

基于修正运动矢量的 H.264 帧间模式快速判决研究

高 振, 罗桂娥

(中南大学 信息科学与工程学院, 湖南 长沙 410083)

摘 要: H.264/AVC 引入的可变块尺寸运动估计技术, 使得帧间模式判决异常复杂。针对这个问题, 提出了一个新概念—相邻宏块模式同构, 并给出一种基于修正运动矢量的模式快速判决算法, 通过找到当前编码块 8×8 模式方向和幅值趋近一致的运动矢量, 进而修正运动矢量, 以使较小块尺寸模式可扩展为较大块尺寸模式。实验结果表明, 该算法在重构图像质量没有明显降低的前提下, 大幅度提高了编码速度。

关键词: H.264/AVC; 运动矢量; 模式判决; 帧间预测

中图分类号: TP37

文献标识码: A

文章编号: 1674-7720(2010)10-0045-04

Research of fast H.264 mode decision based on correcting motion vector

GAO Zhen, LUO Gui E

(School of Information Science and Engineering, Central South University, Changsha 410083, China)

Abstract: Mode decision becomes very complexity when H.264/AVC introduces variable block size about motion estimation. To address the issue, in this paper a novel concept about same mode structure of neighboring macroblock is proposed. Furthermore, a fast mode decision algorithm by gaining the direction and value about motion vector of current macroblock is also proposed, and the mode focuses on 8×8 at the beginning for expanding the block size with correcting motion vector. Simulation shows that the proposed algorithm can efficiently improve the coding speed with similar visual quality.

Key words: H.264/AVC; motion vector; mode decision; inter prediction

H.264/AVC^[1] 是由联合视频小组 JVT (Joint Video Team) 制定的新一代视频编码国际标准, 与以往的编码标准 H.263、MPEG-4 等相比, 它具有显著的编码效率, 其中一个重要因素就是引入了可变块尺寸划分模式等新技术, 如图 1 所示。但据试验统计, 可变块尺寸模式判决占到了整个编码计算复杂度的 70% 以上^[2], 因此简化多种块尺寸模式的选择流程, 对于减小 H.264 编码计算复杂度、实现其实时性应用至关重要。

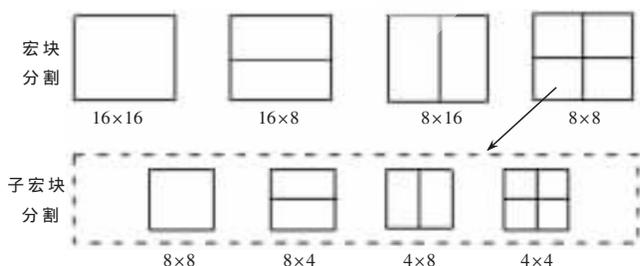


图 1 H.264 帧间预测的的 7 种块划分模式

迄今为止, 大量的研究致力于兼顾编码质量和编码速度的可变块尺寸模式判决快速算法上。参考文献[3]利用视频物体同质性来减少候选模式的数目; 参考文献[4]通过分析编码块的纹理信息及运动特点, 能跳过一些不必要的编码模式, 但这两种算法改进效果一般。本文在深入研究 H.264 视频编码标准及其运动特征和运动矢量的基础上, 针对 P 帧帧间预测, 提出一种基于修正运动矢量的帧间模式快速判决算法, 通过找到当前编码块 8×8 模式方向和幅值趋近一致的运动矢量, 以使较小块尺寸模式可扩展为较大块尺寸模式, 进而提高编码速度。

1 H.264 模式判决分析

H.264 采用树状结构运动补偿, 即把宏块分割成 16×16 、 16×8 、 8×16 、 8×8 等 4 种模式, 该过程称为宏块分割。如果选择 8×8 模式, 则 8×8 模式可进一步分割成 8×8 、 8×4 、 4×8 、 4×4 等 4 种模式, 该过程称为子宏块分割。再加上 H.264 开发的 SKIP^[5] 模式、帧内 16×16 预测

