

基于改进的 Canny 算子的 HEVC 帧内模式选择^①

史媛媛, 曹腾飞, 梁亚舒

(青海大学 计算机技术与应用系, 西宁 810016)

摘要: 针对 HEVC 中帧内预测模式选择算法复杂度问题, 提出了一种基于改进的 Canny 算子的 HEVC 帧内模式选择算法. 首先, 通过改进的 Canny 边缘检测算子提前分析预测块的方向, 将此方向相应的角度预测模式加入到候选模式集中; 其次, 利用图像空间相关性, 将相邻预测单元的最优帧内预测模式添加到候选模式集中, 最后, 通过 RDO 代价进行帧内模式选择. 实验结果表明, 与粗选择模式相比, 该算法将候选模式数目从 35 种减少到不超过 9 种, 算法编码时间平均减少 32%, 而 BD-Rate 损失仅为 0.096%, 并且编码后的主观图像质量基本没有变化.
关键词: HEVC; 帧内预测; 模式选择; 改进的 Canny 算子

Intra Prediction Mode Selection for HEVC Based on Canny Operator

SHI Yuan-Yuan, CAO Teng-Fei, LIANG Ya-Shu

(Department of Computer Technology and Applications, Qinghai University, Xining 810016, China)

Abstract: Intra-modes decision in HEVC has higher complexity. Aiming at this problem, after brief analysis of the intra prediction algorithm, an intra prediction method selection algorithm based on the improved Canny operator is proposed to reduce the HEVC intra-frame prediction complexity. Firstly, The improved Canny operator is used to obtain the direction of the Prediction Unites(PU) in advance, in order to select the corresponding angle prediction model to comprise the candidate list, and the spatial correlation of images is adopted to add the optimal mode of the neighboring PU to the candidate list. At Last, this algorithm takes full use of the advantages of rate distortion cost to select the candidate pattern quick descending law. Experiment proves that the proposed algorithm will reduce the number of candidate patterns from 35 to less than 10 compared with RMD, and effectively reduces the complexity of intra-coding. The average encode speed is accelerated about 32%, while BD-Rate loss was only 0.096% encoding time, and subjective image quality has no obvious change after encoding.

Key words: HEVC; intra-prediction; modes decision; the improved Canny operator

近些年来, 随着高清、超高清视频应用逐渐走进人们的视野, 视频编码技术受到了巨大的挑战. 为此, VCEG 和 MPEG 联合组建视频编码联合组 JCT-VC, 联合制定了新一代的视频编码标准—HEVC/H.265^[1]. HEVC 的核心目标是在 H.264 的基础上, 将压缩效率提高一倍, 即保证在相同编码图像质量的前提下, 视频流的码率减少 50%^[2]. 但 HEVC 在获得高性能的同时, 也无可避免地增加了计算复杂度, 在帧内编码过程中, 采用了二叉树递归编码单元结构以及采用枚举

法遍历 35 种帧内预测模式找出最佳模式^[3]. 然而随着帧内预测方向的大量增加, 不可避免地增加了编码复杂度. 为了更好地应用, 需要对帧内预测进行进一步处理. 帧内预测的精确性能够有效降低帧内编码的复杂度, 提高编码的效率. 近些年来, 针对帧内预测的高计算复杂度, 已有众多学者提出了改进的算法. 文献[4]改进预测单元模式决策算法, 利用三步搜索方法, 减少候选模式数量和帧内模式预测时间. 文献[5]使用 Sobel 算子计算出视频图像中纹理最强的方向,

① 基金项目:基金青海省科技厅应用基础研究项目(2014-ZJ-718);青海大学中青年基金(2014-QGY-23)

收稿时间:2015-12-19;收到修改稿时间:2016-01-27 [doi:10.15888/j.cnki.csa.005293]

