

一种 H. 264 自适应块搜索范围快速运动估计算法

郑庆辉¹⁾ 何加铭^{1),2)} 范志蔚¹⁾ 王军明¹⁾

¹⁾(宁波大学通信技术研究, 宁波 315211) ²⁾(宁波大学省属科技创新基地, 宁波 315211)

摘要 块匹配运动估计在视频编码中有着举足轻重的地位,其性能的优劣在很大程度上影响着输出码流的质量。全搜索是效果最好的运动估计算法,但其巨大的运算量是实际应用无法承受的。为解决这一问题,该文提出了一种基于 H. 264 自适应块搜索范围快速运动估计算法。该算法是利用相邻宏块的运动矢量来确定初始搜索中心位置,并动态地确定搜索区域的大小。与以往固定搜索区域的算法相比,它具有更高的灵活性,还可以与其他任何一种快速运动估计算法相结合。仿真结果表明,在基本不损失性噪比的前提下,根据不同的运动序列,该算法可不同程度地减少运算的复杂度。

关键词 自适应 运动估计 搜索区域 峰值性噪比

中图分类号: TN919.81 **文献标识码**: A **文章编号**: 1006-8961(2005)11-1350-05

A Fast Motion-estimation Algorithm Based on Adaptive Block Search Range for H. 264

ZHENG Qing-hui¹⁾, HE Jia-ming^{1),2)}, FAN Zhi-wei¹⁾, WANG Jun-ming¹⁾

¹⁾(Communication Technology Institute, Ningbo University, Ningbo 315211)

²⁾(Innovation-base, Ningbo University, Ningbo 315211)

Abstract Block matching motion estimation (BMME) plays a very important role in video coding. The performance of BMME greatly affects the quality of the encoded bit stream. Full Search (FS) is the optimal algorithm for BMME. But its tremendous computation payload is unaffordable for practical applications, especially for real-time applications. To overcome this problem, a lot of fast algorithms were developed. This paper presents an adaptive fast motion-estimation algorithm based on H. 264. Taking advantage of motion-vectors of the nearby macro-blocks to find the original searching central position and determine the size of the searching zone. Compared with the former algorithms with fixed searching zone, this algorithm is more flexible and can combine with any other fast motion estimation algorithm. Experimental results show that the algorithm can reduce computation amount variously by different motion-sequence under the premise don't reduce the PNSR.

Keywords self-adaptive, motion-estimation, searching zone, PNSR

1 引言

H. 264 是由 ITU-T 和 MPEG 组成联合专家组 JVT 制定的目前最新的视频编码标准,这一编码标准可以获得很高的编码效率,尤其在低码率视频编码方面比 MPEG4 有明显提高,非常适合于低带宽

质量网络视频的应用需要,但是,由于 H. 264 为了提高编码效率,采用了许多高计算复杂度的算法,使得编解码计算量很大,因而很难用软件方式进行实时编解码,同时为硬件实现编解码也增加了难度。

为了使 H. 264 能比较容易地在低码率、高实时性的应用系统中实现,必须对其编码算法进行优化。通过对 H. 264 编码器各个算法模块进行的分析可

基金项目: 国家科技部科技技术创新基金项目 (04C26213310141)

收稿日期: 2005-09-05; 改回日期: 2005-09-29

第一作者简介: 郑庆辉 (1981 ~), 男。2002 年获宁波大学自动化专业学士学位, 现于宁波大学攻读无线通信方向硕士学位。研究方向为基于 ARM 的 MPEG-4 动态图像压缩及无线传输关键技术。E-mail: zqh0899tq@163.com

以看出,由于运动估计模块的计算量占整个编码器运算量的 80% 以上,因此,若对整个编码器进行优化,运动估计模块应作为首选。

运动估计算法的好坏影响着编码比特率、重建图像质量和速度。如果以预测误差作为衡量标准,则全搜索(full search, FS)是最好、最简单的匹配算法,但由于它的计算量太高,不适合实时应用,因而提出了许多改进的快速算法如三步法(three-step search, 简称 3SS)^[1]、新三步法(new three-step search, 简称 N3SS)^[2]、四步法(four-step search, 简称 4SS)^[4]和 MPEG4 校验模型中的菱形法(diamond search, 简称 DS)^[4]等。由于它们有一个共同的特点就是它们的搜索范围都是确定的,从而必然带来一定的冗余。而现在一些论文中提出的自适应搜索范围算法(extended adaptive range search, 简称 EARS)^[5]则是先求取整帧图像所有块的运动向量的平均值,然后通过这一平均值来确定整帧图像的块搜索范围。但这种算法对于那些运动向量偏差较小的块也带来了搜索冗余,而对于那些运动向量偏差较大的块,则又容易造成图像质量损失。