模式 Intra_16×16 和帧内 4×4 预测模式 Intra_4×4, P 帧帧间预测可供选择的模式共有 10 种 {SKIP、16×16、16×8、8×16、8×8、8×4、4×8、4×4、Intra_16×16、Intra_4×4}。其中 SKIP 是一种特殊的模式,无需进行运动估计。如果在编码过程中能首先判决出 SKIP 模式,将大幅度降低编码计算复杂度。

在视频编码中,10 种模式所采用的概率不同,通过对 2 帧 QCIF 和 2 帧 CIF 格式的视频序列在 H.264 参考编码器 JM8.6 下进行编码统计,得出如表 1 所示的统计结果。

表 1 SKIP 模式在所有模式中所占比例(%)

测试序列	量化参数 QP			
	28	32	36	40
Akiyo(QCIF)	82.9	88.0	91.8	94.4
Foreman(QCIF)	20.6	33.3	44.9	53.8
news(CIF)	72.4	78.2	82.9	86.4
Bus(CIF)	11.9	23.1	35.7	47.8

从表 1 可以看出,SKIP 模式在所有模式中所占比例较高,并且随着量化参数 QP 值的增大而逐渐升高比例。

在模式判决中,H.264 通过率失真优化 RDO (Rate-Distortion Optimization) 比较每种可能模式的率失真代价,然后从它们中间选择代价值最低的模式作为最佳编码模式。拉格朗日率失真代价函数定义如下:

$$J_{mode}(s,c,mode|\lambda_{mode})=SSD(s,c,mode)+\lambda_{mode}R(s,c,mode) \quad (1)$$

式中,SSD 代表原始视频信号和重构信号均方差, λ_{mode} 为拉格朗日算子比特数,mode 为上述 10 种模式中的其中一种,为编码该模式所需比特数。

2 最佳宏块划分模式的快速判决

根据原始视频图像的运动特征,可以将视频图像区域划分为 4 种类型,分别为静止区域、平滑区域、运动剧烈区域和高度激烈运动区域。结合帧间预测的 10 种模式,可以把运动区域类型和预测模式对应起来:候选模式集 P-mode-skip 包括 SKIP 模式,对应静止区域;模式集 P-mode-16 包括 16×16、16×8、8×16 三种模式,对应平滑区域;模式集 P-mode-8 包括 8×8、8×4、4×8、4×4 四种模式,对应剧烈区域;P-mode-intra 包括 Intra_16×16、Intra_4×4 两种模式,对应高度激烈区域。各种模式在所有模式中所占的比例如表 2 所示。

表 2 各种模式在所有模式中所占比例(%)

模式序列	SKIP	16×16	16×8	8×16	8×8	Intra
Akiyo	82.9	5.9	3.2	3.7	4.3	0.0
Foreman	20.6	35.4	17.5	15.8	9.3	1.4
Silent	65.2	12.1	4.8	6.3	10.4	1.2
news	72.4	8.4	4.7	5.6	8.7	0.2

表 2 实验数据表明,在帧间预测过程中,帧内预测模式的使用率非常低,可见在帧间预测过程中,忽略帧内预测模式集 P-mode-intra,对编码质量应该不会造成影响;进一步分析,帧间预测中的 8×4、4×8、4×4 等三种模式的使用率也非常低,故可以归结到 8×8 模式中,P-mode-8 即代表 8×8 模式。省略模式集 P-mode-intra 和模式集 P-mode-8 中 8×4、4×8、4×4 模式的判决,对编码质量构成的影响微乎其微,但对编码复杂度的降低贡献良多。

2.1 自适应 SKIP 模式优先判决

由于 SKIP 模式无需进行运动估计,没有像素残差和运动矢量残差,因此不需要再进行变换和量化等操作,所以它是所有模式中计算复杂度最低的,因此对于提高编码速度具有极大的作用。此外,原始视频图像存在着大量的静态背景区域及 SKIP 的使用比例随着量化参数的增大而不断提高等因素,更利于发挥 SKIP 模式的优势。以往大量学者的研究证明,SKIP 模式具有最低率失真 RD (Rate-Distortion)^[6] 代价值,并且极易与其他模式相区分。基于以上分析,本文采用一种兼顾编码速度和码率的自适应 SKIP 模式优先判决方案,选择 SKIP 模式作为当前编码块的最佳编码模式,只要它满足条件:当前编码块的 SKIP 模式的 RD 代价值 $J(\text{SKIP}) < rTh_{\text{SKIP}}$ 。

