文章编号:1001-9081(2005)08-1811-04

# 基于混合特征的 H. 264/AVC 快速帧内预测模式决策

葛睍睍¹,张俊峰²,郝重阳¹,王 宇¹,席迎来¹ (1. 西北工业大学 电子信息学院,陕西 西安 710072; 2. 平顶山工学院 计算机科学与工程系,河南 平顶山 467000) (cindysaigxx@126.com)

摘 要:H. 264/AVC 中采用率失真优化(RDO)代价函数对多种帧内预测模式进行全搜索决策,为了克服这种决策方法运算复杂度较大的弊端,提出了一种基于混合特征的快速帧内预测模式决策算法。该方法针对最耗时的 4×4 亮度块模式决策环节,强调了 DC 模式的重要性,并将其与混合代价函数特征相结合形成混合决策信息,以此对候选模式进行预先筛选,从而加快了决策速度。实验结果表明,该方法可以使图像序列的 I 帧编码速度提高近 30%,且与原全搜索算法相比,其峰值信噪比

关键词:H. 264/AVC; 帧内预测; 混合特征; 快速帧内预测模式决策

中图分类号: TP317.4;TN919.8 文献标识码:A

(PSNR)的减少和比特率的增加均小到可以忽略不计。

# Mixed features-based fast intra-prediction mode decision for H. 264/AVC

GE Xian-xian<sup>1</sup>, ZHANG Jun-feng<sup>2</sup>, HAO Chong-yang<sup>1</sup>, WANG Yu<sup>1</sup>, XI Ying-lai<sup>1</sup>

(1. School of Electronic Information, Northwestern Polytechnical University, Xi'an Shaanxi 710072, China;

2. Department of Computer Science and Engineering, Pingdingshan Industrial College, Pingdingshan Henan 467000, China)

Abstract: Due to the use of Rate Distortion Optimization (RDO) for the decision of multiple intra-prediction modes, a full search method used in reference software of H. 264/AVC comes at a cost in greatly increased complexity at the encoder. In order to improve the time efficiency of the encoder, a mixed features-based on fast intra-prediction mode decision method aiming at 4 × 4 luminance blocks was proposed. DC mode was emphasized in this method, and it was combined with the mixed features of cost functions to form mixed decision information. Advance modes were filtered by virtue of these mixed features, and the decision approach was speeded up. The experiment results show that the new scheme is able to achieve a reduction of up to 30% encoding time for I frames of image sequences, with negligible Peak Signal-Noise Ratio (PSNR) loss and bit rate increase compared with the original H. 264/AVC reference software.

Key words: H. 264/AVC; intra-prediction; mixed features; fast intra-prediction mode decision

# 0 引言

新一代视频编码标准 H. 264/AVC<sup>[1]</sup> 由 ITU-T VCEG (Video Coding Experts Group)和 ISO/IEC MPEG (Moving Pictures Experts Group)共同建立的联合视频专家组 JVT(Joint Video Team)制定并于 2003 年正式获得通过。H. 264 的主要目标是提高编码效率和提高网络适应性。它提供了高质量的视频,具有较强的抗误码特性和网络友好性,可以应用于低延时模式下的实时视频会议、视频电话以及没有延时限制的视频存储、视频流等。相比较于以前的诸如 H. 263 + 和 MEPG-4 等视频压缩标准,最新的 H. 264 视频编码标准能提供比前述标准高 1.5~2.0 倍的压缩效率。

H. 264/AVC 的显著特点包括 4×4 的整数离散余弦变换 (Discrete Cosine Transform, DCT)、更精确的 I 帧 (intra 帧, 帧内预测帧) 帧内预测编码、1/4 像素运动估计、多参考帧以及 P 帧 (inter 帧, 帧间预测帧)的多种块尺寸运动预测,这些特点

使得 H. 264 的编码效率大幅度提高<sup>[2]</sup>。其中,帧内预测编码 是指利用相邻宏块的空间相关性进行预测编码,以提高帧内 编码的效率,这种预测编码不仅可以应用于编码 I 帧,也是编码 P 帧、B 帧(双向预测帧)时的候选模式<sup>[3,4]</sup>。

