

# H. 264 帧内 $4 \times 4$ 块预测模式选择快速算法研究

韩青 李莉 应骏

(上海师范大学通信工程系, 上海 200234)

**摘要** 在 H. 264 视频编码过程中, 编码时间受诸多因素的影响, 如帧间/帧内模式选择、运动估计 (ME)、率失真优化 (RDO) 等。为了以较快速度和较好质量进行编码, 针对 H. 264 帧内模式选择, 提出了一种适用于 H. 264 帧内  $4 \times 4$  块预测的模式选择快速算法。该算法利用帧内  $4 \times 4$  块最优预测模式与和它相邻的预测模式之间率失真代价 (RD Cost) 的高相关性, 以及绝对变换误差和 (SATD) 与率失真 (RD) 性能之间的强相关性, 有效地跳过一些不太可能的预测模式, 从而使帧内  $4 \times 4$  块模式选择过程只需进行 4 次率失真代价计算即可。实验结果显示, 该算法在编码性能和编码速度之间取得了很好的折衷。

**关键词** H. 264 帧内预测 视频编码

**中图分类号:** TN919.81 **文献标识码:** A **文章编号:** 1006-8961(2007)10-1745-04

## On Fast Intra Prediction Mode Decision Algorithm for $4 \times 4$ Blocks in H. 264

HAN Qing, LI Li, YING Jun

(Telecommunication Department, Shanghai Normal University, Shanghai 200234)

**Abstract** In the video-encoding process, the encoding time is greatly influenced by many factors, such as inter/intra mode decision, motion estimation (ME), etc. In order to encode the video with faster speed and better quality, this paper only focuses on the intra mode decision in H. 264 and presents a fast intra prediction mode decision algorithm for  $4 \times 4$  blocks in H. 264. The proposed algorithm uses: (i) high correlation of RD cost between best prediction mode and its neighborhood prediction mode, (ii) strong correlation between sums of absolute transform differences (SATD) and rate-distortion (RD) performance, efficiently skipping some unlikely prediction modes. Thus, it can complete  $4 \times 4$  intra prediction by only computing RD cost for four times. The simulation result shows that the proposed algorithm is a good compromise between the coding performance and the encoding speed.

**Keywords** H. 264, intra prediction, video coding

## 1 引言

为满足视频编码压缩效率高、网络友好性强的要求, ITU-T 视频编码专家组 (VCEG) 和 ISO/IEC 运动图像专家组 (MPEG) 组成了联合视频组 (JVT), 共同制定了最新的视频压缩编码国际标准 H. 264<sup>[1]</sup>。

类似于已往的视频编码标准, H. 264 的总体结果仍为基于块的混合编码, 但增加了许多新特性, 如率失真优化 (rate distortion optimization, RDO) 技术、 $4 \times 4$  整数变换、I 帧图像的帧内预测编码、 $1/4$

像素运动补偿、P 帧图像的多参考帧和多种块尺寸的编码、环内去块滤波和高效的熵编码等等<sup>[2]</sup>。这些新特性使 H. 264 比以往的视频压缩标准具有更高的压缩效率和更强的网络友好性, 但同时也增加了 H. 264 的编解码复杂度, 其在比 MPEG-4 提高约 50% 压缩效率的同时, 增加了十几倍的计算复杂度<sup>[3]</sup>, 因此为了满足实时应用的要求, 必须在其计算复杂度和编码效率之间取一个折衷。

帧内预测是 H. 264 计算复杂度高的主要因素之一。为了用最少的比特率获得最好的编码效果, 帧内预测采用全搜索方法, 通过计算每一种可能模

基金项目: 上海市教委基金项目 (06D2007)

收稿日期: 2007-06-01; 改回日期: 2007-08-01

第一作者简介: 韩青 (1982 ~ ), 男, 硕士研究生。主要从事视频编解码及无线通信研究。E-mail: 1000131774@smail.shnu.edu.cn

式的率失真代价 (rate distortion cost) 来选取最优的帧内预测模式, 但由于这大大增加了 H. 264 编码器的计算复杂度, 因此研究帧内预测模式选择的快速算法具有重要的应用价值。

目前, H. 264 帧内预测模式选择快速算法的研究可分为以下两类:

第 1 类是简化代价函数, 如文献 [4] 提出了一种改进的代价函数, 它利用整数变换绝对误差和 (sum of absolute integer transform difference, SAITD) 得到了更好的编码效果, 但其在编码效果提高的同时, 会比用绝对变换误差和 (sum of absolute transform difference, SATD) 计算带来更高的计算复杂度;