对于权重因子 r 来说,将根据视频图像运动特征自动调整,本文设置 3 个参数 {0.9, 0.7, 0.5}, 静止区域选择权重因子 0.9,平滑区域选择 0.7,剧烈运动区域选择 0.5。如果当前编码块选择 SKIP 作为最佳编码模式,它将不进行余下模式的判决。

2.2 相邻宏块模式同构判决

实验表明,当前编码块与相邻已编码块存在着高度的空间相关性。根据实验数据,本文引入相邻宏块预测模式同构 SMSNM (Same Mode Structure about Neighboring Macroblocks) 概念:如果当前编码块 E 的相邻已编码宏块 A、B 和 C (如图 2 所示) 都选择预测模式集 P-mode-16 或 P-mode-8, 则 E 直接在模式集 P-mode-16 或 P-mode-8 中选择最佳编码模式。



图 2 当前编码块及其相邻已编码块

实验数据来源于 Akiyo、Foreman、news 等测试序列,共计 500 个宏块。实验表明,当 E 的相邻宏块 A、B、C 都采用 P-mode-16 集或 P-mode-8 集时,E 采用 P-mode-16 集或 P-mode-8 集的比例高达 93% 以上。

2.3 修正运动矢量模式判决

JM8.6 参考编码器编码帧间宏块的模式选择顺序如

图形、图像与多媒体

下: 16×16、16×8、8×16、8×8、8×4、4×8、4×4、SKIP、Intra_4×4、Intra_16×16, 然后运用 RD 策略选择其中模式 RD 代价最小者为 E 的最佳编码模式。该遍历顺序极大地迟缓和了编码速度。实验发现, 当前编码宏块的 8×8 模式的 mv 值相近甚至相同的比例较高。基于以上实验发现, 本文提出修正 mv 模式快速判决算法。算法具体描述如下:

(1) 对 E 进行 8×8 模式判决, 记录其 4 个 mv 值, 分别记录为 P-mode8-mv(0)、P-mode8-mv(1)、P-mode8-mv(2) 和 P-mode8-mv(3), 如图 3 所示。

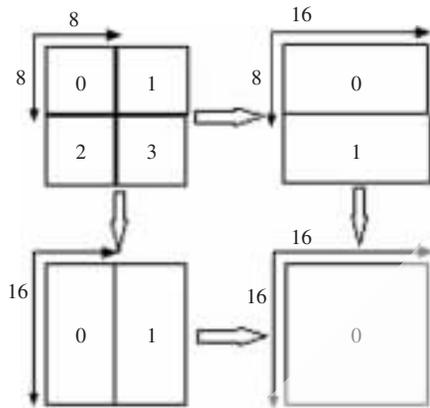


图 3 修正运动矢量扩展模式块尺寸

(2) 设置判决阈值

$$P\text{-mode-}mv_x(i) - P\text{-mode-}mv_x(j) \leq Th_x \quad (2)$$

$$P\text{-mode-}mv_y(i) - P\text{-mode-}mv_y(j) \leq Th_y$$

其中, $i, j \in (0, 1, 2, 3)$ 。

如果满足条件 condition1:

$$P\text{-mode-}mv_x(0) - P\text{-mode-}mv_x(1) \leq Th_x$$

$$P\text{-mode-}mv_y(0) - P\text{-mode-}mv_y(1) \leq Th_y$$

&&

$$P\text{-mode-}mv_x(2) - P\text{-mode-}mv_x(3) \leq Th_x$$

$$P\text{-mode-}mv_y(2) - P\text{-mode-}mv_y(3) \leq Th_y$$

或者满足条件 condition2:

$$P\text{-mode-}mv_x(0) - P\text{-mode-}mv_x(2) \leq Th_x$$

$$P\text{-mode-}mv_y(0) - P\text{-mode-}mv_y(2) \leq Th_y$$

&&

$$P\text{-mode-}mv_x(1) - P\text{-mode-}mv_x(3) \leq Th_x$$

$$P\text{-mode-}mv_y(1) - P\text{-mode-}mv_y(3) \leq Th_y$$

则 8×8 模式可扩展为 16×8 或 8×16 模式; 如果扩展后 16×8 或 8×16 模式的 mv 仍然满足公式(2), 则可以把 16×8 或 8×16 模式进一步扩展成 16×16 模式。