为进一步提高帧内预测编码效率,对于亮度分量,H. 264 对含有较多空域细节信息的宏块采用 4×4 预测(9 种模式), 而对于较平坦的区域采用 16×16 的预测模式(4 种模式),另 外还提供了不经预测和变换量化的步骤 I-PCM 模式<sup>[5]</sup>;对于 色度分量有 4 种 8×8 的预测模式,其预测方向和 16×16 亮 度块预测模式一致(仅模式号不同)。

由于多种预测模式的存在,这就要求在编码 I 帧时对这些模式进行择优,而择优的方法很多,在本文所参考的 JM7.6<sup>[6]</sup>代码中用了全搜索的方法,这种方法在引入率失真优化(Rate Distortion Optimization, RDO)后,计算量比较大,因此可以对决策步骤进行优化,形成一些快速帧内模式决策算法,例如基于边缘方向直方图的方法<sup>[7]</sup>和基于分组像素点的

收稿日期:2005-02-04;修订日期:2005-05-11 基金项目:航空科学基金资助项目(02I53071);国家重点实验室基金资助项目(51473010103HK0311);高等学校博士点基金资助项目(20040699015)

作者简介:葛晛晛(1982-),女,安徽蚌埠人,硕士研究生,主要研究方向:视频编码、多媒体通信; 张俊峰(1967-),男,河南平顶山人,剧高级工程师,主要研究方向:基于网络的计算机技术; 郝重阳(1940-),男,河北保定人,教授,博士生导师,主要研究方向:图像工程、虚拟现实,多媒体; 王宇(1981-),男,河南平顶山人,博士研究生,主要研究方向:多媒体通信,通信网络; 席迎来(1977-),男,陕西人,博士研究生,主要研究方向:视频编码、多媒体通信.

方法<sup>[8-10]</sup>,这些方法有各自的优点和弊端。本文在空域和变换域联合决策方法<sup>[11]</sup>的基础上提出了一种基于混合特征的快速决策方法,该方法主要针对关键的 4×4 亮度块预测模式,强调了 DC 模式的重要性,并利用了混合代价函数特性预先对 9 种候选模式进行筛选。实验结果表明,该方法在保证图像保真度和码率性能的前提下,较大幅度地加快了帧内模式决策速度。

# 1 全搜索(FS)的帧内预测模式决策算法

JM7.6 程序中采用全搜索的方法来进行帧内预测模式的 决策,步骤为:

- 1)根据一种预测模式产生 4×4 的亮度预测块,计算原始 4×4 块和预测块间的残差块;
  - 2)计算代价函数:
  - (1)非 RDO:

$$Cost_{4\times4} = \sum_{i=1}^{16} coeff_i \mid SATD(diff) \mid + 4P\lambda(Q_P)$$
 (1)

其中,diff = 原始块s - 预测块p,SATD为对残差块diff做哈达马变换后变换系数的绝对值之和,也可不选择做哈达马变换,这时SATD转化为SAD,SAD是对diff块各像素点绝对值的直接求和。将SATD值与 $4P\lambda(Q_p)$ 相加为代价函数 $Cost_{4\times 4}$ 。  $\lambda(Q_p)$ 。是量化因子 $Q_p$ 的指数函数的近似:

$$\lambda(Q_P) = \lambda_{\text{mod } e} = 0.85 \times 2^{(Q_P - 12)/3}$$
 (2)

对于  $most\_probable\_mode$ (最可能模式,指当前块的左邻和上邻块最佳预测模式中的较小模式),P=0;对其他模式, P=1。

(2) 若选用 RDO, 可通过计算 intro4 × 4 亮度块的 RDO 作为代价函数进行择优:

$$Cost_{4\times4} = RDO = J_{mod e}(s, c, MODE \mid \lambda_{mod e}) = SSD(s, c, MODE) + \lambda_{mod e} \cdot R(s, c, MODE)$$
(3)

其中:

$$SSD = \sum_{(x,y) \in A} |s(x,y,t) - s'(x,y,t)|^{2}$$
 (4)

