

# 一种基于 H.264/AVC 的帧内预测模式快速选择算法

肖广<sup>1)</sup> 滕国伟<sup>2)</sup> 王国中<sup>2)</sup>

<sup>1)</sup>(华东师范大学,上海 200062) <sup>2)</sup>(上海广电(集团)公司中央研究院,上海 200233)

**摘要** 在 H.264/AVC 中,帧内编码是采用帧内预测来提高编码效率,其色度块尺寸为  $8 \times 8$ ,亮度块包括  $4 \times 4$  和  $16 \times 16$  两种尺寸。由于每种块可提供多种预测模式,因此这种多模式选择大大增加了编码的复杂性。针对这种多模式选择,为了加快编码速度,利用色度块和亮度块以及亮度  $4 \times 4$  和  $16 \times 16$  块模式选择之间的关系,提出一种基于 RDO(rate distortion optimization)的不同判决依据的快速帧内预测模式选择算法。实验证明,在信噪比和码率变化不大的同时,该算法可使编码速度平均提高 6 倍。

**关键词** H.264/AVC 帧内预测 模式选择

中图分类号: TN919.81 文献标识码: A 文章编号: 1006-8961(2005)11-1375-04

## A Fast Intra-prediction Mode Selection Algorithm Based on H.264/AVC

XIAO Guang<sup>1)</sup>, TENG Guo-wei<sup>2)</sup>, WANG Guo-zhong<sup>2)</sup>

<sup>1)</sup>(East China Normal University, Shanghai 200062)

<sup>2)</sup>(Central Research Academy, SVA Group, Shanghai 200233)

**Abstract** In H.264/AVC, intra-prediction is used to improve the coding efficiency of I-frame. There are one size of  $8 \times 8$  for the chroma block and two sizes including  $4 \times 4$  and  $16 \times 16$  for the luma block. Each kind of block provides many prediction modes, so that this multi-mode selection increases the coding complexity greatly while improving coding efficiency. In this paper, a fast intra-prediction mode selection algorithm with different criterion based on RDO (Rate distortion optimization) is proposed using the mode selection relation of chroma block and luma block and the relation of luma  $4 \times 4$  block and luma  $16 \times 16$  block. Experiment results show that our proposed algorithm can averagely speed up to 6 times while maintaining similar PSNR and bit rate.

**Keywords** H.264/AVC, intra-prediction, mode selection

### 1 引言

H.264/AVC 作为新一代的国际视频编码标准已得到人们越来越多的关注。与编码标准 MPEG-2、H.263 和 MPEG-4 相比,虽然其总体结构仍为基于块的混合编码,但已包含许多新特征:如 VCL 层和 NAL 层分离、帧内预测、1/4 像素精度运动估计、自适应块尺寸运动补偿、多参考帧运动补偿预测、低复

杂 16bit 的整数变换和量化、环内去块滤波器和高效的熵编码等。这些新特征使得 H.264/AVC 能够显著提高编码效率且具有网络友好性,可有效用于各种网络和应用环境<sup>[1]</sup>。但也极大地增加了编码算法的复杂性,文献[2]比较了 JVT(joint video team)与 MPEG-4 的运算复杂性,JVT 在比 MPEG-4 提高约 50% 压缩效率的同时,增加了十几倍运算复杂性,这样就很难实时应用。

本文在深入分析 H.264/AVC 给出的帧内模式

基金项目:国家技术创新基金项目(01BK-272);上海市教育考试院基础研究项目(2003-18)

收稿日期:2005-08-16;改回日期:2005-09-06

第一作者简介:肖广(1969~),男。上海市教育考试院高级工程师,现为华东师范大学计算机系博士研究生。研究方向为多媒体通信。

E-mail:leoxiao@8163.net.cn

选择算法的基础上,充分利用色度和亮度块模式选择间的关系以及亮度宏块和  $4 \times 4$  块之间模式选择与出现概率的关系,提出了一种基于 RDO (rate distortion optimization) 的不同判决依据的快速帧内模式选择算法,从而使帧内编码速度提高数倍。