第 2 类是缩小模式选择范围, 如文献 [5] 提出了一种基于局部边缘方向信息的快速帧内模式选择算法。它先采用 Sobel 梯度算子计算当前块的边缘信息, 然后通过累加当前块中属于同一方向像素点的边缘矢量来构造不同模式下的边缘方向直方图 (edge direction histogram), 以便确定最可能的预测模式。然而, 帧内模式和边缘方向高相关性的假设并不总是成立, 这不但会额外增加比特开支, 并会降低编码图像的峰值信噪比 (peak signal noise ratio, PSNR)。在全 I 帧时, 该算法在 PSNR 值下降 0.27dB 的同时, 编码比特率增加了 4.0%。

本文提出了一种基于 H. 264 的帧内 4 × 4 块模式选择快速算法, 该算法仅需进行 4 次率失真代价计算即可完成模式选择。仿真结果表明, 该算法具有很好的性能。

## 2 H. 264 帧内 4 × 4 块预测模式选择

H. 264 采用了与 MPEG-4 不同的帧内预测方法, 它利用相邻像素的相关性, 采用了新的帧内预测模式, 最大程度地减少了图像的空间冗余信息。对于亮度 4 × 4 块共有 9 种预测模式, 其中模式 2 为直流 (DC) 预测模式 (如图 1 所示)。每种模式预测值的计算方法在 H. 264 标准中有详细的描述<sup>[1]</sup>。一个 4 × 4 块包含标记为  $p_{i,j}$  ( $i=1\sim 4, j=1\sim 4$ ) 的 16 个像素点, 像素 A ~ M 为邻近块的已编码的重构像素点 (如图 2 所示)。对于每个 4 × 4 块 (除了边缘块用 DC 预测), 每个像素都可用 A ~ M 的不同加权来预测。这种在空间域上进行的预测编码算法, 可以除去相邻块之间的空间冗余度, 可取得更为有效的压缩效率。

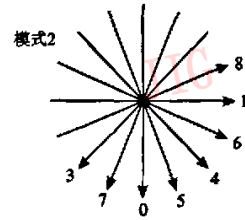


图 1 H. 264 帧内 4 × 4 预测模式方向

Fig. 1 Direction of intra 4 × 4 prediction modes in H. 264

M	A	B	C	D	E	F	G	H
I	$p_{1,1}$	$p_{1,2}$	$p_{1,3}$	$p_{1,4}$				
J	$p_{2,1}$	$p_{2,2}$	$p_{2,3}$	$p_{2,4}$				
K	$p_{3,1}$	$p_{3,2}$	$p_{3,3}$	$p_{3,4}$				
L	$p_{4,1}$	$p_{4,2}$	$p_{4,3}$	$p_{4,4}$				

图 2 帧内 4 × 4 块及其相邻像素

Fig. 2 A 4 × 4 block and its neighboring pixels

此外, 为提高所选模式的编码效率, 4 × 4 块在编码时应参考空间相邻块的编码模式 (如图 3 所示)。通过已编码邻块 A、B 的模式可推断当前块 C 的最可能模式, 如果 C 的实际模式与可能模式一致, 则仅需 1bit 确认, 如不一致, 再对它的实际模式进行编码。

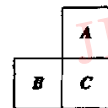


图 3 当前 4 × 4 块的邻块位置

Fig. 3 Adjacent blocks of current 4 × 4 block

在参考软件中<sup>[6]</sup>, 为了选取一个帧内 4 × 4 块的最优预测模式, 需要采用全搜索方法先分别计算 9 种可能模式的率失真代价, 然后选择率失真代价最小的那种模式作为最佳的帧内 4 × 4 块预测模式。

虽然全搜索算法能得到最优的预测模式, 但却是以高昂的计算复杂度为代价的。

## 3 帧内 4 × 4 预测模式选择快速算法

帧内 4 × 4 块的率失真计算公式如下:

$$J(s, c, Mode | QP, \lambda_{Mode}) = SSD(s, c, Mode | QP) + \lambda_{Mode} \cdot R(s, c, Mode | QP) \quad (1)$$

其中,  $QP$  为量化参数;  $\lambda_{Mode}$  为拉格朗日参数, 选取值为  $\lambda_{Mode} = 0.85 \times 2^{(QP-12)/3}$ ;  $SSD(\cdot)$  为当前 4×4 块亮度信号  $s$  与它的重构块  $c$  之间的差值平方和 (sum of squared differences, SSD)。  $R(s, c, Mode | QP)$  表示以选取的模式进行编码后的码率。