在纹理最强方向和与该方向最接近的 2 个方向以及 DC 方向模式搜索最优的预测方向, 将 H.264 中的 9 种帧内预测方向减少为 4 种预测方向, 从而大量节省计算时间. 文献[6]提出基于相邻 PU 的加速帧内模式决策来减少粗选择时候选模式的数量以及提前终止编码单元的划分来降低帧内预测的复杂度. 文献[7]将率失真(RD)代价作为阈值参数, 利用候选模式集中预测模式被选中概率快速递减的规律, 基于 RD 代价进行帧内预测块划分和帧内预测模式选择. 文献[8]融合关联帧编码单元的深度信息及当前帧相邻编码单元的度信息, 从而预测当前编码单元的深度范围. 文献[9]采用一种基于 Canny 彩色算子的快速帧内预测模式选择算法, 避免遍历 35 种帧内预测模式, 从而降低了预测的复杂度. 但是现实场景受到光照、信息传输等因素的影响会引入一定的噪声. 同时, 一些重要的边缘细节会由于噪声干扰导致信息产生误差.

因此针对以上问题, 本文从图像的角度分析, 将视频分割成一帧帧的图像, 思考图像的基本特征, 即边缘信息. 边缘信息在图像分析、识别领域是一个经典的分析方法, 边缘检测的效果直接影响到图像纹理识别的性能. 在视频编码中考虑到图像的信息可以降低编码的复杂度, 提高效率. Canny 边缘检测算子^[10,11]是一种提取图像边缘信息效果较为良好的方法, 学者们也针对 Canny 算子进行了一定的优化^[12,13].

因此, 本文提出一种基于改进的 Canny 算子的 HEVC 帧内模式选择, 通过对图像边缘梯度信息以及空间域相关性的分析, 降低帧内预测的复杂度.

1 HEVC 帧内预测模式选择

HEVC 标准继承了 H.264/AVC 标准的很多基本要素, 其技术框架仍然基于以块为单元的预测加变换的混合编码框架, 但相比于 H.264 中的 9 种帧内预测模式, H.265 提供了 35 种帧内预测模式(如图 1 所示). 为了找出最佳模式, HEVC 不得不遍历这 35 种模式来计算每一个 PU 的率失真代价, 这样导致了高计算复杂度. 具体过程如下:

首先粗略的进行预测模式选择, 对 35 种角度预测模式使用基于 HAD (*Hadamard transform Absolute Difference*)的代价函数计算代价, 具体公式如下:

$$Cost = SATD + \lambda \times bits \quad (1)$$

其中, SATD 表示使用预测模式计算的残差平方和, λ

表示拉格朗日系数; bits 表示使用当前预测模式下进行编码的码率.

然后按代价值最小的前 N 种模式构成一组候选模式. 候选模式集中的候选模式数量如表 1 所示.

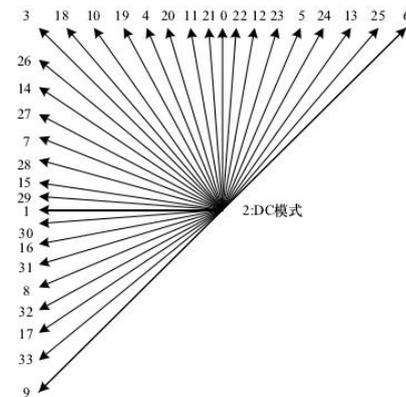


图 1 帧内模式预测方向

表 1 角度帧内预测方向模式和候选模式数量

| 预测单元大小 | 角度帧内预测方向模式 | 候选模式数量 |
|--------|------------|--------|
| 4×4 | 17 | 8 |
| 8×8 | 34 | 8 |
| 16×16 | 34 | 3 |
| 32×32 | 34 | 3 |
| 64×4 | 34 | 3 |

2 算法介绍

帧内预测模式计算运算量巨大. 帧内模式快速选择算法的研究作为提高编码效率的研究重点之一. 通过上述 HEVC 中较为耗时模块的分析, 降低帧内预测算法复杂度可以从以下 2 个方面进行研究:

1) 通过实验观察和统计, 视频是由众多的图像帧构成, 那么帧内预测的最佳模式也是与图像本身的信息有着密切的联系. 因此我们可以通过对图像物理边缘的梯度信息进行分析, 从而降低帧内预测的复杂度.