## 2 H. 264/AVC 帧内预测编码的模式选择

为了提高 I 帧图像的编码效率, H. 264/AVC 采用了帧内预测技术。虽然其宏块尺寸依然是  $16 \times 16$ , 然而由于其基本处理块单元相对于其他标准由  $8 \times 8$  降为  $4 \times 4$ , 因此亮度宏块支持  $4 \times 4$  (I4 预测块尺寸) 和  $16 \times 16$  (I16 预测块尺寸) 两种帧内预测块尺寸, 而色度仅支持  $8 \times 8$  (I8 预测块尺寸) 块尺寸。图 1 给出了  $4 \times 4$  块的帧内预测示意图, 包括 9 种预测模式, 除了图 1 (b) 所示的 0 ~ 1、3 ~ 8 等 8 种方向预测模式外, 2 为直流模式。图 1 (a) 中  $a \sim p$  的值可用相邻块的  $A \sim Q$  来预测。但亮度  $16 \times 16$  块和色度  $8 \times 8$  块仅支持垂直 (V)、水平 (H)、直流 (D) 和平坦模式 (P)。

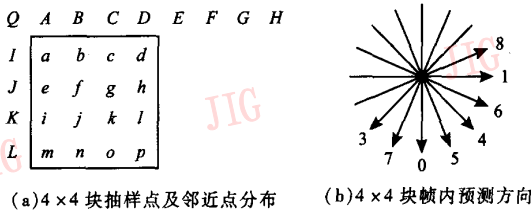


图 1  $4 \times 4$  块帧内预测  
Fig. 1 Intra prediction of  $4 \times 4$  block

为了保证在获得尽可能好的编码质量的同时, 能够将编码后的数据比特降至最小, H. 264/AVC 采用如式 (1) 所示的 RDO 判决<sup>[3]</sup>:

$$J_{\text{MODE}} = D_{\text{REC}} + \lambda_{\text{MODE}} R_{\text{REC}} \quad (1)$$

式中,  $D_{\text{REC}}$  表示编码后总的失真,  $R_{\text{REC}}$  表示编码后总的码率。  $\lambda_{\text{MODE}} \geq 0$  为拉格朗日参数, 然后通过确定编码模式 MODE, 就可使编码后的失真和码率总和最小。针对宏块帧内预测模式选择, MODE 为帧内预测模式的一种, 如果编码后总失真为  $D_{\text{REC}}$ , 编码后的总码率为  $R_{\text{REC}}$ , 则可通过确定 MODE 来使二者的和最小。  $D_{\text{REC}}$  可由式 (2) 给出的误差平方和 (sum of squared difference, SSD) 衡量:

$$SSD = \sum_{(x,y) \in A} |s[x,y,t] - \hat{s}[x,y,t]|^2 \quad (2)$$

式中,  $s$  为原始块像素,  $\hat{s}$  为重建块像素,  $A$  为当前块。式 (1) 中拉格朗日参数由量化参数  $Qp$  决定:

$$\lambda_{\text{MODE}} = 0.85 \cdot 2^{(Qp-12)/3} \quad (3)$$

$R_{\text{REC}}$  包括宏块类型、子块帧内预测模式 (I4 预测块尺寸情况下) 和残差等总的编码比特。 H. 264/AVC 为提高模式选择的编码效率, 在  $4 \times 4$  块编码时参考空间相邻块的编码模式 (如图 2 所示)。通过已编码邻块  $A$ 、 $B$  的模式可以推断当前块  $E$  的最可能模式, 如果  $E$  的实际模式与可能模式一致, 则仅需一比特确认, 如不一致, 再对它的实际模式进行编码。