由于绝对变换误差和与率失真性能具有很强的相关性, 而且 SATD 的计算只需执行简单的加和移位操作, 所以可以利用 SATD 来检测最可能的预测模式<sup>[8]</sup>, 检测计算如下式所示:

$$MostProbableMode = \arg \min_i (SATD_i) \quad (2)$$

$$SATD = \sum |T\{I(x, y) - P(x, y)\}| \quad (3)$$

其中,  $I(x, y)$  和  $P(x, y)$  分别为当前块和相应的预测块点  $(x, y)$  的值。  $T(\cdot)$  表示哈达玛 (Hadamard) 变换。

虽然用 SATD 可以检测出最可能的预测模式, 但 SATD 最小的模式并不一定是率失真性能最好的模式。如何合理地利用 SATD 计算复杂度低的优点, 并克服 SATD 预测最优模式不精确的缺点, 是本文算法的关键。

通过观察发现, 与最优预测模式相邻的 2 个预测模式的率失真代价也不大, 即帧内 4×4 块最优预测模式与和它相邻的预测模式之间的率失真代价有很高的相关性。例如, 若最优模式为模式 0 (垂直), 则模式 5 (垂直向右) 和模式 7 (垂直向左) 的率失真代价也比较小; 若最优模式为模式 1 (水平), 则模式 6 (水平向下) 和模式 8 (水平向上) 的率失真代价也不大。那就意味着可以通过初始化搜索来跳过一些可能性不大的预测模式, 并以此来降低帧内 4×4 模式选择的计算复杂度。算法的主要步骤如下:

(1) 计算模式 0 (垂直)、模式 1 (水平) 这两种预测模式的率失真代价, 因为它们作为预测模式的可能性很高;

(2) 选取第 1 步中率失真代价较小的预测模式, 并计算与之相邻的两种预测模式的 SATD;

(3) 选取第 2 步中 SATD 较小的预测模式, 并计算与之相邻的余下的预测模式的 SATD;

(4) 选取第 2、第 3 步中 SATD 最小的预测模式, 并计算它的率失真代价;

(5) 计算模式 2 (直流) 的率失真代价;

(6) 比较第 1、第 4、第 5 步中计算出的率失真代价, 选取具有最小率失真代价的模式作为当前帧内 4×4 块的预测模式。

## 4 实验结果和分析

将本文提出的快速算法嵌入到参考软件 JM8.6 中<sup>[6]</sup>, 并与参考软件 JM8.6 进行比较测试。本文选取 Coastguard, Container, Stefan 等 3 个纹理复杂度不同的 CIF 视频序列做测试实验。实验中采用 CAVLC 进行熵编码, 量化参数 (quantization parameter, QP) 分别选为 28、32、36、40。对每个视频序列选取 100 帧, 以全 I 帧的编码方式进行编码。为了便于比较, 用  $\Delta PSNR$  表示相对于 JM8.6 全搜索算法亮度信噪比的变化, 用  $\Delta br$  表示码率变化百分比,  $\Delta t$  表示总的编码时间变化百分比, 其中正值表示增加, 负值表示减少。表 1 ~ 表 3 分别为 3 个测试序列的实验结果。图 4 ~ 图 6 分别为 3 个测试序列的 RD (rate distortion) 曲线图, 其中 NEW 代表本文提出的快速算法, ORG 代表参考软件 JM8.6 中的算法。

表 1 序列 Coastguard 的实验结果

Tab.1 Results of 'Coastguard'

QP	$\Delta br$ (%)	$\Delta PSNR$ (dB)	$\Delta t$ (%)
28	1.056	-0.07	-45.406
32	1.400	-0.04	-43.860
36	1.552	-0.02	-42.882
40	1.258	-0.01	-39.381

表 2 序列 Container 的实验结果

Tab.2 Results of 'Container'

QP	$\Delta br$ (%)	$\Delta PSNR$ (dB)	$\Delta t$ (%)
28	1.206	-0.06	-42.288
32	1.342	-0.06	-42.664
36	1.811	-0.07	-40.689
40	1.940	-0.06	-39.983

表 3 序列 Stefan 的实验结果

Tab.3 Results of 'Stefan'

QP	$\Delta br$ (%)	$\Delta PSNR$ (dB)	$\Delta t$ (%)
28	1.101	-0.15	-42.294
32	1.297	-0.14	-43.621
36	1.505	-0.13	-42.926
40	1.860	-0.08	-39.978