2) 又由于连续的帧构成视频, 那么视频具有空间相关性, 全方位的考虑到空间域, 可以提高帧内预测的精度. 下面将从以下 2 方面进行分析:

1) 通过实验观察和统计, 视频是由众多的图像帧构成, 那么帧内预测的最佳模式也是与图像本身的信息有着密切的联系. 因此我们可以通过对图像物理边缘的梯度信息进行分析, 从而降低帧内预测的复杂度.

2) 又由于连续的帧构成视频, 那么视频具有空间

相关性, 全方位的考虑到空间域, 可以提高帧内预测的精准度. 下面将从以下 2 方面进行分析:

2.1 改进的 Canny 算子获取图像边缘的梯度信息

在通常情况下, 图像是通过数字化处理的, 因此在其处理过程中必然存在一定的噪音信息. 由于图像边缘和噪声都是高频信号, 边缘信息很容易被噪声污染, 所以直接对含有噪声的图像运用梯度算子, 可能丢失重要的边缘信息, 影响接下来的梯度信息判断. 本文算法首先进行的是高斯滤波, 提取图像边缘信息, 其目的就是对原始图像进行平滑处理, 以去除或减弱图像中的噪声. 所以本文首先用式(2)对输入的图像进行高斯低通平滑处理.

$$G = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \tag{2}$$

σ 是高斯滤波函数的标准差.

接着用一阶偏导有限差分来计算梯度幅值方向. 但由于图像的离散性, 图像的灰度梯度值不能完全由 2×2 的领域有限差分, 并且受到噪音影响较大, 但在 3×3 领域中可以克服这些缺点. 下表 2 为 Canny 算子利用 3×3 领域的图像像素坐标, 图 2 为边缘检测的 4 个方向.

表 2 图像像素坐标

| | | |
|-------------|-----------|-------------|
| $[i-1,j]$ | $[i-1,j]$ | $[i-1,j+1]$ |
| $[i,j-1]$ | $[i,j]$ | $[i,j+1]$ |
| $[i+1,j-1]$ | $[i+1,j]$ | $[i+1,j+1]$ |

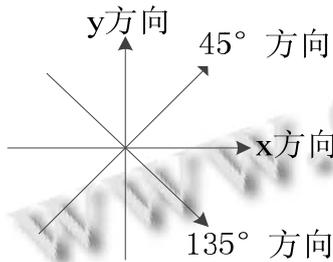


图 3 边缘检测的 4 个方向

下式中, $G_x(i,j)$ 和 $G_y(i,j)$ 是在点 (i,j) 处 x 方向和 y 方向的偏导数.

$$G_x(i,j) = I(i,j+1) - I(i,j-1) \tag{3}$$

计算 y 方向的偏导数:

$$G_y(i,j) = I(i+1,j) - I(i-1,j) \tag{4}$$

计算 45°方向的偏导数:

$$G_{45^\circ}(i,j) = I(i-1,j+1) - I(i+1,j-1) \tag{5}$$

计算 135°方向的偏导数:

$$G_{135^\circ}(i,j) = I(i+1,j+1) - I(i-1,j-1) \tag{6}$$

计算水平方向的差分:

$$G_x = G_x(i,j) + [G_{45^\circ}(i,j) + G_{135^\circ}(i,j)] / 2 \tag{7}$$

计算垂直方向的差分:

$$G_y = G_y(i,j) + [G_{45^\circ}(i,j) - G_{135^\circ}(i,j)] / 2 \tag{8}$$

则梯度幅值为:

$$G(i,j) = \sqrt{G_x^2 + G_y^2} \tag{9}$$

梯度方向为:

$$g(i,j) = \arctan \frac{G_x}{G_y} \tag{10}$$

得到当前像素点中的梯度信息之后, 用其来判断图像的预测角度. 利用逻辑的方法合并得到最后的图像的边缘梯度信息. 最终的梯度对应该像素点的最大梯度幅值的方向. 帧内预测模式与梯度的对应关系如表 3 所示. 假设 $0.599 \leq g(i,j) \leq 0.742$, 所以对应预测模式为模式 16. 表 3 中 Range 对应的是 $g(i,j)$.

