

文章编号:1001-9081(2006)12-2874-03

一种 H. 264/AVC 帧间块模式抉择快速算法

李俊杰^{1,2}, 齐 华², 管 莉², 樊养余^{1,2}, 郝重阳^{1,2}

(1. 西北工业大学 电子与信息研究所, 陕西 西安 710072;

2. 西安虚拟现实工程技术研究中心, 陕西 西安 710072)

(junprince@56.com)

摘要:传统的 H. 264/AVC 帧间模式抉择采用的是所有模式全搜索的算法, 计算量过大。针对这种情况提出了一种利用均匀区域、零运动矢量和块合并原理相结合的快速模式抉择算法, 并用 JM 模型进行实验, 结果表明, 该算法可以在 PSNR 降低很少(不超过 0.03dB)、码率增加不大(小于 4%)的情况下提高编码效率, 平均编码时间减少了 60% 以上。

关键词: H. 264/AVC; 均匀区域; 零运动矢量; 块合并

中图分类号: TP37; TP391.41 **文献标识码:** A

Fast inter-prediction mode decision for H. 264/AVC

LI Jun-jie^{1,2}, QI Hua², GUAN Li², FAN Yang-yu^{1,2}, HAO Chong-yang^{1,2}

(1. Institute of Electronic and Information Engineering, Northwestern Polytechnical University, Xi'an Shanxi 710072, China;

2. Xi'an Technology Research Center of Virtual Reality, Xi'an Shanxi 710072, China)

Abstract: A full search method used in reference software of H. 264 comes at a cost of greatly increased complexity of encoder. A new scheme was proposed to optimize block mode decision and improve the efficiency of the encoder in H. 264/AVC significantly. Homogeneous area, zero vector and block merging method were combined in this scheme. The experimental results of JM model show that the new scheme is able to achieve a reduction of 60% encoding time on average, with a negligible average peak signal-noise ratio (PSNR) loss (<0.03dB) and bit rate increase (<4%) compared with the original H. 264/AVC reference software.

Key words: H. 264/AVC; homogeneous area; zero vector; merged block

0 引言

H. 264 帧间预测^[1]采用的是一种基于块的运动补偿预测算法, 与以往标准相比最大的不同是使用变尺寸的块模式, 块尺寸从 16 × 16 到 4 × 4, 如图 1 所示共有 7 种不同尺寸的块模式, 按照从大到小分别命名为模式 1 ~ 7。

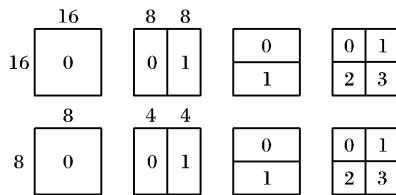


图 1 块模式划分

当选定一个比较大的块模式时, 只需要较小的码字进行编码传送, 但也意味着难于达到最优匹配, 也就是说在图像的残差部分包含着比较多的能量, 包含着图像的细节部分。选定一个较小的块尺寸, 可以显著降低在残差部分的信息, 但是付出的代价是产生比较高的码率来表示 MVD 和对应的分块策略, 因此在不同的条件下需要选择合适的分块策略。一般的情况下, 在低频信息较多的区域, 使用较大的分块策略, 而在有详细信息的地方采用比较小的分块策略^[2]。

H. 264 参考代码^[3]中采用的模式抉择算法是全搜索算法, 该算法遍历每种块模式, 选择代价函数最小的模式作为最佳模式, 但由于需要对每种块模式分别进行运动预测, 计算量非常大。因此本文提出了一种基于均匀区域^[4]、零运动矢量和块合并原理相结合的模式抉择快速算法。

1 基本原理

1.1 均匀区域判决原理

通过实验发现大模式所占的比例相当高, 并且这种大模式往往处于比较平坦的区域, 可以称之为均匀区域, 表 1 (编码 100 帧) 的实验结果表明, 对于运动不是很剧烈的序列, 大块模式几乎占了一多半, 因此若事先判决出该宏块是均匀区域, 就可以确定其最佳模式是大模式, 避免不必要的小模式搜索。

表 1 16 × 16 大模块的分布比例 (QP = 28)

图像序列	模式 0 和 1 (%)	其余模式 (%)
Foreman. qcif	55.05	44.95
Carphone. qcif	63.24	36.76
Grandma. qcif	90.25	9.75
Tennis. qcif	50.37	49.63

均匀区域的判决可以采用计算其方差、标准差、偏差等表

收稿日期: 2006-06-20; 修订日期: 2006-08-15

基金项目: 国家博士点基金资助项目(20040699015); 西北工业大学 2006 年度种子基金资助项目(M016630/06017)