为原始块 s 与重构块 s' 的均方残差 R(s,c,MODE) 为残差块编码后的码流位数。

- 3) 对所有9种4×4模式重复1)~2)步,然后选择最小的 Cost<sub>4×4</sub>;
- 4) 对宏块内的 16 个 4 × 4 块重复步骤 1) ~ 3),对所有 Cost<sub>4×4</sub> 求和;
- 5)根据一种16×16的预测模式产生一个预测宏块,对残差宏块的每个4×4块进行哈达马变换;
- 6) 从 16 个 4 × 4 块中提取每个 DC 系数的 1/4 形成另一个 4 × 4 块,对 DC 系数块进行哈达马变换;
- 7) 对所有哈达马变换系数的绝对值求和(除步骤 5) 中的 DC 系数),取和的 1/2 作为代价函数值;
- 8) 对 4 种  $16 \times 16$  模式重复步骤 5) ~ 7),选择最小的代价函数值  $Cost_{16\times 16}$ ;
  - 9) 检测步骤 4) ~ 8) 的代价函数值,如果:

$$Cost_{16\times 16} \ge \sum_{i=1}^{16} Cost_{4\times 4,i} + 24\lambda(Q_P)$$
 (5)

则当前块选取4×4的 intra 预测模式,否则,选取16×16的预测模式;

10) 对于色度块,先做4种8×8模式的预测,然后用4×4亮度块的第一种代价函数进行模式决策。

由此可见,一个宏块内的模式组合数为:

$$M8 \times (M4 \times 16 + M16) \tag{6}$$

其中 M8 = 4、M4 = 9、M16 = 4,分别代表  $8 \times 8$  色度块、  $4 \times 4$  亮度块和  $16 \times 16$  亮度块的帧内预测模式个数,所以将要计算 592 个不同的代价函数以选择最佳模式,计算复杂度很大,尤其是在运用 RDO 后,虽然保真度大幅度提高了,但耗时过大,需要采用一些快速算法,在保证或提高其编码精确度的基础上对其进行优化。

# 2 基于混合特征的帧内预测模式决策算法

基于混合特征的帧内预测模式决策算法主要针对关键的 4×4 亮度块的9 种预测模式,突出了 DC 模式的重要性,并将 其与空域、变换域特征相结合,以形成混合特征进行联合决策。

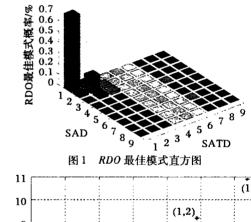
## 2.1 混合特征的提取

$$SAD = \sum_{(x,y) \in A} |D(x,y,t)| = s(x,y,t) - p(x,y,t) | (7)$$

$$SATD = \sum_{(x,y)\in A} |T\{D(x,y,t)\}|$$
 (8)

对三个 *QCIF* 图像序列的300 帧进行测试,针对4×4亮度 块预测模式得出以下结论:

1) 以  $4 \times 4$  亮度块的 9 种模式的 SAD 和 SATD 幅度等级为 x 轴、y 轴,建立图像序列的 RDO 最佳模式的直方图,发现大部分最佳模式落在(SAD, SATD) 的  $3 \times 3$  窗内,如图 1 所示。



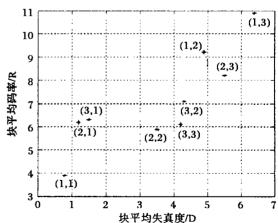


图 2 各幅度等级(SAD, SATD) 的平均码率(R) 失真(D) 关系

- 2) 由(1,1) ~ (3,3) 各等级模式的 R-D 关系(如图 2), 发现模式(1,1) 近似于 RDO 的最佳模式,其次为模式(2,1), (3,1),…
- 3) 不在这个 3 × 3 窗范围内的 RDO 最佳模式有 60% 为 DC 模式(2 号模式);
- 4) DC 模式为 SATD 或者 SAD 的次小模式或第三小模式时,选择 DC 模式;
  - 5) SATD 或者 SAD 的三个较小模式的幅值相等时选择 DC

模式。

由以上结论建立如下的决策思路:

定义阈值变量  $T_1, T_2, T_1'$  和  $T_2'$ :

$$T_1 = \frac{SATD_2 - SATD_1}{SATD_1} \tag{9}$$

$$T_2 = \frac{SATD_3 - SATD_1}{SATD_1} \tag{10}$$

$$T_1' = \frac{SAD_2 - SAD_1}{SAD_1} \tag{11}$$

$$T_2' = \frac{SAD_3 - SAD_1}{SAD_1} \tag{12}$$

其中, $SATD_1$  和  $SAD_1$  分别为 SATD 和 SAD 的最小值,  $SATD_2$  和  $SAD_2$  分别为 SATD 和 SAD 的次小值, $SATD_3$  和  $SAD_3$  分别为 SATD 和 SAD 的第三小值。