表 3 梯度与帧内预测模式

| Mode | Angle | Range | Mode | Angle | Range |
|------|------------|------------------------------------|------|-------------|---------------------------------------|
| 2 | $8\pi/32$ | $(0.906, 1.103] \parallel (G > 0)$ | 19 | $-9\pi/32$ | $(-1.348, -1.103]$ |
| 3 | $7\pi/32$ | $(0.742, 0.906]$ | 20 | $-10\pi/32$ | $(-1.668, -1.348]$ |
| 4 | $6\pi/32$ | $(0.599, 0.742]$ | 21 | $-11\pi/32$ | $(-2.114, -1.668]$ |
| 5 | $5\pi/32$ | $(0.473, 0.599]$ | 22 | $-12\pi/32$ | $(-2.795, -2.114]$ |
| 6 | $4\pi/32$ | $(0.358, 0.473]$ | 23 | $-13\pi/32$ | $(-3.992, -2.795]$ |
| 7 | $3\pi/32$ | $(0.25, 0.358]$ | 24 | $-14\pi/32$ | $(-6.741, -3.992]$ |
| 8 | $2\pi/32$ | $(0.148, 0.25]$ | 25 | $-15\pi/32$ | $(-20.355, -6.741]$ |
| 9 | π | $0.049, 0.148]$ | 26 | $-16\pi/32$ | $(-1, -20.355] \parallel (20.355, 1)$ |
| 10 | 0 | $(-0.149, 0.149]$ | | | |
| 11 | $-\pi/32$ | $(-0.148, 0.049]$ | 27 | $-17\pi/32$ | $(6.741, 20.355]$ |
| 12 | $-2\pi/32$ | $(-0.25, -0.148]$ | 28 | $-18\pi/32$ | $(3.992, 6.741]$ |
| 13 | $-3\pi/32$ | $(0.599, 0.742]$ | 29 | $-19\pi/32$ | $(2.795, 3.992]$ |
| 14 | $-4\pi/32$ | $(0.906, 1.103]$ | 30 | $-20\pi/32$ | $(2.114, 2.795]$ |
| 15 | $-5\pi/32$ | $(0.742, 0.906]$ | 31 | $-21\pi/32$ | $(1.668, 2.114]$ |
| 16 | $-6\pi/32$ | $(0.599, 0.742]$ | 32 | $-22\pi/32$ | $(1.348, 1.668]$ |
| 17 | $-7\pi/32$ | $(0.599, 0.742]$ | 33 | $-23\pi/32$ | $(1.103, 1.348]$ |
| 18 | $-8\pi/32$ | $(0.599, 0.742]$ | 34 | $-24\pi/32$ | $(0.906, 1.103] \parallel (G \leq 0)$ |

由于求得的纹理方向存在一定误差, 可能会影响角度预测模式的选择, 为了提高这种方法的可靠性,

根据式(11)添加所选角度预测模式的相邻角度预测模式 M_1 和 M_2 到候选集中.

$$\begin{aligned} M_1 &= \left((D_{Canny} + 29) / 32 \right) + 2 \\ M_2 &= \left((D_{Canny} - 1) / 32 \right) + 2 \end{aligned} \quad (11)$$

2.2 基于空间相关性的帧内预测模式

视频由连续的图像序列帧组成, 具有空间相关性和时间相关性 2 种特征. 由于帧与帧之间时间间隔非常短, 前后帧之间存在着大量的冗余信息, 这种关系称为时间相关性; 一个像素通常与其周围的像素点存在关联, 这种特征称为空间相关性. 根据图像的空间相关性, 相邻 PU 块的最优帧内预测模式对当前 PU 块帧内预测模式的选取有重要的参考价值. 分别获取两个相邻 PU 块(如图 3 所示)的最优预测模式加入到当前 PU 块的最佳预测候选模式集中.