本文提出了一种适用于 H. 264 的自适应块搜索范围快速运动估计算法(new adaptive range search, 简称 NARS)。该算法由于利用了图像运动的帧间、帧内相关性来找到其初始搜索中心位置^[6],并动态地确定每个块的搜索区域的大小,从而其可以根据帧率和运动幅度的不同,不同程度地减少运算复杂度。同时这种算法还可以与其它任何一种快速运动估计算法相结合,并能够在基本不损失信噪比的前提下,提高这种搜索算法的运算速度。

2 自适应块搜索范围快速运动估计算法(NARS)

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、动态调整搜索区域的大小以及根据设定的阈值提前结束搜索等步骤。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,然后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,就根据相邻宏块的运动矢量来动态地确定当前块的搜索区域的大小,以便减少搜索点数和减少运

算的复杂度。算法流程如图 1 所示。

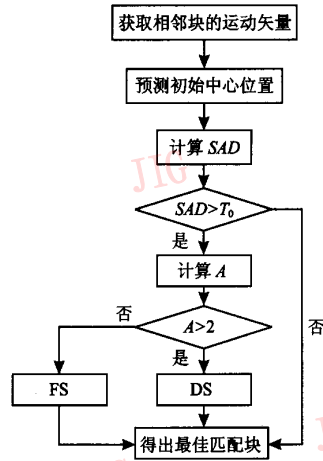


图 1 NARS 算法的实现流程图

Fig. 1 Flow chart of the NARS method

2.2 衡量最佳匹配的准则

衡量最佳匹配的准则有很多种^[7],其中绝对差和(sum of absolute differences, 简称 SAD)运算量最小,用得也最多,本文采用 SAD 来进行最佳匹配的衡量。SAD 定义为

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

2.3 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,由于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 $V_{i,j}$,那么找到初始搜索中心位置 (x_c, y_c) 后,就可以使最终得到的运动矢量更接近于全局最优值^[8],甚至优于用 FS 搜索算法找到的运动矢量。初始搜索中心位置的计算可由式(2),式(3)得到,式(2)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。

$$v_{i,j} = \begin{cases} \text{当参考帧为 I 帧时,} \\ 0.375v_{i,j-1,t} + 0.375v_{i-1,j,t} + \\ 0.125v_{i-1,j-1,t} + 0.125v_{i-1,j+1,t}; \\ \text{当参考帧为 P 帧时,} \\ 0.5v_{i,j,t-1} + 0.15v_{i,j-1,t} + 0.15v_{i-1,j,t} + \\ 0.1v_{i-1,j-1,t} + 0.1v_{i-1,j+1,t}; \end{cases} \quad (2)$$

$$x_c = i + v_x(i, j), y_c = j + v_y(i, j) \quad (3)$$

2.4 静止宏块的判定

通常情况下,都是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 T_0 。 T_0 的确定见式 (4),SAD 的定义见式 (1)。若当前宏块与其在参考帧中的初始对应宏块所得的 $SAD \leq T_0$,则其初始运动矢量就为 $v(i, j)$,此时就可跳过此宏块直接对下一宏块进行运动估计。

$$T_0 = \begin{cases} \text{当参考帧为 I 帧时,} \\ 0.375SAD_{i,j-1,t} + 0.375SAD_{i-1,j,t} + \\ 0.125SAD_{i-1,j-1,t} + 0.125SAD_{i-1,j+1,t}; \\ \text{当参考帧为 P 帧时,} \\ 0.5SAD_{i,j,t-1} + 0.15SAD_{i,j-1,t} + 0.15SAD_{i-1,j,t} + \\ 0.1SAD_{i-1,j-1,t} + 0.1SAD_{i-1,j+1,t}. \end{cases} \quad (4)$$

