

文章编号:1001-9081(2007)09-2150-02

基于 UMHexagonS 的快速帧间模式选择算法

白茂生,田裕鹏,田晓冬

(南京航空航天大学自动化学院,南京 210016)

(xiaoyaobms@163.com)

摘要:对基于“非对称十字型多层次六边形格点搜索”(UMHexagonS)的帧间模式选择算法进行了改进,应用了优先判断 SKIP 模式和基于图像细节的分块模式选择算法。实验结果表明,该方法在保证图像失真度和码率性能的前提下,单帧编码时间较 UMHexagonS 减少了 57.48%,较大幅度增加了帧间模式的决策速度。

关键词:H.264/AVC;帧间预测;UMHexagonS;视频编码

中图分类号:TP391.41 **文献标志码:**A

Fast inter-mode decision algorithm based on UMHexagonS

BAI Mao-sheng, TIAN Yu-peng, TIAN Xiao-dong

(College of Automation Engineering,

Nanjing University of Aeronautics and Astronautics, Nanjing Jiangsu 210016, China)

Abstract: Fast inter-mode decision algorithm based on "Unsymmetrical cross Hexagon grid Search" (UMHexagonS) was improved in this paper. The priority estimation of SKIP mode and the inter-mode decision based on image detail were employed. Experimental results show that the one-frame scramble time of fast inter-mode decision, compared with UMHexagonS algorithm, is reduced by 57.48% with negligible loss of peak signal-to-noise ratio.

Key words: H.264/AVC; interprediction; UMHexagonS; video coding

0 引言

H.264/AVC^[1]是由 ISO/IEC MPEG 和 ITU-T VCEG 联合制定的最新的视频编码标准。在相同的条件下,与 H.263+ 和 MPEG-4 相比,H.264/AVC 能减少 50% 的码流。然而,H.264 编码效率的提升是以计算复杂度的增加为代价的。与 MPEG-4 简单配置相比,H.264 主要配置下解码复杂度提高了 3 倍以上,而编码复杂度更是增加了 10 倍。

1 H.264/AVC 帧间编码

帧间预测是从已被编码的重建的一帧或多帧中产生预测模板。这个模板是通过在参考帧中进行采样点的平移得到的。在 H.264 中仍然采用了从 H.261 起就被广泛运用的基于块的运动补偿。和早期的标准不同的是 H.264 支持更多更细的块大小划分以及亚像素运动矢量估计。

尽管 H.264/AVC 标准中并未指定运动估计算法,但由于运动估算的速度和准确性对整个编码器性能的影响较大,因此 H.264 一直在寻找一种更优秀的快速运动估计算法。目前 H.264 已正式采纳了“非对称十字型多层次六边形格点搜索”(UMHexagonS)^[2]算法,它的运算量相对于快速全搜索算法可节约 90% 以上,同时能保持较好的率失真性能。

H.264/AVC 支持 16×16 、 16×8 、 8×16 、 8×8 、 8×4 、 4×8 、 4×4 的 7 种分块大小的运动补偿^[3]。为了提高压缩效率,H.264/AVC 允许 B 帧和 P 帧中的宏块按照 intra 模式进行编码。所以尽管 P 帧和 B 帧采用帧间编码,但是 H.264 编码器还是会检查 intra 4×4 、intra 16×16 中的所有模式。另外

加上 SKIP 模式,故帧间编码时须对 10 种模式进行比较,来获得最优的编码性能。H.264/AVC 采用了率失真优化 (Rate Distortion Optimization, RDO)^[4] 技术,比较各种模式的性能。然而这增加了计算的复杂度。如果事先对待测模式进行筛选,获得一个相对较小的待测模式集,便可有效提高编码速度。因此研究 RDO 模式下的快速帧间模式选择算法具有重要的应用价值。

本文在 UMHexagonS 算法的基础上,对帧间编码模式进行了改进。实验结果表明,该方法在保证图像失真度和码率性能的前提下,较大幅度地加快了帧间模式的决策速度。

2 UMHexagonS 算法

在已有的块匹配算法当中,全搜索 (FS) 精度是最高的,但是其计算时间长,运算量大,很难满足实时性要求。UMHexagonS 以其高速和高性能被 H.264 正式采用。其基本步骤如下:

步骤 1:起始点预测

有 5 种运动向量预测模式:

1) 中值预测:利用空间相关性,取已求出的当前块的左、上、右上邻块的运动矢量的中间值 $MV_{pred_MP} = Median(MV_A, MV_B, MV_C)$;

2) 上层预测:按照 16×16 (模式 1) 到 4×4 (模式 7) 的顺序,将已求出的当前位置的上一级模式运动向量作为当前模式的运动向量的预测值 $MV_{pred_up} = MV_{UpLayer}$;