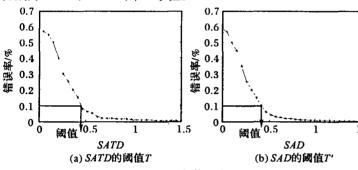


图 3 阈值的选取

以错误判决率为依据,分别选取 SATD 和 SAD 的 阈值 T 和 T',如图 3 所示,以实验中的 foreman. qcif 图 像序列为例:  $T \approx 0.43$ ,  $T' \approx 0.4$ 。

先对 SATD,再对 SAD 进行阈值检测。若  $T_1$  大于阈值 T 则直接选取  $SATD_1$  模式和 DC 模式为候选模式,其余若  $T_1$ ′大于 T′则选取  $SAD_1$  模式和 DC 模式为候选模式。

由结论 4 和 5,为保证决策精度,对于通过的上述检测步骤的块,若其 SATD 或 SAD 的三个较小模式的幅值相同,即  $T_1 = T_2 = 0$  或  $T_1' = T_2' = 0$ ,则直接选择 DC 模式;对于不满足的块,若其 SATD 或 SAD 的三个较小模式中有 DC 模式,则选择 DC 模式。

### 2.2 系统模型

按照以上思路,搭建4×4亮度块帧内预测模式决策系统 框图如图4。

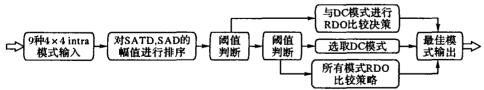


图 4 模式决策系统框图

#### 2.3 决策步骤

- 1)根据一种预测模式产生4×4亮度预测块, 计算原始4×4块和预测块间的残差块;
- 2) 计算 4 × 4 块 9 种帧内预测模式的 SATD 值,如公式(8)。对 9 种模式的 SATD 进行排序,提取 SATD 值较小的三种模式;
- 3) 计算 SATD 阈值变量  $T_1$  和  $T_2$ ,若  $T_1$  大于训练阈值  $T_1$ 则直接选取 SATD 最小模式和 DC 模式为候选模式,再计算两种模式的 RDO 值(公式(3)),进行比较判断,选取最终的最佳预测模式,到步骤 8);

- 4) 计算 SAD 阈值变量  $T_1'$  和  $T_2'$ ,若  $T_1'$  大于 T',则将 SAD 最小模式与 DC 模式进行比较决策,到步骤 8);
- 5) 若阈值变量不满足3) 和4) 的条件,则比较 $T_1$  和 $T_2$  以及 $T_1$  和 $T_2$  ,若两者相等且均为0,则表示至少有三种模式的SATD 或者SAD 值一致,选择DC 模式为最佳模式,到步骤8):
- 6) 若 SATD 或 SAD 的较小模式中有 DC 模式,则选取 DC 模式为最佳模式,到步骤 8);
  - 7) 计算所有模式的 RDO, 比较后进行模式决策;
- 8) 模式决策结束,输出最佳 4 × 4 亮度块帧内预测模式, 然后进行帧内预测。

# 3 实验结果及分析

用本文提出的基于混合特征的快速帧内预测模式决策算法对六种 QCIF 格式的图像序列进行测试,其中快运动序列包括 foreman, coastguard 和 news,慢运动序列包括 motr\_dotr,akiyo 和 container。实验中采用 I 帧、P 帧和 B 帧编码,且帧内预测模式也是 P 帧、B 帧编码时的候选模式,帧速率为 30Hz。将 foreman 和 motr\_dotr 两序列的决策时间和性能参数的变化记录如表 1(与原全搜索算法进行比较)。其中, QP 为编码量化系数, PSNR (dB)为本文算法与原全搜索算法相比图像序列峰值信噪比的变化量, Bitrate (%)为序列比特率的变化百分比, Time (%)为 I 帧编码时间的变化百分比。

