4)$$

假设搜索窗大小为 $\pm w$, 则参考位置数为 $(2w+1)^2$, 引入缩放因子 a , 利用式(5)、(6)得到调整后搜索窗的水平宽度 W 和垂直高度 H . 式(5)中, t_x 、 t_y 分别为 t 的水平方向和垂直方向的大小. 需要指出的是, 当参考帧为 I 帧时, 不需要对搜索窗口进行调整.

$$(2w+1)^2 = (at_x) \times (at_y) \quad (5)$$

$$W = at_x, \quad H = at_y \quad (6)$$

2.5 预测质量可控的渐进式块匹配准则(AQM)

衡量最佳匹配的准则有很多种, 其中 SAD 运算量最小, 用的最多. SAD 定义为

$$d = SAD_{(x,y)}(u,v) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |F_i(x+i, y+j) - F_{i-1}(x+i+u, y+j+v)| \quad (7)$$

式(7)中, $F_i(x+i, y+j)$ 为第 i 帧中位于 (x, y) 位置的块中 (i, j) 处的像素值, u 、 v 分别为水平和垂直方向的偏移量.

AQM 准则把每一个 16*16 宏块分割成 16 个 4*4 的小块, 如图 1 所示, 各小块对应位置如图中标号所示. 同时, 把小

块 1 分成若干个子块, 图中标注相同值的像素点 (不需要一定相邻) 构成小块 1 的一个子块. 若小块 1 被分割成 L 个子块, 则可得到 d_1, \dots, d_L 个失真值, 下标 1, \dots, L 对应各子块中像素的标注值. 小块 1 的各子块和每一个小块的 SAD 值 (失真 d) 的计算见式(7).

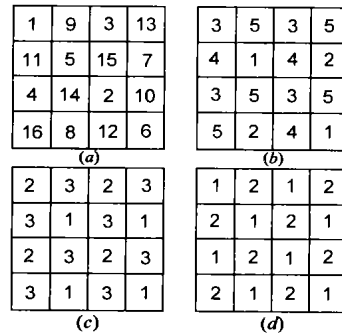


图 1 宏块分割及小块 1 的三种分割方式

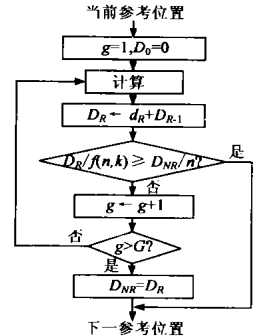


图 2 预测质量可控的渐进式块匹配准则实现流程图

AQM 匹配准则的实现过程如图 2 所示, 图中 G 是一个宏块所分的级数, 若小块 1 被分成 5 个子块, 则 $G = 5 + 16 - 1 = 20$. 假设经过若干参考位置后, 得到这些参考位置中的一个最小失真值 D_{MN} , 在当前参考位置下不妨先设 $g = 1, D_0 = 0$, 利用式(7)计算出 d_g , 则 $D_g \leftarrow d_g + D_{g-1}$, 若 $(D_g/f(n, k)) \geq (D_{MN}/N^2)$ 暂定为参考位置宏块中已参与基本运算的像素数目 n , 认为该参考位置不是所搜索过位置中的最优位置, 转入下一参考位置; 否则 $g \leftarrow g + 1$, 重复上述过程, 直到搜索完搜索窗中所有需要参考的位置, 找到使 D_{MN} 最小的位置点.

将式 $(D_g/f(n, k)) \geq (D_{MN}/N^2)$ 变形为 $D_g \geq (D_{MN}f(n, k)/N^2)$, 可以看出 $f(n, k)$ 越大, AQM 准则在执行起始时式 $D_g \geq (D_{MN}f(n, k)/N^2)$ 成立的可能性就越小, 这意味着更多的像素点有机会参与匹配运算, 从而提高预测质量, 但是也降低了搜索速度. 因此, $f(n, k)$ 是控制预测质量与搜索速度之间矛盾的一个关键因素, 称为调节函数, 其定义见式(8).