作者简介: 李俊杰(1981-), 男, 河南孟津人, 硕士研究生, 主要研究方向: 视频压缩、DSP; 齐华(1963-), 女, 陕西咸阳人, 副教授, 主要研究方向: 通信与信息处理; 管莉(1967-), 女, 黑龙江宾县人, 博士研究生, 主要研究方向: 人工智能、虚拟现实等; 樊养余(1960-), 男, 陕西西安人, 教授, 博士生导师, 主要研究方向: 模式识别、虚拟现实; 郝重阳(1940-), 男, 河北保定人, 教授, 博士生导师, 主要研究方向: 图形图像、虚拟现实。

示图像纹理特性的统计量,也可以采用边缘检测方法进行判决^[5]。为了简化计算,这里采用了计算各像素与平均像素值的绝对误差之和的方法进行判决,计算公式如式(1),(2),如果计算出该块的值小于某个阈值,则判决该块采用模式 1 作为最佳块模式进行运动预测。表 2 的数据是 $QP = 28$ 时的实验结果,它表明该算法的准确性高达 85% 以上。

$$mean = \frac{1}{256} \sum_{i=0}^{15} \sum_{j=0}^{15} s(i,j) \quad (1)$$

$$\sum_{i=0}^{15} \sum_{j=0}^{15} |s(i,j) - mean| < TH \quad (2)$$

表 2 均匀区域最佳模式为模式 1 的概率

序列	命中率 (%)
Foreman. qcif	85.87
Carphone. qcif	87.71
Grandma. qcif	99.64
Tennis. qcif	89.73

另外需要注意的是,该算法的准确度与量化系数有关,量化系数越大准确性越高,一般量化系数大于 20 时,准确性高于 50%。图 2 是 tennis 测试序列的量化系数与算法准确度的关系曲线。

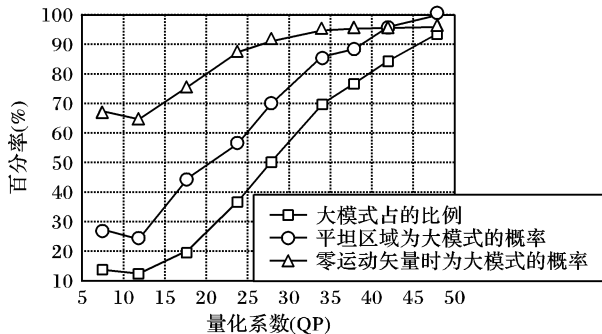


图 2 量化系数与大模式分布的曲线

1.2 零运动矢量决策原理

由于大模式相对于小模式难以达到最佳匹配,因此如果用大模式进行运动搜索得到的运动矢量为零,那么模式 1 为最佳模式的概率应该非常大。为此做了相应的实验,并得到表 3 的数据,结果证实了该设想的准确度(实验编码 100 帧数据, QP 取值 28)大于 83% 以上。因此可以事先对大模块进行运动搜索,如果获得的运动矢量为零,就可以判定该宏块的最佳块模式为模式 1,同时进一步判决该宏块的最佳模式是否为 SKIP 模式。图 2 显示了量化系数与该算法准确性的关系曲线,该曲线表明量化系数减小时,算法的准确性略有降低,但是最低仍然在 65% 以上,可见该算法是很准确的。

表 3 运动矢量为零时大模式为最佳模式的概率 ($QP = 28$)

序列	命中率 (%)
Foreman. qcif	83.96
Carphone. qcif	85.78
Grandma. qcif	99.03
Tennis. qcif	95.27

1.3 小模块合并原理

直观上可以认为,具有相似运动矢量的宏块采用大尺寸块进行运动估计编码更有效一些^[6],所以当两个小尺寸块具有相同或非常相似运动趋势时,将它们合并成大的块再进行运动估计和编码,会使编码效率有所提高。如图 3 所示,两个相邻小尺寸块 A, B 的运动矢量分别为 $MV1, MV2$, 两个块被合并后,合并块被标记为 C , 块 C 的运动矢量为 MV 。如果运动矢

量 $MV1, MV2$ 非常相似,使得块合并后编码效率优于小尺寸块,那么合并块 C 的运动矢量 MV 将会保持与 $MV1, MV2$ 的相似性,而且会以很大的概率落在矢量 $MV1, MV2$ 附近的区域内,实验证明该假设成立的概率在 90% 以上^[6]。

图 3 为块合并示意图,并且采用 SAD 表示合并块 C 最优运动矢量 MV 对应的块匹配误差; SAD_1 与 SAD_2 表示子块 A, B 最优运动矢量 $MV1, MV2$ 对应的块匹配误差; SAD_3, SAD_4 分别表示子块 A, B 对应平均运动矢量 MV' 的块匹配误差,其中 $MV' = (MV1 + MV2)/2$ 。