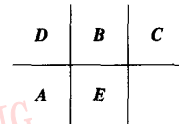


图 2 当前块的邻块位置  
Fig. 2 Neighbor block location of current block

校验模型  $Jm72^{[4]}$  给出了一种全搜索的算法, 它先以色度模式数为外循环, 然后依次扫描亮度的所有模式。这种算法能够扫描到所有模式, 因此编码质量最高, 运算量也是最大。其搜索的模式总数为  $4 \times (16 \times 9 + 4)$ , 共 592 个。这意味着为了得到一个宏块的最佳模式, 需要计算 592 个不同的  $J_{\text{MODE}}$ , 因此研究快速的帧内模式选择算法是提高编码速度所必需的。

下一节通过分析色度和亮度之间模式选择的关系、亮度宏块 I4 和 I16 模式选择的关系, 并由相邻块的相关性来联合确定整个宏块的模式判决, 这样不仅可使搜索范围实质上涵盖了全搜索, 而且在降低编码质量的同时, 可成倍提高 I 帧的编码速度。

## 3 基于 RDO 的快速模式选择算法

### 3.1 色度块和亮度块的关系

从上述的全搜索算法及其实验结果可以观察到, 由于亮度宏块无论选择 I16 预测块尺寸, 还是 I4 预测块尺寸, 色度宏块模式都是一致的, 因此色度预测模式可以独立判决, 并能够为亮度宏块模式判决提供信息。对于亮度 I16 预测块尺寸和色度 I8 预测块尺寸, 尽管二者模式选择不能达到完全一致, 但仍存在着联系: 当色度块选择 D 模式, 则亮度可以选择 D、H、V、P 模式; 当色度块选择 H 模式, 则亮度可以选择 D、H、P 模式; 当色度块选择 V 模式, 则亮

度可以选择 D、V、P 模式;当色度块选择 P 模式,则亮度可以选择 D、P 模式。而对于亮度 I4 预测块尺寸和色度 I8 预测块尺寸,色度 I8 的模式是亮度宏块 I4 预测块尺寸下绝大多数  $4 \times 4$  块可能选择的模式。

### 3.2 亮度宏块 I16 和 I4 的关系

对于 I16 和 I4 预测块尺寸, V、H、D 模式相对于其他模式具有优先选择性。I4 预测块尺寸情况下,通过相邻块的模式就可以判决得到当前块的最优先选择模式,其中与预测方向邻近的模式具有优先性。在实际的实现过程中,由于选择 I4 预测块尺寸的概率远远大于 I16 预测块尺寸,因此在判决过程上优先检测 I4 预测块尺寸。但由于 I16 和 I4 预测块尺寸随宏块纹理复杂度变化的 RDO 曲线(如图 3 所示)是不同的,因此在模式选择时,还应考虑宏块的纹理复杂度。

从图 3 可知,在纹理平坦区域, I16 预测块尺寸生成的重建图像能更好地与原始图像相匹配,故率失真较小,选择 I16 预测块尺寸概率大于 I4 预测块尺寸;但随着纹理复杂度的逐渐增加, I16 预测块尺寸在细节内容上率失真逐渐增加,到某个阈值  $T_1$  时大于 I4 的率失真,这时 I4 预测块尺寸更适合宏块编码模式选择。

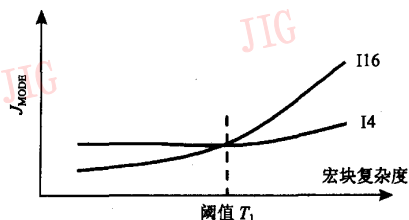


图 3 I16 和 I4 的率失真变化  
Fig. 3 RDO change of I16 and I4

在确定的 16 个  $4 \times 4$  块最佳模式中,由于概率最大的两种模式很可能与 I16 预测块尺寸的最佳模式一致,因此可以和色度模式一起检验,以便最终确定 I16 预测块尺寸的最佳模式。