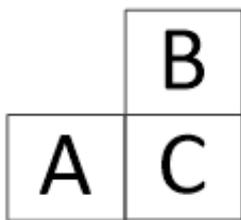


图 4 当前 PU 的相邻 PU 块

再将 4 种特殊预测模式—DC 模式(0)、平面模式(1)、水平模式(10)和垂直模式(26)加到最佳预测候选模式集中. 将候选模式中重复的模式删掉, 若此时候选模式集中模式有由图像物理边缘梯度信息获得的 3 个角度预测模式, 4 个特殊预测模式, 2 种相邻 PU 块的帧内预测模式, 这 9 种预测模式是互不重复的, 则最佳预测候选模式集中最多有 9 种帧内预测模式. 最后将最佳预测模式候选集进行率失真代价优化计算, 从中选出率失真代价最小的作为最佳预测模式.

2.3 帧内预测的详细算法步骤

整体算法描述(图 4)如下:

1) 对当前输入图像进行高斯低通滤波平滑处理, 得到平滑后的图像 G .

2) 利用改进的 *Canny* 梯度算子对平滑后的图像进行处理, 如公式(10)所示, 得到当前预测块的梯度的信息, 然后根据这个梯度信息, 在表 3 中找到预判的模式选择信息.

3) 选取所对应的角度预测模式 D_{canny} , 根据式

(11)增加相应角度预测模式 M_1 和 M_2 , 共同组成候选模式集;

4) 添加相邻 PU 块的最优预测模式, 以及 DC 模式(0), 平面模式(1), 水平模式(10)和垂直模式(26)到 RMD 候选模式列表中, 并对获得的候选模式集进行筛选, 去除其中重复的帧内预测模式;

5) 在得到的候选模式集中选出 N 个最优帧内预测模式, 用 RDO 代价模型来计算 N 个候选模式的代价信息, 选出代价最小的最佳预测模式.

2.4 实验结果与分析

为验证所提出方法的效率, 本文以 HEVC 的软件测试模型 *HM10.0* 为参照进行实验. 实验平台为 Inter(R)酷睿双核 CPU, 主频 2.60GHz, 内存 4.00GB, 操作系统 *Windows7*. 在此实验平台上对视频序列 *RaceHorse* 和 *BasketballPass* 进行验证, QP 量化参数为: 22, 27, 32, 37 和 42. *RaceHorse* 序列场景较为平滑, 活动量少; *BasketballPass* 序列场景比较复杂, 活动量大. 这两个视频序列对比鲜明, 特点突出, 对于算法的验证的效果比较好.

实验结果如图 5 所示.



(a)RaceHorse 原图



(b)改进的 Canny 算子效果图



(c)BasketballPass 原图



(d)改进的 Canny 算子效果图

图 5 实验结果

比较编码效率的参数指标有 PSNR 增量($\Delta PSNR$)、码率增量($\Delta Bitrate$)和编码时间增量($\Delta time$), 分别用以下公式进行计算:

$$\Delta PSNR = PSNR_{proposed} - PSNR_{HM10.0} \quad (12)$$

$$\Delta Bitrate = \frac{Bitrate_{proposed} - Bitrate_{HM10.0}}{Bitrate_{HM10.0}} \times 100\% \quad (13)$$

$$\Delta time = \frac{time_{proposed} - time_{HM10.0}}{time_{HM10.0}} \times 100\% \quad (14)$$

其中, $PSNR_{proposed}$ 表示本文算法编码后的视频峰值信噪比; $PSNR_{HM10.0}$ 表示采用 HM10.0 算法编码后的视频峰值信噪比; $Bitrate_{proposed}$ 表示采用算法编码后的视频的平均码率; $Bitrate_{HM10.0}$ 表示采用 HM10.0 算法编码后的视频的平均码率; $time_{proposed}$ 表示本文算法编码的编码时间; $time_{HM10.0}$ 表示采用 HM10.0 算法编码的编码时间.