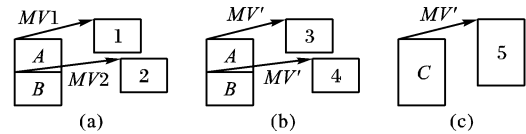


图 3 块合并图示

块匹配程度的代价函数^[2]如下:

$$J_i = SAD_i + \lambda_{motion} \cdot R(MVD_i) \quad (3)$$

式中 λ_{motion} 表示拉格朗日系数, $R(MVD_i)$ 表示编码块 i 的运动矢量与预测矢量的差值所需要的位长。假设用 J_A, J_B 表示相邻的两个块 A, B 的最优块匹配程度代价函数,用 J_A', J_B' 表示相邻的两个块 A, B 对应平均运动矢量的块匹配程度代价函数,用 J_C 表示块 A, B 合并后的块 C 的最优块匹配程度代价函数,那么如果将块 A, B 进行合并后,只有式(4)的条件满足,才认为合并是有效的。

$$J_C \leq J_A + J_B \quad (4)$$

通过分析可知:

$$J_C \leq J_A' + J_B' - \lambda_{motion} \cdot R(MVD') \leq J_A + J_B \quad (5)$$

因此只要满足:

$$J_A' + J_B' - \lambda_{motion} \cdot R(MVD') \leq J_A + J_B \quad (6)$$

就可以满足(4)式,为了提高匹配的效率,可以增加一个效率因子 k ,即:

$$J_A' + J_B' - k \cdot \lambda_{motion} \cdot R(MVD') \leq J_A + J_B \quad (7)$$

其中: $1 < k < 2$ 。经过试验测试得知取 $k = 1.45 \sim 1.55$ 时效果比较好。

另外可以认为:

$$R(MVD') \approx \frac{1}{2} (R(MVD_1) + R(MVD_2)) \quad (8)$$

则(7)式可化为:

$$SAD_3 + SAD_4 - SAD_1 - SAD_2 \leq k \cdot \frac{\lambda_{motion}}{2} \cdot (R(MVD_1) + R(MVD_2)) \quad (9)$$

于是,块合并的有效性判断准则可以如(9)式描述,其中效率因子 k 取 $1.45 \sim 1.55$ 。

合并运算时 16×16 的宏块被分为四个 8×8 的块,分别被标记为 a, b, c, d 。假如式(9)成立,则可判断该两块可以被合并。如果用 $M(a, b) = 1$ 来表示式(9)条件成立,那么块合并的判断过程可描述如下:

- 1) 若 $(M(a, b) = 1 \&\& M(c, d) = 1)$, 则采用模式 2;
- 2) 若 $(M(a, c) = 1 \&\& M(b, d) = 1)$, 则采用模式 3;
- 3) 若 $(M(a, b) = 1 \&\& M(c, d) = 1 \&\& M(a, c) = 1 \&\& M(b, d) = 1)$, 则采用模式 1;

4) 否则把 8×8 的块分成 4 个 4×4 的块进行合并判断,并根据判断结果合并成新块。

通过合并小块获得最佳块模式以后,对该模式进行运动预测获得最佳运动矢量。

2 快速抉择算法的计算步骤

根据上面的基本原理,本文快速模式抉择算法归纳为下

面的计算步骤:

步骤 1: 对模式 1 进行运动搜索, 获得运动矢量。

步骤 2: 计算该宏块是否为均匀区域。

步骤 3: 如果运动矢量为零或者该宏块为均匀区域, 则判断该宏块采用模式 1; 否则进入步骤 4。

步骤 4: 分别对该宏块内的四个 8×8 块进行小模式合并运算; 如果可以合并则对最佳块模式进行运动搜索, 算法结束, 否则进入步骤 5。

步骤 5: 分别对各 8×8 块内的 4×4 块进行小模式合并运算, 如果可以合并则对新合并的模式进行运动搜索, 算法结束; 否则判定模式 7 为最佳模式。

3 实验结果

对本文提出的快速算法进行了实验仿真。计算机配置为: P4 1.6GHz, 256M 内存; 实验环境: 编码器为 JM8.6, 帧速率为 30 帧/s, 运动矢量搜索范围为 $(-16, 16)$, 码率控制关闭, 熵编码 CAVLC, 编码帧数为 100 帧, 视频序列编码的 GOP 格式为 IPPPP…。实验结果如表 4~表 6 和图 4 所示。

表 4 快速算法与原算法决策的模式分布比例

模式	原算法 (%)	快速算法 (%)
模式 0	0.37	0.55
模式 1	54.68	53.78
模式 2	11.47	7.10
模式 3	15.81	12.43
模式 4	11.41	21.97
模式 5	2.55	1.08
模式 6	2.90	1.72
模式 7	0.81	1.37

表 5 Mobile 的参考帧分布比例

参考帧	原算法 (%)	快速算法 (%)
参考帧 1	47.15	47.43
参考帧 2	14.58	14.76
参考帧 3	15.17	15.53
参考帧 4	12.10	11.32
参考帧 5	11.00	10.95