### 3.3 RDO 判决准则

帧内预测模式判决准则应该遵守公式(1),为了减少运算的复杂性,可根据具体情况作相应的简化。在本算法中,由于色度模式选择不依赖亮度的模式选择,且其  $R_{\text{REC}}$  只包含色度模式和色度残差两部分,因此采用 SATD (sum of absolute transform

difference) 就可以反映不同模式间 RDO 判决  $J_{\text{MODE}}$  的差别。在亮度 I4 预测块尺寸下的  $4 \times 4$  块的 RDO 判决  $J_{\text{MODE}}$  中,  $R_{\text{REC}}$  包括块模式和残差的熵编码。因为  $4 \times 4$  块尺寸较小,其块模式编码的比特数和不同模式间的残差编码比特数的差异会导致  $J_{\text{MODE}}$  变化很大,故选择  $J_{\text{MODE}}$  作为 I4 预测块尺寸下的模式判决依据。由于  $4 \times 4$  块的  $J_{\text{MODE}}$  的分布比较集中,因此可在分布范围中定义一个阈值  $T_0$ 。如果在某一模式下,其  $J_{\text{MODE}}$  小于  $T_0$ ,则可认为是最佳模式。对于 I16 预测块尺寸,由于  $R_{\text{REC}}$  主要包括宏块模式和残差的熵编码,故 SATD 能够反映  $J_{\text{MODE}}$  的变化,并可以作为模式选择的依据。

如 3.2 节所述,对亮度 I4 预测块尺寸的检测应优先于 I16,因此当 I4 预测块尺寸情况下所有  $4 \times 4$  块的模式确定后,即可通过比较重建宏块和原始宏块来计算整个宏块的复杂度,若大于图 3 所示的阈值  $T_1$ ,则不必再进一步检测 I16 预测块尺寸;否则需进一步计算 I16 预测块尺寸的  $J_{\text{MODE}}$ ,并与 I4 的  $J_{\text{MODE}}$  比较,以便最终确定宏块的最佳模式。

### 3.4 快速选择算法

根据上述的各种关系,得到的帧内预测模式快速选择算法的步骤如下:

(1) 以 SATD 为判决依据来确定色度宏块帧内的最佳模式;

(2) 进行 I4 预测块尺寸的帧内预测,首先根据相邻块的模式确定可能模式,并将之和色度块的最佳模式相比较,以确定最有可能块模式,然后计算最可能模式的  $J_{\text{MODE}}$ ,并与阈值  $T_0$  比较,如果小于  $T_0$ ,即可确认为最佳模式,跳转到步骤 7,否则跳转到步骤 3;

(3) 计算另一模式(上面被排除的模式)的  $J_{\text{MODE}}$ ,首先与步骤 2 计算的  $J_{\text{MODE}}$  比较来得到最小  $J_{\text{MODE}}$  值的模式,然后再与  $T_0$  比较,如果小于  $T_0$ ,则确认为最佳模式,跳转到步骤 7;

(4) 选择最小  $J_{\text{MODE}}$  的模式作为候选模式,并计算与其预测方向相邻的两个模式的  $J_{\text{MODE}}$ ,同时比较这 3 种模式的  $J_{\text{MODE}}$ ,如果邻近模式的  $J_{\text{MODE}}$  值都大于候选模式的  $J_{\text{MODE}}$ ,则跳转步骤 7,否则到步骤 5;

(5) 如果最小  $J_{\text{MODE}}$  的模式是 V、H 或 D,则计算这 3 种模式在上述过程未曾计算的  $J_{\text{MODE}}$ ,同时比较得到最小  $J_{\text{MODE}}$  的模式,跳转到步骤 7,如果是其他模式,跳转到步骤 6;

(6) 计算其他剩余模式,若得到最小  $J_{\text{MODE}}$  模式,

跳转到步骤 7;

(7) 该最小  $J_{MODE}$  模式为该块的最佳模式;

(8) 循环步骤 2~7, 可得到 16 个  $4 \times 4$  块的最佳模式, 同时计算宏块复杂度, 如果大于阈值  $T_1$ , 则该宏块为 I4 预测块尺寸, 结束判决, 否则计算 I4 模式情况下整个宏块的  $J_{MODE}$ , 跳转到步骤 9;

(9) 根据色度的最佳模式, 确定 I16 预测块尺寸的可能模式;