表 4 BasketballPass 序列的编码性能

| QP | $\Delta PSNR/dB$ | $\Delta Bitrate/\%$ | $\Delta time/\%$ |
|----|------------------|---------------------|------------------|
| 22 | -0.025 | 0.080 | -27.0 |
| 27 | -0.025 | -0.030 | -29.0 |

| | | | |
|----|--------|-------|-------|
| 32 | -0.030 | 0.160 | -35.0 |
| 37 | -0.030 | 0.120 | -32.0 |
| 42 | -0.040 | 0.150 | -37.0 |
| 平均 | -0.030 | 0.096 | -32.0 |

表 5 RaceHorse 序列的编码性能

| QP | $\Delta PSNR/dB$ | $\Delta Bitrate/\%$ | $\Delta time/\%$ |
|----|------------------|---------------------|------------------|
| 22 | -0.020 | 0.200 | -26.0 |
| 27 | -0.020 | 0.260 | -24.0 |
| 32 | -0.030 | 0.290 | -29.0 |
| 37 | -0.040 | 0.260 | -30.0 |
| 42 | -0.050 | 0.300 | -28.0 |

3 结语

为降低算法遍历 HEVC 帧内 35 种预测模式的计算复杂度, 本文提出了一种基于改进的 Canny 算子边缘检测的快速帧内预测模式选择算法. 首先利用高斯低通函数平滑图像之后, 后经改进的 Canny 算子计算图像的梯度信息, 以判断当前预测块的物理边缘的方向, 再利用图像的空间相关性, 将相邻预测单元的最优帧内预测模式添加到候选模式集中, 本文算法将粗模式选择过程的候选模式数目从 35 种减少到不超过 9 种, 快速确定了当前块的预测方向, 减少了遍历帧内预测的模式数, 从而降低了帧内压缩算法的复杂度. 实验结果表明, 本文方法在提高率失真性能的同时还保持了主观图像质量.

参考文献

- Sullivan GJ, Ohm J, Han WJ, et al. Overview of the high efficiency video coding (HEVC) standard. IEEE Trans. Circuits Syst. Video Technol, 2012, 22(12): 1649-1668.
- Ohm J, Sullivan GJ, Schwarz H, et al. Comparison of the coding efficiency of video coding standards including high efficiency video coding (HEVC), IEEE Trans. Circuits Syst. Video Technol, 2012, 22(12): 1669-1684.
- Lainema J, Ugur K. Angular intra prediction in high efficiency video coding(HEVC). Proc. of the 13th IEEE International Workshop on Multimedia Signal Processing. Hangzhou, China: [s.n.]. 2011. 1-5.
- 雷海军, 杨忠明. HEVC 快速帧内模式决策算法. 计算机工程, 2014, 40(4): 270-276.
- Pan F, Lin X, Rahardja S, et al. Fast mode decision algorithm

- for intra prediction in H.264/AVC video coding. IEEE Trans. on Circuits and Systems for Video Technology, 2005, 15(7): 813–822.
- 6 Zhao L, Fan XP, Ma SW, Zhao DB. Fast intra-encoding algorithm for high efficiency video coding. Signal Processing: Image Communication, 2014, 29: 935–944.
- 7 张新晨,肖秀秀,赵娅,等.一种 HEVC 帧内预测快速算法.计算机工程,2014,40(10):17–24.
- 8 姚晓敏,王万良,岑跃峰.一种面向 HEVC 的编码单元深度决策算法.计算机工程,2015,41(1):240–244.
- 9 顾梅花,杨婷婷,刘新慧.HEVC 帧内彩色边缘检测快速模式选择算法.电子设计与工程,2015,23(8):132–135,139.
- 10 王植,贺赛先.一种基于 Canny 理论的自适应边缘检测方法.中国图象图形学报,2010,18(2):957–961.
- 11 赵宏中,张彦超.基于 Canny 边缘检测算子的图像检测算法.电子设计工程,2004,9(8):75–78.
- 12 雒涛,郑喜凤,丁铁夫.改进的自适应阈值 Canny 边缘检测.电光工程,2009,36(11):106–117.
- 13 李二森,张保明,周晓明.自适应 Canny 边缘检测算法研究.测绘科学,2008,33(6):109–121.