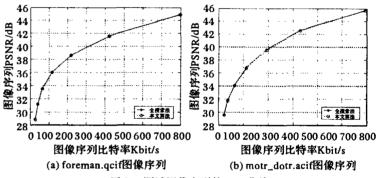


图 5 测试图像序列的 R-D 曲线

由表格数据可知,与原全搜索算法相比,本文方法对快速和慢速运动图像序列均有很好的优化效果。I 帧编码时间减少量接近 30%; PSNR 的降低 ≤0.08,有时还有增加,所以可以忽略;比特率的增加 < 6%,有时还有减少,因此也可以忽

略。还可以通过序列的 R-D 曲线更直观地分析算法的性能参数,仍以foreman 和 motr\_dotr 两序列为例,如图 5 所示。由图 5,本文算法和原全搜索算法的R-D曲线非常近似,几乎重合。由此可见,本文算法较大幅度

地提高了 I 帧编码速度,对图像序列性能的影响却微乎其微,

表 1 测试图像序列的性能参数

(a) foreman. qcif 图像序列				(b) motr_dotr, qeif 图像序列			
QP	PSNR/dB	Bitrate(%)	Time(%)	QP	PSNR/dB	Bitrate(%)	Time(%)
16	0.00	0. 69	- 25. 8	16	- 0.04	0.38	- 20. 1
20	- 0.06	- 0. 59	- 28. 5	20	- 0.03	0.55	- 28.9
24	- 0. 08	- 0.36	- 27. 2	24	- 0. 08	0.95	- 22. 9
28	- 0. 02	0. 21	- 23. 7	28	- 0. 08	0. 92	- 33. 6
32	- 0. 03	4. 10	- 25.3	32	- 0. 01	2.40	- 25. 0
36	0.05	2. 54	- 25. 4	36	- 0. 05	0.08	- 24. 1
40	- 0. 02	5. 10	- 21. 9	40	- 0. 07	5. 29	- 23. 9

因此,本文提出的基于混合特征的帧内预测模式决策算法对原算法的优化效果非常显著。

### 4 结语

针对 H. 264/AVC 参考代码中帧内预测模式决策算法运算复杂度较大的弊端,提出了一种基于混合特征的快速帧内预测模式决策算法。该方法强调了 DC 模式的重要性,并结合混合代价函数特性预先对候选模式进行筛选,由此加速了帧内预测模式决策中最耗时的 4×4 亮度块模式决策环节。通过对测试图像序列的性能参数和 R-D 曲线的分析表明,该方法在保证图像保真度和码率性能的前提下,将 I 帧编码速度提高了近 30%,对原算法具有较显著的优化效果,进而提高了编码系统的效率。为了提高整个系统的效率,今后还可以将本算法应用于解码系统,以进一步增加本算法的实用性。参考文献:

- [1] Draft ITU-T Recommendation and Final Draft International Standard of Joint Viedo Specification (ITU-T Rec. H. 264 | ISO/IEC 14496-10 AVC) JVT Doc. JVT-G050 [Z]. 2003.
- [2] THOMAS W, GARY JS. Overview of the H. 264/AVC Video Coding Standard [J]. IEEE Transactions on Circuits and Systems for Video

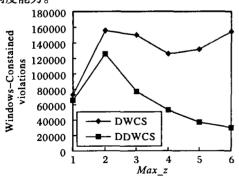
- Technology, 2003, 13(7): 560 576.
- [3] PENGY, HYEYCT, AIEXIS MT, et al. Fast Mode Decision and Motion Estimation for JVT/H. 264[A]. IEEE International Conference on Multimedia&Expo[C]. Maryland, 2003, III: 853 – 856.
- [4] LIM KP, WU S, WU DJ, et al. Fast INTER Mode Selection[A]. JVT 1020, 9th JVT meeting[C]. 2003.
- [5] RICHARDSON I EG. H. 264/MPEG-4 Part 10 White Paper[S]. 2003
- [6] Joint Video Team(JVT), reference software JM7.6[S].
- [7] PAN F, LIN X, SUSANTO R, et al. Fast Mode Decision for Intra Prediction Al. JVT G013, 7th JVT meeting Cl. 2003.
- [8] MENG BJ, OSCAR CA, WOMG CW, et al. Efficient Intra-Prediction Algorithm in H. 264 [A]. IEEE International Conference on Multimedia&Expo[C]. Maryland, 2003, III: 837 ~ 840.
- [9] MENG BJ, OSCAR CA. Fast Intra-Prediction Model Selection for 4 ×4 Blocks in H. 264[A]. ICASSP 2003[C]. 2003. 389 - 392.
- [10] MENG BJ, OSCAR CA, WOMG CW, et al. Efficient Intra-Prediction Mode Selection for 4 × 4 Blocks in H. 264 [A]. ICME 2003 [C]. 2003. 521 524.
- [11] KIM CS, SHIH HH, KUO JC-C. Fast Intra-Prediction Mode Selection for H. 264 Codec[A]. SPIE International Symposium ITCOM 2003[C]. Orlando, Florida, 2003.