2.5 搜索区域的自适应调整

尽管减少运动估计算法复杂度的方法很多,但本文重点把研究重心放在搜索区域的自适应调整上。由 2.3 节即可得出初始运动中心位置的预测,同时可以知道相邻宏块的运动矢量。根据 SAD 分布的空间方向性和运动矢量中心偏移性的思想^[9],若当相邻块的运动矢量绝对值较大时,预测出来的当前块的运动矢量与最后计算得出的运动矢量之间的偏差的收敛域会比较大,这时就需要放大搜索范围才能找到最佳匹配块;当相邻块的运动矢量都比较小时,则情况恰恰相反。另外,根据搜索区域的不同,还可以与具体的快速运动估计算法相结合,以便在保证图像质量的前提下,达到减少运算复杂度的目的。

NARS 的实现过程为:当初始搜索中心位置测定后,即可根据相邻块运动向量的绝对值的平均值来动态确定搜索区域的大小 A 。具体计算公式如下:

$$A = \lfloor 1.5 \times D \rfloor \quad (5)$$

$$\text{其中, } D = \frac{\sum_{i=0}^{N_{MV}} |V_i^{\text{motion}}|}{N_{MV}} \quad (6)$$

V_i^{motion} 为第 i 帧运动矢量。

分析以上实现过程可知当帧率较高,相应的运动幅度较小时,由于一帧图像中绝大部分宏块的运动矢量都比较小,所以采用本文的算法后,绝大部分宏块的搜索区域可以限定在较小的范围内 ($A \leq 2$),本文对这些宏块采用 FS 搜索算法,虽大大地减少了匹配的点数,降低了计算复杂度,但并不影响重建图像的质量;当帧率相对比较低运动幅度较大时,一小

部分宏块的搜索区域会比较大 ($A \geq 3$),本算法对这部分宏块采用 DS 搜索算法,由于一帧图像中大部分宏块的运动矢量还是比较小的,因而匹配的点数也会大大地减少,但没有帧率高、运动幅度小的情况下减少得多,图像质量也有一定的损失,也就是在性能的提高上有所下降。

3 仿真实验结果

实验时,在 JVT 的 H.264/MPEG 4-AVC 编码参考模型 JM5.0 上实现本文所提算法。测试环境是根据 JVT 制定的编码器测试环境扩展而来。为验证本文算法效果,选择具有广泛代表性的 7 个标准测试序列进行了仿真测试实验,这些序列为: Saleman、Carphone、Claire、Foreman、Grandmother、Flower-garden、Table-tennis,其中, Table-tennis 为大运动序列, Carphone、Foreman、Salesman 为中等运动序列, Claire、Grandmother 为小运动或几乎静止序列, Flower-garden 则包含了较多的细节与镜头的平移。采用的帧率是 30Hz,量化步长为 28,搜索范围是 ± 32 点,单帧参考。编码后的序列除首帧外,其余各帧均为 P 帧。为了便于比较,还在 JM5.0 上实现了全搜索法、三步法、新三步法、四步法和菱形法 5 种经典算法。

测试结果如表 1、表 2,图 2 ~ 图 5 所示。

表 1 测试序列不同算法重建图像的 PSNR 比较

Tab.1 Average PSNR performance of the discussed methods

测试序列	算 法					
	FS	3SS	N3SS	4SS	DS	NARS
Carphone	35.39	35.11	35.26	35.19	35.32	35.38
Claire	39.5	39.42	39.4	39.4	39.46	39.49
Flower-garden	35.46	35.13	35.43	35.33	35.45	35.46
Foreman	35.58	35.43	35.58	35.45	35.52	35.55
Grandmother	37.2	37.15	37.2	37.2	37.19	37.2
Salesman	36.24	36.15	36.21	36.19	36.22	36.24
Table-tennis	35.78	35.41	35.67	35.63	35.66	35.7

表 2 测试序列不同算法平均搜索点数比较

Tab.2 Average searching points of the discussed methods

测试序列	算 法					
	FS	3SS	N3SS	4SS	DS	NARS
Carphone	4225	41	19.5	18.27	17.86	17.23
Claire	4225	41	17.56	17.24	15.68	10.38
Flower-garden	4225	41	19.16	17.96	16.74	17.58
Foreman	4225	41	19.36	18.21	16.83	15.26
Grandmother	4225	41	17.46	17.16	15.46	10.86
Salesman	4225	41	17.89	17.48	16.24	10.52
Table-tennis	4225	41	19.29	18.16	18.05	17.36