$$f(n, k) = (1-k)n + kN^2, \quad 0 \leq k \leq 1 \quad (8)$$

分析式(8), 当 n 为常量的情况下, $f(n, k)$ 是 k 的单增函数, 其最小值为 n , 最大值为 N^2 , 即式 $D_g \geq (D_{MN}f(n, k)/N^2)$ 变为 $D_g \geq D_{MN}$, 此种情况与对该参考位置进行全搜索算法所得的结果相同, 重建图像质量最好. 当要求搜索速度较快时, 我们就通过减小 k 值牺牲一定程度的预测质量来实现, 反之亦然. 一般来说, 在给定预测误差的容限后, AQM 准则在执行过程中 k 值是不变的.

通过 AQM 准则找到搜索窗口中参考位置一半点数中的最优匹配位置后, 需要对其进行一步搜索 (即窗口为 ± 1). 对该窗口中的所有宏块采用简单的宏块亚采样方式, 使每一宏块只有一半像素点参与式(7)中的 SAD 计算.

3 实验结果

实验中, 所有序列均为 CIF 格式, 长度为 100 帧, 采用 TM5 码率控制方式, 码率为 384kb/s. 表 1、表 2 中, AQS(a)、AQS(b)、AQS(c) 分别表示采用图 4 中方式 (a)、(b)、(c) 分割小块

1 进行渐进式块匹配的 AQS 算法, 并假设 $k=0$ 。从表 1 可以看出, AQS 不仅对运动缓慢的序列具有很高的 PSNR, 同时对运动剧烈的序列也具有很高的 PSNR, 接近于 FS 的 PSNR。表 2 中, 比较了 AQS 与其它算法的速度, AQS 的速度明显高于其它快速算法。表 1、表 2 中, 与 AQS(a)、AQS(b) 相比, AQS(c) 的 PSNR 最高, 速度最慢, 原因在于其参与匹配运算的像素点数较多。表 3 中, 列出了在不同 k 值下 AQS(a) 的 PSNR 值, k 越大, PSNR 就越大, 意味着重建图像的质量就越好, 但搜索速度下降, 实现了预测质量和搜索速度之间的可控性, 方式(b)、方式(c)在不同 k 值下的 PSNR 和速度的变化规律类似于方式(a)。图 3 中, 采用 AQS 得到的 Mobile&Calendar 序列获得接近于 FS 的重建图像质量, 好于 TSS 得到的结果。图 4 给出了使用三种算法预测 Foreman 序列图像的部分动态矢量图(主要集中于面部的转动, 运动矢量比较复杂), 可以发现, AQS(c) 与 FS 产生的运动矢量图很接近, 优于 TSS(特别是大运动矢量的情况下)。

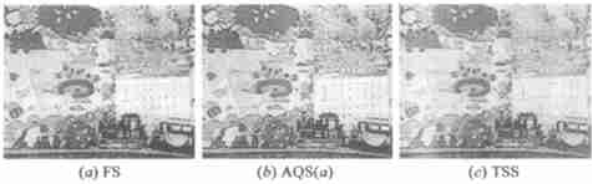


图 3 序列 Mobile&Calendar 在各种算法下的重建图像

表 1 不同序列使用各种算法所得 Y 信号的平均 PSNR (单位: dB)

序列 \ 算法	FS	TSS	NTSS	DS	4SS	AQS (a)	AQS (b)	AQS (c)
Akiyo	44.25	44.22	44.25	44.25	44.25	44.25	44.25	44.25
Coastguard	33.67	33.59	33.57	33.60	33.29	33.60	33.63	33.67
Foreman	31.83	29.72	29.65	29.75	29.13	30.04	30.30	31.57
Mobile&Calendar	29.77	28.07	28.21	28.54	28.26	28.95	29.21	29.53
Basketball	30.24	28.70	28.33	28.68	28.30	29.41	29.64	29.97

表 2 同序列使用各种算法的搜索速度比较
(与 FS 搜索速度相比的比值)