### (上接第1782页)

 $Max_z > 2$  时明显下降,因而在z 窗口增大时,两者的差距更加明显。

### 3) Max\_W 对窗口违例数的影响

实验在  $Max_z = 3$  以及 U = 1.1 的情况下模拟出  $Max_w$  与总窗口违例数的变化关系,如图 5 所示。从图 5 可看出,总体上 DWCS 算法的窗口违例数明显要高于 DDWCS 算法。而随着  $Max_w$  的增加,DWCS 算法下的窗口违例数的下降不明显,而 DDWCS 算法的窗口违例数却在  $Max_w > 0.3$  时呈明显的下降趋势,从而使 DDWCS 算法的优势更加明显。这说明新的算法对 w 的变化更为敏感。这是因为 DDWCS 算法在 w 限制放宽的情况下,能够更加容易地利用节约的带宽满足任务的窗口限制,这也就说明了 DDWCS 算法比 DWCS 算法有着更强的调度能力。



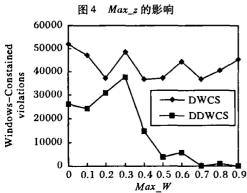


图 5 Max\_W 的影响

# 4 结语

本文提出了一种新的受窗口限制的帧间相关周期模型,同时针对该模型的帧间相关的特点,提出了基于动态双窗口限制调度算法,实验结果表明新算法比旧算法有更好的调度能力,具有非常明显的优势。利用该模型人们可以在帧间相关的实时应用系统中,采用动态双窗口限制调度算法来对QoS 进行控制以及进行任务调度。

同时随着计算机以及网络多媒体技术的迅速发展,对实时操作系统的发展也提出了新的要求,譬如需增加对网络功能的支持以及相应的服务质量的支持,因此研究支持 QoS 的实时操作系统是当前一个研究热点。本文所研究的调度模型及算法也完全可应用于支持 QoS 的实时操作系统中。

#### 参考文献:

- [1] LIU CL, LAYLAND JW. Scheduling Algorithms for Multiprogramming in a Hard Real-Time Environment[J]. Journal of the ACM, 1973, 20(1): 40-61.
- [2] HAMDAOUI M, RAMANATHAN P. Evaluating Dynamic Failure Probability for Streams with (m, k)-Firm Deadlines [J]. IEEE Transactions on Computers, 1997, 45(12): 1325 – 1357.
- [3] WEST R, POELLABAUER C. Analysis of a Window Constrained Scheduler for Real-Time and Best-Effort Packet Streams [A]. Proceedings of the 21st IEEE Real-time Systems Symposium (RTSS) [C]. Orlando: IEEE, 2000. 239 248.
- [4] WEST R, SCHWAN K. Dynamic Window-Constrained Scheduling for Multimedia Applications [A]. 6th Proceedings of the IEEE International Conference on Multimedia Computing and Systems (ICMCS) [C]. Florence, Italy: IEEE, 1999, 87 - 91.
- [5] BERNAT G, BURNS A, LIAMOSI A. Weakly Hard Real-Time Systems[J]. IEEE Transactions on Computers, 2001, 50(4): 308 321
- [6] MOK AK, WANG W. Window-Constrained Real-Time Periodic Task Scheduling A]. In Proceedings of the 22st IEEE Real-Time Systems Symposium C]. London; IEEE, 2001. 15 - 24.
- [7] MOK AK, CHEN D. A Multiframe Model for Real-Time Tasks[J]. IEEE Transactions on Software Engineering, 1997, 23(10): 635 -645.
- [8] 黄文广. 一个改进的实时任务模型—周期多帧任务模型[J]、计算机研究与发展,2001,38(2):234-239.