3) 对应块预测:将前一帧相应位置的运动向量作为当前块的运动向量的预测值 $MV_{pred_CP} = MV_{CB}$;

收稿日期:2007-03-20;修回日期:2007-06-08。

作者简介:白茂生(1983-),男,甘肃金昌人,硕士研究生,主要研究方向:视频编解码、DSP 应用;田裕鹏(1964-),男,山东潍坊人,副教授,硕士,主要研究方向:视频编解码、图像处理;田晓冬(1978-),男,河南漯河人,硕士研究生,主要研究方向:视频编解码。

4) 相邻参考帧预测:利用时间相关性,将前面参考帧的运动向量按时间进行缩放:

$$MV_{pred_NRP} = MV_{NR} \times \frac{t - t' + 1}{t - t'}$$

其中 MV_{NR} 为在时间 t' 参考帧 N 在当前位置的运动向量;

5) 原点预测.考虑到在固定背景上的物体边界的情况:

$$MV_{pred_O} = MV_{(0,0)}$$

对以上的运动向量进行菱形搜索,来获得最佳的起始搜索点。

类似 MV ,相邻块的 SAD 在时间和空间上也有很强的相关性,因此对 SAD 也进行预测,利用阈值公式来判断此处不甚满意、满意还是很满意区来跳转到步骤2或步骤3或步骤4。

步骤2:不甚满意区的块搜索

1) 以最佳点起始搜索点为中心,用非对称十字型搜索模板进行搜索;获得当前最佳点,判断此处是否属于满意或很满意区,跳到相应的步骤3或步骤4;或继续搜索。

2) 以当前最佳点为中心,在 $(-2,2)$ 的方形区域中进行全搜索;获得当前最佳点,判断此处是否属于满意或很满意区,跳到相应的步骤3或步骤4;或继续搜索。

3) 用不断扩大一倍直径的大六边形模板进行搜索,直至搜索到能符合相应阈值而进入步骤3或步骤4的点为止;否则也结束步骤2的搜索进入步骤3。

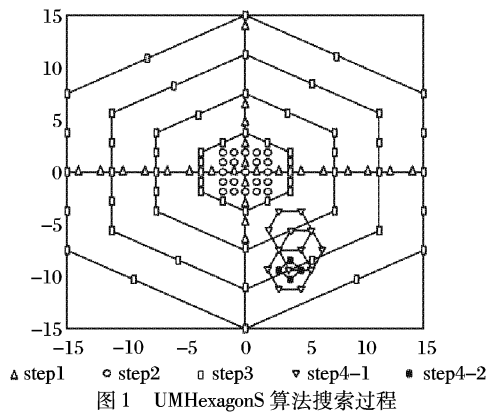
步骤3:满意区的块搜索

以当前最佳点为中心进行六边形搜索,直至最佳点为六边形的中心。

步骤4:很满意区的搜索

以当前最佳点为中心,进行菱形搜索,直至最佳点为菱形的中心。

整个搜索过程如图1所示。



3 改进的快速帧间模式判别算法

3.1 SKIP模式预先判别

在通常的视频序列中,相邻图像之间普遍存在着时间相关性,即相邻图像的大部分区域的变化是缓慢的。在这部分变化缓慢的区域中,SKIP模式占有相当大的比重。因此,在对所有其他模式进行遍历,以找出最佳模式前,对SKIP模式进行提前预判可以极大地减少为其他模式进行运动搜索以及计算率失真代价所消耗的开销。

SKIP模式是直接从前一参考宏块复制,代替当前帧的宏块,一般在宏块无运动向量或宏块做整体运动即只有一个运动向量时采用SKIP模式编码,因而计算简单,速度快。

H.264/AVC中,因为图像的纹理不同,故采用不同的分块模式。纹理特征可以用宏块间的残差表示^[5]。对运动补

偿后,残差小的采用大的分块模式,残差大的采用小的分块模式。纹理特征可以从帧间宏块间的 SAD 获得。

$$SAD = \frac{1}{256} \sum_{i=1}^{16} \sum_{j=1}^{16} |x_{i,j} - y_{i,j}| \quad (1)$$

其中 $x_{i,j}, y_{i,j}$ 表示当前帧和前一帧 (i,j) 处的像素灰度值。另外定义一阈值 T_0 ,判断是否采用SKIP模式,而跳过其他模式。

$$T_0 = \frac{1}{\omega} \times \frac{1}{M \times N} \sum_{i=1}^{16} \sum_{j=1}^{16} |x_{i,j} - y_{i,j}| \quad (2)$$