序列 \ 算法	FS	TSS	NTSS	DS	4SS	AQS (a)	AQS (b)	AQS (c)
Akiyo	1	11.96	16.15	11.53	16.07	43.21	23.56	12.51
Coastguard	1	12.45	12.37	11.96	12.08	48.57	28.92	16.53
Foreman	1	12.11	11.98	11.31	11.64	45.29	31.57	22.41
Mobile&Calendar	1	8.72	9.73	12.41	10.85	47.72	36.13	23.56
Basketball	1	8.75	11.40	14.78	11.43	53.73	40.88	26.42

表 3 不同序列采用 AQS 方式(a)时在不同 k 值下所得 Y 信号的 PSNR (单位: dB) 及搜索速度比较(与 FS 搜索速度相比的比值)

序列 \ 算法	AQS(0.00)		AQS(0.05)		AQS(0.10)		AQS(0.15)	
	PSNR	速度	PSNR	速度	PSNR	速度	PSNR	速度
Akiyo	44.25	43.21	44.25	26.53	44.25	20.14	44.25	16.51
Coastguard	33.60	48.57	33.67	30.16	33.67	19.03	33.67	16.47
Foreman	30.04	45.29	30.27	36.18	31.21	27.51	31.81	22.63
Mobile&Calendar	28.95	47.72	29.34	35.18	29.51	27.09	29.76	19.72
Basketball	29.41	53.73	29.75	38.70	30.08	27.15	30.24	20.89

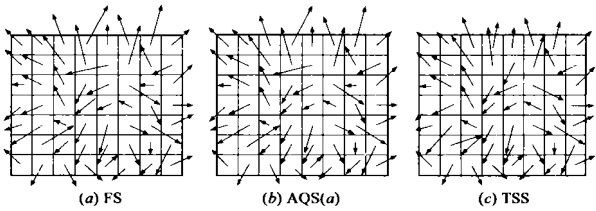


图 4 Foreman 序列在各种算法下的部分运动矢量图

4 结论

本文提出一种预测质量可控的快速运动估计搜索算法, 引入调节函数在预测质量和搜索速度之间增加了可控性。实验结果表明该算法大大减少了计算量, 明显提高了运算速度, 具有很好的重建图像质量, 很适合于在低码率情况下进行优质快速的视频编码。

参考文献:

[1] JAIN J R, JAIN A K. Displacement measurement and its application in interframe image coding [J]. IEEE Trans Commun, 1981, COM-29 (12): 1799- 1808.

[2] FURIA, et al. An efficient block matching algorithm for motion compensated coding [J]. IEEE ICASSP PROC, 1987, 12(1): 1063- 1066.

[3] LI R, et al. A new three-step search algorithm for block motion estimation [J]. IEEE Trans Circuits Syst Video Technol, 1994, 4(4): 438- 443.

[4] PO L M, MA W C. A novel four-step search algorithm for fast block motion estimation [J]. IEEE Trans Circuits Syst Video Technol, 1996, 6(3): 313- 317.

[5] THAM J Y, et al. A novel unrestricted center biased diamond search algorithm for block motion estimation [J]. IEEE Trans Circuits Syst Video Technol, 1998, 8(4): 369- 377.

[6] KOGA T, et al. Motion compensated interframe coding for video conferencing [J]. Telecommunications Proc Nat, 1981, 1(1): 531- 535.

[7] 李华, 俞斯乐, 等. 一种用于运动估计的预测搜索算法 [J]. 信号处理, 1996, 12(1): 46- 49.

[8] LIU B, ZACCARIN A. New fast algorithm for the estimation of block motion vectors [J]. IEEE Trans Circuits Syst Video Technol, 1993, 3(2): 148- 157.

[9] CHEUNG C K, PO L M. Normalized partial distortion search algorithm for block motion estimation [J]. IEEE Trans Circuits Syst Video Technol, 2000, 10(3): 417- 422.

作者简介:



鲁照华 男, 1979 年 11 月出生于内蒙古自治区包头市, 博士, 主要研究方向为多媒体通信。

李 华 女, 1942 年 5 月出生于天津市, 现为天津大学教授、博士生导师, 主要研究方向为数字高清晰度电视、数字视音频测量技术、数字图像处理。