(10) 统计 I4 预测块尺寸情况下 16 个子块的模式, 即可得到概率最大的两种模式, 如果属于 V、H 或 D, 则它们是宏块的可能模式; 如果为其他模式, 则对应于 I16 预测块尺寸的 P 模式;

(11) 比较上述确定的 I16 预测块尺寸可能模式, 即可得到 I16 预测块尺寸的最佳模式, 并计算其  $J_{MODE}$ , 然后与 I4 预测块尺寸的情况下的宏块最小  $J_{MODE}$  比较, 即可得到最佳块尺寸和其对应的最佳预测模式。

由上述的算法可知, 通常情况下需要检测的色度模式为 4 种, 亮度 I16 预测块尺寸最少有 2 种或最多 4 种, 亮度 I4 预测块尺寸的每个  $4 \times 4$  块最少有 1 种或最多 9 种, 虽然这几乎涵盖了所有的模式, 但总体检测的模式数极大地降低。

## 4 实验和本文算法结果

将这种快速算法在校验模型 Jm72 上进行实验, 测试了 3 个纹理复杂程度不同的 CIF 序列 Mobile、Bus 和 Paris。实验中熵编码采用了 CABAC, 量化参数  $Q$  分别为 28、32 和 36。然后观察在全搜索算法下 I 帧中亮度  $4 \times 4$  块的  $J_{MODE}$  和宏块复杂度的分布, 并将阈值  $T_0$ 、 $T_1$  固定为 200、10。为简化比较。

具体结果分别由表 1、表 2 和表 3 给出。观察该 3 个表的最右一列可见, 本文的快速帧内预测模式选择算法不仅比 H. 264/AVC 原有的算法可提高编码速度大约 6 倍多, 而且在速度提高的同时, 平均的信噪比只约降低 0.15dB 左右, 码率仅稍有变化。

表 1 Mobile 序列测试结果

Tab. 1 Results of mobile sequence

$Q_p$	比特率变化率 (%)	峰值信噪比变化 (dB)	速度变化 (倍)
14	2.5	-0.23	5.17
28	4.2	-0.19	5.51
36	5.7	-0.19	6.08

表 2 Bus 序列测试结果

Tab. 2 Results of bus sequence

$Q_p$	比特率变化率 (%)	峰值信噪比变化 (dB)	速度变化 (倍)
14	2.7	-0.17	5.45
28	4.3	-0.11	5.94
36	5.8	-0.09	6.64

表 3 Paris 序列测试结果

Tab. 3 Results of paris sequence

$Q_p$	比特率变化率 (%)	峰值信噪比变化 (dB)	速度变化 (倍)
14	2.4	-0.12	5.56
28	4.5	-0.09	5.93
36	6.7	-0.09	6.57

## 5 结论

本文通过对色度块和亮度块之间以及亮度 I16 和 I4 之间关系的分析, 指出了对色度模式、亮度 I4 和 I16 模式应使用不同的判据, 并在此基础上提出了用于 H. 264/AVC 帧内预测模式的一种基于 RDO 的快速模式选择算法。实验结果表明, 该算法能够成倍地提高编码速度。下一步拟将阈值  $T_0$ 、 $T_1$  随不同序列的复杂性变化而变化, 以便在保持编码质量的同时, 进一步提高编码速度。

## 参考文献 (References)

- Wiegand T, Sullivan G J, Bjøntegaard G, et al. Overview of the H. 264/AVC video coding standard [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 560 ~ 576.
- Ravasi M, Mattavelli M, Clerc C A. Computational Complexity Comparison of MPEG4 and JVT Codecs [S]. JVT-D153r1-L, Joint Video Team of ISO/IEC MPEG & ITU-T VCEG, Klagenfurt, Austria, 2002.
- Wiegand T, Schwarz H, Joch A, et al. Rate-constrained coder control and comparison of video coding standards [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 688 ~ 703.
- H. 264/AVC Reference software. Jm72 [DB/OL]. download from: <http://bs.hhi.de/suehring/tml/>.