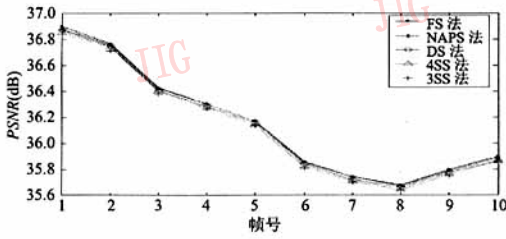


图 2 Foreman 序列几种不同算法所得重建图像的 PSNR 的比较(10 帧)

Fig. 2 Comparison of PSNR for "Forman" (10 frames)

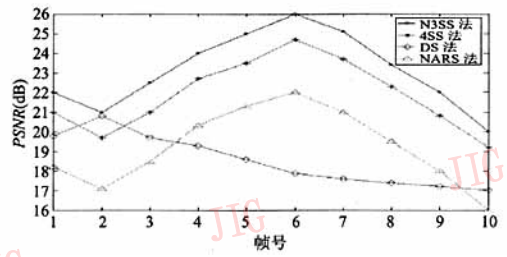


图 3 Foreman 序列几种不同算法比较次数的比较(10 帧)

Fig. 3 Comparison of operation counts for "Forman" (10 frames)

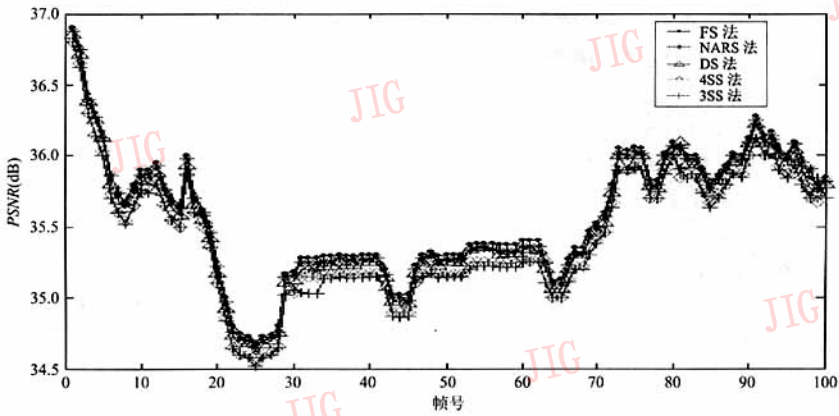


图 4 Foremanmt 几种不同算法所得重建图像的 PSNR 的比较(100 帧)

Fig. 4 Comparison of PSNR for "Forman" (100 frames)

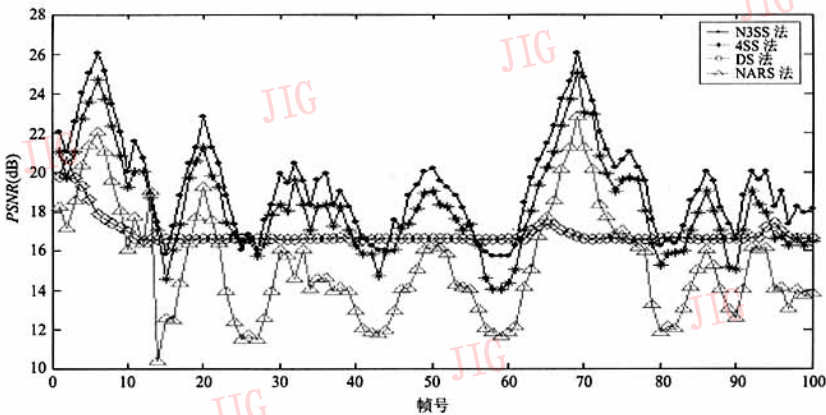


图 5 Foreman 序列几种不同算法比较次数的比较(100 帧)

Fig. 5 Comparison of operation counts for "Forman" (100 frames)

分析表 1 可以得出, NARS 算法所得图像的 PNSR 与 FS 算法最接近。同时, 当测试序列的运动越小时, 利用 NARS 算法损失的信噪比越少; 当测试序列的运动比较大时, 则用该算法所得图像的信噪比损失会有小幅度增大, 其主要原因是由于当运动向量比较大时 ($A \geq 3$), 由于本文采用 DS 算法, 从而引入了一点的质量损失。