其中 M, N 分别为一帧图像的高和宽, ω 为校正系数,对 QP 为 28,32,36,40 下,经实验统计得到经验值分别为 5.0,4.8,4.6,4.4。

当计算一个宏块的 $SAD < T_0$ 时,采用SKIP模式,而跳过其他模式的搜索和率失真优化。

3.2 改进的块类型选择

众所周知,对于图像中平滑的区域,其能量主要集中在低频分量上,所以高频分量可以用来衡量图像的细节程度^[6]。

对于一个 $M \times N$ 块,AC系数总能量可以用式(3)表示:

$$E_{AC} = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (X_{i,j})^2 - (X_{0,0})^2 \quad (3)$$

其中 $X_{i,j}$ 代表 $M \times N$ 块各像素的DCT系数, $X_{0,0}$ 代表该 $M \times N$ 块的DC系数。

当一个 $M \times N$ 块像素值 $Y_{i,j}$,其上下左右的像素值与其相差为 255(灰度值用 8 位表示,最大值为 255,最小为 0),表明此时块细节最丰富。根据能量守恒定律,变换后的总能量等于变换前的总能量,所以对于大小为 $M \times N$ 的块,其AC系数的最大能量为:

$$E_{ACmax} = \frac{M \times N}{2} [(X_{max})^2 + (X_{min})^2] - (M \times N) \left(\frac{X_{max} + X_{min}}{2} \right)^2 \quad (4)$$

其中 $X_{max} = 255, X_{min} = 0$ 。

把 E_{ACmax} 作为一个标准化因子,对 E_{AC} 和 E_{ACmax} 取自然对数,得到:

$$T_1 = \frac{\ln E_{AC}}{\ln E_{ACmax}} \quad (5)$$

此时 T_1 被限制在 $[0,1]$ 之间,确定出一个阈值 R ,取 $R = 0.78$,如果 $R > T_1$,表示为高细节块,反之表示为低细节块。

选 8×8 作为一个判断的基准,如果低细节块大于等于 3,此时选 $Mode1$ 作为候选模式,否则选择 $Mode2$ 作为候选模式。

$Mode1 = \{SKIP, 16 \times 16, 16 \times 8, 8 \times 16\}$;

$Mode2 = \{SKIP, 16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8, 8 \times 4, 4 \times 8, 4 \times 4, intra16 \times 16, intra4 \times 4\}$;

算法步骤如下:

步骤1:计算一个 16×16 块的 SAD ,判断是否满足SKIP模式,满足终止其他模式搜索与率失真优化。

步骤2:不满足SKIP模式,计算宏块中 4 个 8×8 的AC系数,判断是否低细节块大于等于 3 个,跳转到步骤3;否则,跳转到步骤4。

步骤3:选择 $Mode1$,终止其他分块模式搜索和率失真优化。

步骤4:选择 $Mode2$,得到最佳分块模式。

(下转第 2156 页)

表 1 本文算法与 BeamLab 所用算法运行时间对比

图像大小	BeamLab	等效二进楔形	有 1 024 个角度的二进楔形	有 180 个角度的二进楔形	有 4 个角度的二进楔形	二进方块
16 × 16	6.72s	< 0.02s	1s	0.2s	< 0.01s	< 0.01s
32 × 32	45.34s	0.05s	2s	0.3s	< 0.01s	< 0.01s
64 × 64	330.41s ≈ 5.5min	0.2s	8s	1.6s	0.04s	< 0.01s
128 × 128	2 676.3s ≈ 4.6min	1.2s	37s	6s	0.2s	0.04s
256 × 256	27 918s ≈ 7h45min	8s	140s	24s	1s	0.5s
512 × 512	> 12h	59s	722s	120s	10s	6s

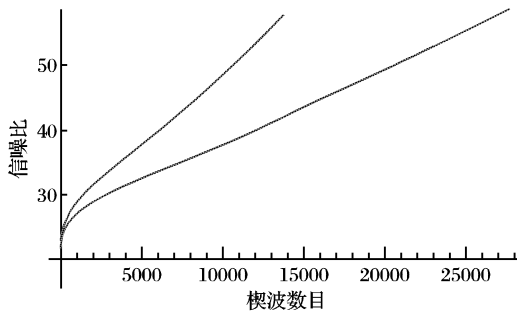


图 4 信噪比与楔波数目的关系 (Barbara)

本文提出了一种多边形区域上矩计算和楔波近似的高效快速算法。该算法能够计算建立在任意自适应局部退化函数基上的对比实验,是一种高效灵活的工具。其中最有用的特性是对于不同的 λ 能够快速计算出结果;在比较适中的时间内计算出率失真曲线;还能够针对不同的参数模型进行仿真实验。

参考文献:

- [1] WEDGELETS D L. Nearly minimax estimation of edges [J]. *Annals of Statistics*, 1999, 27 (3): 859 - 897.
- [2] WILLETT R M, NOWAK R D. Platelets: A multiscale approach for recovering edges and surfaces in photon-limited medical imaging [J]. *IEEE Transactions on Medical Imaging*, 2003, 22(4): 332 - 350.
- [3] YANG L, ALBREGTSEN F. Fast and exact computation of Cartesian geometric moments using discrete Green's theorem [J]. *Pattern Recognition*, 1996, 29(7): 1061 - 1073.
- [4] WAKIN M B. Image compression using multiscale geometric edge models [D]. Houston: Rice University, 2002.
- [5] ROMBERG J K, WAKIN M B, BARANIUK R G. Multiscale wedgelet image analysis: fast decompositions and modeling [C]// *IEEE ICIP 02*. [S. l.]: IEEE Press, 2002.
- [6] WAKIN M B, ROMBERG J K, CHOI H, *et al.* Image compression using an efficient edge cartoon + texture model [C]// *IEEE Proceedings of the Data Compression Conference (DCC '02)*. [s. n.]: IEEE Press, 2002: 43 - 52.

(上接第 2151 页)

4 试验结果

表 1 Fast inter 和 UMHexagonS 的比较

格式	序列	$\Delta r/\%$	$\Delta PSNR/dB$	$\Delta kbits/\%$
QCIF	Foreman	-60.06	-0.08	+2.89
	Silent	-71.44	-0.06	+5.45
	News	-55.98	-0.11	+3.53
	Coastguard	-59.97	-0.02	+1.63
CIF	Container	-46.87	-0.02	+1.65
	Mobile	-31.20	-0.07	+0.32
	Pairs	-42.69	-0.08	+1.55
	Tempete	-57.50	-0.10	+3.96
	Highway	-75.46	-0.11	+2.32
	Sign_irene	-80.62	-0.13	+4.88

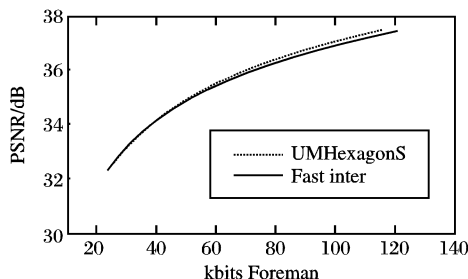


图 2 Foreman RD 曲线

该算法应用于 H. 264/AVC JM8. 6^[7] 参考软件中。在 IPPP, baseline 档次下, 参考帧为 1, 分别对 5 个 QCIF 和 5 个 CIF 150 帧测试序列进行试验, 就编码时间、比特率、PSNR 指标, 对 UMHexagonS 算法和在其基础之上的改进的 Fast inter 算法进行了比较。如表 1 所示: 当 $QP = 28$, Δt 为 Fast inter

较 UMHexagonS 节省的编码时间; $\Delta PSNR$ (dB) 为 Fast inter 较 UMHexagonS 下降的峰值信噪比, $\Delta kbits$ 为 Fast inter 较 UMHexagonS 提高的比特率。试验结果表明, 改进的帧间模式判别算法, 平均编码时间仅为 UMHexagonS 的 42.52%, 平均峰值信噪比下降仅为 0.076 dB, 平均比特率增加仅为 2.818%。图 2 表示 Foreman 视频序列两种算法的 RD 曲线, 从中可以看出本算法编码时间的减少是以峰值信噪比 (PSNR) 的下降, 单帧比特数的增加为代价的, 但基本上保证图像率失真和码率性能。

参考文献:

- [1] ITU-T Rec. H. 264/ISO/IEC14496-10 AVC (Doc JVT - G050), Draft ITU-T recommendation and final draft international standard of joint video specifications [S].
- [2] CHEN Z B, XU J F, HE Y. Fast motion estimation for JVT [C]// *JVT-G016*, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6). Pat-taya: [s. n.], 2003
- [3] RICHARDSON I. H. 264 and MPEG-4 video compression [M]. London: John Wiley & Sons, 2003.
- [4] LIM KP, SULLIVAN G, WIEGAND T. Text description of joint model reference encoding methods and decoding concealment methods [C]// *JVT-N046*, JVT of ISO/IEC MPEG and ITU-T VCEG. Hong Kong: [s. n.], 2005.
- [5] JING X, CHAU L P. Fast approach for H. 264 Inter mode decision [J]. *Electronics Letters*, 2004, 40(17): 1051 - 1052.
- [6] YU A C. Efficient block - size selection algorithm for inter - frame coding in H. 264 /MPEG AVC [C]// *2004 IEEE international Conference on Acoustics, Speech, and Signal Processing*. [S. l.]: IEEE Press, 2004: 169 - 172.
- [7] SCHÄFFER K. H. 264/AVC reference software [EB/OL]. [2007 - 02 - 10.] http://iphome.hhi.de/suehring/tml/download/old_jm/.