其中阈值 Th_x 和 Th_y 设置为(1, 1), 即 8×8 模式的 mv 在 x 轴、y 轴方向分别偏移一个像素; 扩展后的 16×8、8×16 或 16×16 模式, 其 mv 值由其扩展前的模式 mv 平均值获得, 即为:

$$P\text{-mode}16 \times 8\text{-}mv(0) = (P\text{-mode}8\text{-}mv(0) + P\text{-mode}8\text{-}mv(1)) / 2$$

《微型机与应用》2010 年第 10 期

$$P\text{-mode}16 \times 8\text{-}mv(1) = (P\text{-mode}8\text{-}mv(2) + P\text{-mode}8\text{-}mv(3)) / 2 \quad (3)$$

通过修正 mv 模式快速判决, 将不需要再对 16×8、8×16、16×16 等模式进行运动估计。

3 算法具体过程描述

算法由 3 部分组成, 分别为自适应 SKIP 模式优先判决、相邻宏块模式同构判决和修正 mv 模式判决。算法具体步骤描述如下:

(1) 如果 E 处于当前编码帧的第一行或第一列, 则跳转至步骤(5);

(2) 检测 SKIP 模式, 如果其 RD 代价小于 rTh_{SKIP} 阈值, 则选择 SKIP 模式为最佳编码模式, 跳转至步骤(6); 否则跳转至步骤(3);

(3) 检测 E 的 3 个相邻已编码宏块, 如若满足 SMSNM 概念, 则在相应的 P-mode-16 或 P-mode-8 模式集中选择最佳编码模式, 跳转至步骤(6), 否则跳转至步骤(4);

(4) 分割 E 至 8×8 模式, 检测其 mv 是否满足 condition1 或 condition2, 如若满足则进行模式块尺寸的扩展, 否则跳转至步骤(5);

(5) E 在所有模式中遍历, 选择其 RD 代价最小者作为当前宏块最佳编码模式;

(6) 结束当前模式判决。

算法流程图如图 4 所示。

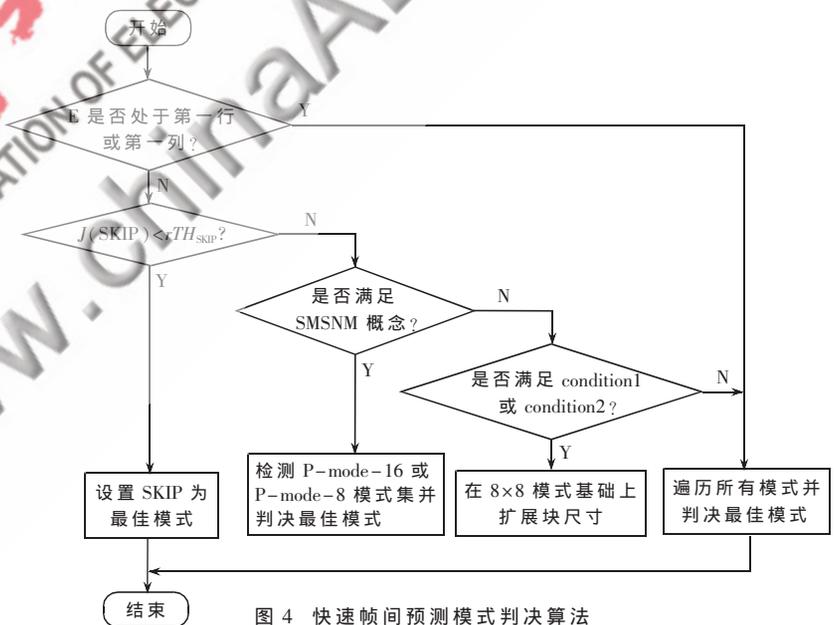


图 4 快速帧间预测模式判决算法

4 实验结果及性能分析

在实验中用到 3 个测度指标验证本文算法性能, 与参考编码器标准算法相对应, 分别为重建图像质量的降低百分比 $\Delta PSNR$ (dB)、输出码率的升高百分比 ΔBR (%) 和减少编码时间百分比 ΔT (%)。