分析表 2 可得, 在平均搜索点数的比较上, NARS 算法所需要的比较次数最少, 因为当前块的运动矢量大部分都是靠近预测出来的初始搜索中心位置的缘故。当搜索范围等于 1 时, NARS 算法最多只需要 9 个搜索点; 而当搜索范围等于 2 时, 最多只需要 25 个搜索点。由于引入静止块的方法以后, 搜索点数还会继续减少, 从而使最终的搜索点数明显小于其他各类搜索算法。

通过对图 2 ~ 图 5 的分析, 由于可以更加清楚地看到 Foreman 测试序列每帧图像的信噪比和平均搜索点数的情况, 从而可更加直观地看到 NARS 算法的优越性。

从上面的实验结果可以看出, 不论是对小的运动序列, 还是对大的运动序列, 本文给出的算法在比较次数、匹配精度和重建图像质量方面都优于前面提出的各类快速算法。本文算法在匹配精度上虽与全搜索相当, 但计算量却有很大的降低。它可根据相邻块的运动相关性来预测出当前块的运动向量, 并可通过一定的计算公式来动态地确定块搜索范围的大小。同时, 由于它还能根据具体情况与各类快速算法相结合, 因此适用面非常广。

4 结 论

本文提出了一种新的适用于 H. 264 自适应块搜索范围的快速运动估计搜索算法, 并引入了根据相邻块运动矢量的绝对值来动态地确定每个块的搜索区域大小的概念。实验结果表明, 在基本不影响图像重建质量的前提下, 该算法大大地减少了计算

量, 不仅明显提高了运算速度, 而且具有适用面广、鲁棒性强的特点。

参考文献 (References)

- 1 Iinuma K, Hirano A, Baik H, et al. Motion compensated inter-frame coding for video conferencing [A]. In: National Telecommunications Conference [C], New Orleans, LA, USA, 1981.
- 2 LI Ren-xiang, ZENG Bing, LIU Ming. A new three step search algorithm for block motion estimation [J]. IEEE Transactions on Circuits System Video Technology, 1994, 4(4): 438 ~ 442.
- 3 PO Lai-man, MA Wing-chung. A novel four-step search algorithm for fast block motion estimation [J]. IEEE Transactions on Circuits System Video Technology, 1996, 6(3): 313 ~ 317.
- 4 ZHU Shan, MA Kai-kuang. A new diamond search algorithm for fast block matching motion estimation [J]. IEEE Transactions on Image Processing, 2000, 9(2): 287 ~ 290.
- 5 LI Xiang, WU Guo-wei. A fast motion estimation algorithm based on adaptive search range for H. 264 [J]. Journal of Image and Graphics, 2004, 9(4): 471 ~ 476. [李翔, 吴国威. 一种适用于 H. 264 的基于自适应搜索范围的快速运动估计算法 [J]. 中国图象图形学报, 2004, 9(4): 471 ~ 476.]
- 6 LI Wei, LE Li-luan, LI Bo. A fast motion estimation algorithm based on predicted start [J]. Chinese Journal of Computers, 2001, 24(10): 1110 ~ 1114. [李伟, 乐立鸾, 李波. 基于起点预测和 SAD 分布的快速运动估计算法 [J]. 计算机学报, 2001, 24(10): 1110 ~ 1114.]
- 7 DING Wen-guang, JI Wen-ping, GUO Bao-long. Visual C++ 6.0 for digital image coding [M]. China Machine Press, 2004. [丁文广, 计文平, 郭宝龙等. Visual C++ 6.0 数字图像编码 [M]. 机械工业出版社, 2004.]
- 8 LU Zhao-hua, LI Hua, CHU Jing-hui. A fast motion Estimation search algorithm with adjustable predictive quality [J]. Acta Electronica Sinica, 2004, 32(6): 980 ~ 982. [鲁照华, 李华, 褚晶辉. 一种预测质量可控的快速运动估计搜索算法 [J]. 电子学报, 2004, 32(6): 980 ~ 982.]
- 9 XUE Jing-zhu, SHEN Lan-sun. An efficient block matching motion estimation algorithm for H. 264/AVC [J]. Acta Electronica Sinica, 2004, 32(4): 583 ~ 586. [薛金柱, 沈兰荪. 一种基于 H. 264/AVC 的高效块匹配搜索算法 [J]. 电子学报, 2004, 32(4): 583 ~ 586.]