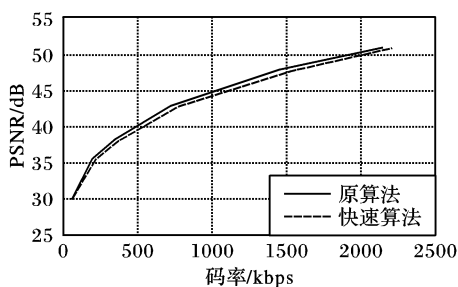


图 4 原算法与快速算法的 RD 曲线对比

由表 4 可以看出, 快速算法和原算法模式抉择后各模式的分布比例很相近, 特别是大模式的分布比例非常接近, 但小模式的比例有些差异, 这是因为本算法主要是从大模式的角度来提高编码速度, 对小模式的判决可能有一定的误差; 图 4 表明快速算法与原算法的 RD 曲线十分接近; 表 5 表明快速算法与原算法模式决策后参考帧的分布情况几乎完全一样; 由表 6 可以看出快速算法在不同量化系数时均可以保证 PSNR 降低很小 (平均不超过 0.03dB)、码率提高不大 (平均不超过 4%) 的同时提高编码效率, 编码时间平均减少 60% 以上。

表 6 快速算法与原算法编码时间和 PSNR 的对比

QP	图像序列	时间变化 (%)	PSNR/dB	码率 (%)
8	Foreman	-51.63	-0.02	+2.96
	Grandma	-64.38	-0.02	+1.13
	Table tennis	-46.51	-0.01	+2.93
	Mobile	-46.08	-0.01	+3.77
18	Foreman	-53.93	-0.03	+2.86
	Grandma	-65.97	-0.01	+2.39
	Table tennis	-54.28	-0.02	+3.44
	Mobile	-48.00	-0.02	+2.49
28	Foreman	-62.43	-0.01	+3.40
	Grandma	-69.68	-0.01	+2.92
	Table tennis	-59.16	-0.04	+3.58
	Mobile	-52.41	-0.02	+4.13
38	Foreman	-66.67	+0.01	+2.97
	Grandma	-69.53	-0.01	+1.12
	Table tennis	-62.71	-0.02	+4.23
	Mobile	-63.47	-0.02	+3.07

本文的算法是一种很有效的模式抉择快速算法, 可以在对编码质量产生很小影响的情况下提高编码效率。

参考文献:

- [1] Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H. 264 | ISO/IEC 14496-10 AVC) [S]. JVT of ISO/IEC MPEG and ITU-T VCEG, 2003. 131-152.
- [2] RICHARDSON I EG. H. 264/MPEG-4 Part 10 White Paper [Z]. 2003. 170-172.
- [3] JM Reference Software. JM8.6 [CP/DK]. <http://iphome.hhi.de/suehring/tml/>, 2005-10.
- [4] LIM KP, WU S, WU DJ, *et al.* Fast INTER Mode Selection [A]. JVT I020, 9th JVT meeting [C]. 2003. 2-7.
- [5] PURI A, HANG H, SCHILLING D. Interframe Coding with Variable Block Size Motion Compensation [A]. Proceedings of GLOBECOM'87 [C]. 1987. 65-69.
- [6] YU-KUANG TU, JAR-FERR YANG. Fast variable-size block motion estimation using merging procedure with an adaptive threshold [A]. Proceedings of ICME [C]. 2003. 789-792.
- [3] KRISHNAMURTHY B, WANG J. On network-aware clustering of Web clients [A]. Proceedings of ACM SIGCOMM [C]. 2000.
- [4] ZHAO BY, DUAN Y, HUANG L, *et al.* Brocade: landmark routing on overlay networks [A]. Proceedings of First International Workshop on Peer-to-Peer Systems [C]. 2002.
- [5] XU ZC, MAHALINGAM M, KARLSSON M. Turning Heterogeneity into an Advantage in Overlay Routing [A]. Proceedings of INFOCOM [C]. 2003.
- [6] ZHANG X, LIU J, ZHANG Q, *et al.* Measure: a group-based network performance measurement service for peer-to-peer applications [A]. GLOBECOM [C]. 2002.

(上接第 2828 页)

应实际网络中的关系的问题, 将资源本地化, 提高了系统的性能。该覆盖网适用于媒体分发以及应用层流媒体, 具有广阔的应用前景。

参考文献:

- [1] 中国计算机学会学术工作委员会. 中国计算机科学技术发展报告(2004) [R]. 北京: 清华大学出版社, 2004. 121-122.
- [2] STOICA I, MORRIS R, KARGER D, *et al.* Chord: a scalable peer-to-peer lookup service for Internet applications [A]. Proceedings of ACM SIGCOMM [C]. 2001.