(1) 实验软件平台: Windows 2003 操作系统, JM8.6 标准编码器, Visual C++ 编译器。

图形、图像与多媒体

Image Processing and Multimedia Technology

(2)实验硬件平台:P4,2.93 GHz,512 MB 内存。

(3)编码参数设置:帧间预测参考帧数量为1帧,编码采用 IPPP 模式,量化参数为28,使用RDO,帧数30 Hz,运动搜索范围为 ± 16 ,熵编码采用CAVLC编码。

(4)实验数据比较:实验中对Akiyo、Table、Foreman 3个QCIF测试序列和Silent、news、bus 3个CIF序列进行测试,测试结果如表3所示。

表3 QCIF和CIF序列测试结果

序列	$\Delta PSNR/dB$	$\Delta BR/(%)$	$\Delta T/(%)$
Akiyo	-0.150	+0.42	-62.43
Table	-0.084	+0.81	-50.44
Foreman	-0.046	+0.96	-36.12
Silent	-0.043	+0.87	-43.82
news	-0.170	+1.38	-54.50
bus	-0.018	+0.65	-37.74

从表3可以看出,与参考编码器的算法相比,本文所提出的算法在信噪比(PSNR)平均仅降低0.085 dB和比特率平均仅升高0.848%的情况下,平均编码时间减少了47.51%,大幅度降低了编码器的计算复杂度。再者Akiyo和news序列速度提高尤为明显,因为低速运动的序列较多地采用了编码速度较快的SKIP模式和较大块尺寸模式作为最佳编码模式,并且其中使用到的较小块尺寸的mv趋近于相同,故可以扩展成较大块尺寸模式,所以更加出色地发挥本文算法的优势。

本文在深入研究帧间预测运动特征和运动矢量的基础上,提出一种基于修正运动矢量的帧间模式快速判决算法,通过找到当前编码块 8×8 模式方向和幅值趋近一致的mv,进而修正mv,以使较小块尺寸模式可扩展为较大块尺寸模式。再者,本文在依据大量具有代表

性统计数据的基础上,提出的SMSNM概念对于这类快速算法具有一定的参考价值。实验数据表明,该算法在保证编码质量的同时,运动估计速度有了显著的提高,更利于视频压缩的实时性应用,是一种综合性能良好的帧间模式快速判决算法。

参考文献

- [1] WIEGAND T, SULLIVAN G T. Overview of the H.264/AVC video coding standard[J]. IEEE Transactions on Circuit and System for Video Technology, 2003,13(7):560-576.
- [2] ZENG Huang Qiang, CAI Can Hui, MA Kai Kuang. A novel fast mode decision for the H.264/AVC Based on local macroblock motion activity[J]. ICIG, 2007(4):263-26.
- [3] LIM K P, WU S, WU D J, et al. Fast inter mode selection[S]. Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, JVT I020, 2003.
- [4] WU D, WU S, LIM K P, et al. Block inter mode decision for fast encoding of H.264[J]. ICASSP, 2004(3):181-184.
- [5] PAN L J, HO Y S. Fast mode decision algorithm for H.264 inter-prediction[J]. Electronics Letters, 2007,43:1351-1353.
- [6] ZHOU Xin Xin, YUAN Chun, LI Chun Hua, et al. Fast mode decision for P-slices in H.264/AVC Based on Probabilistic Learning[J]. ICACT, 2009(11):1180-1184.

(收稿日期:2010-03-10)

作者简介:

高振,男,1985年生,硕士研究生,主要研究方向:视频压缩编码及流媒体。

罗桂娥,女,1962年生,教授,主要研究方向:智能仪器仪表的研究与开发、信息融合技术、数字图像处理技术。