实验结果表明, 本文提出的帧内 4×4 预测模式

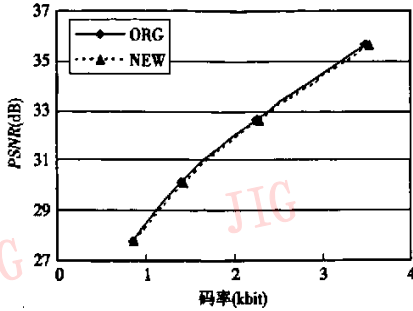


图 4 Coastguard 的 RD 曲线

Fig. 4 RD-curve of Coastguard

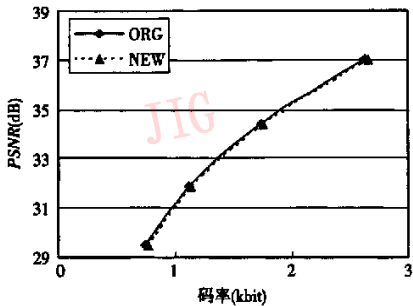


图 5 Container 的 RD 曲线

Fig. 5 RD-curve of Container

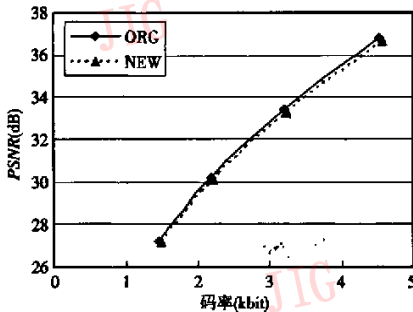


图 6 Stefan 的 RD 曲线

Fig. 6 RD-curve of Stefan

选择快速算法的编码速度不仅比 H. 264/AVC 原有算法平均提高了 42.4%，而且在编码图像峰值信噪比几乎不变的情况下，码率只增加 1.44% 左右。

从实验数据可见，无论是对于高运动特征，还是低运动特征的视频序列，码率的增加几乎相近，这是因为在帧内编码中只利用当前帧的信息，而不利用其他帧的信息。实验结果表明，随着量化参数的增加，编码时间减少的比例将逐渐降低，其原因是随着量化参数的增加，选择帧内  $16 \times 16$  预测块的可能性将不断增大，而本文是针对帧内  $4 \times 4$  预测模式选

择进行优化，而帧内  $16 \times 16$  预测模式选择还是使用参考软件中的全搜索算法。

## 5 结论

本文根据 H. 264 帧内  $4 \times 4$  块最优预测模式与它相邻的预测模式之间率失真代价的相关性，以及绝对变换误差和与率失真性能之间的相关性，提出了一种适用于 H. 264 帧内  $4 \times 4$  块预测的快速模式选择算法。该算法与全搜索算法相比，只需进行 4 次率失真代价运算，从而大幅度降低了视频编码帧内预测的计算复杂度。实验结果显示，本文提出的快速算法与 H. 264 参考软件相比，在编码图像峰值信噪比基本不变、码率只增加 1.44% 左右的情况下，编码速度平均提高了 42.4%。

## 参考文献 (References)

- 1 Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification [S]. ITU-T Rec. H. 264 | ISO/IEC 14496-10 AVC. JVT-G050, March, 2003.
- 2 Wiegand Thomas, Sullivan Gary J, Bjontegaard Gisle, 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.
- 3 Massimo Ravasi, Marco Mattavelli, Christophe Clerc. A Computational Complexity Comparison of MPEG-4 and JVT Codecs [S]. JVT-D153r1-L in ISO/IEC JTC1/SC29/WG11 and ITU-T SC16 Q.6, JVT 4<sup>th</sup> Meeting [C], Klagenfurt, Austria, 2002.
- 4 Tseng Chao-huang, Wang Hung-ming, Yang Jar-ferr. Improved and Fast Algorithms for Intra  $4 \times 4$  Mode Decision in H. 264/AVC [A]. In: Proceedings of IEEE International Symposium on Circuits and Systems [C/OL], <http://ieeexplore.ieee.org/iel5/9898131469/01465040.pdf?tp=&arnumber=1465040&iانumber=31469>, 2005, 3: 2128 ~ 2131.
- 5 Pan Feng, Lin Xiao, Susanto Rahardja, et al. Fast Mode Decision for Intra Prediction [S]. JVT-G013r1 in ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, JVT 7<sup>th</sup> Meeting [C], Pattaya II, Thailand, 2003.
- 6 Joint Video Team (JVT). Reference software JM8. 6 [S/OL]. <http://iphome.hhi.de/suehring/tml/download/>.
- 7 Kyeong Ho Yang, Jaecin Arnaud, Jayant Nikil S. A Normalized Rate-distortion Model for H. 263-compatible Codecs and Its Application to Quantizer Selection [A]. In: Proceedings of International Conference on Image Processing [C], Santa Barbara, CA, USA, 1997: 41 ~ 44.