

基于 SAD 和 SATD 的 H.264 快速帧内预测算法

谢翠兰, 郑艺玲

(柳州职业技术学院信息工程系, 柳州 545006)

摘 要: H.264 在空域上进行多模式帧内预测, 用 RDO 准则选择最优模式, 提高了压缩率, 但是增加了计算复杂度。该文提出一种快速帧内预测算法, 利用 SAD 值对 Intra_4×4 方式进行预判, 根据 SATD 特征排除可能性小的 Intra_4×4 预测模式。实验结果表明, 该算法在保持峰值信噪比基本不变和码率略增的情况下, 可节省 69.3% 的帧内编码时间。

关键词: 视频编码; 帧内预测; 快速算法

Fast Intra Prediction Algorithm Based on SAD and SATD for H.264

XIE Cui-lan, ZHENG Yi-ling

(Department of Information Engineering, Liuzhou Vocational & Technical College, Liuzhou 545006)

【Abstract】 H.264 provides several predicting modes for intra prediction in the spatial region, and the optimal mode is selected by RDO technique. That makes compression performance improved, but the complexity and computation load increase drastically. A fast intra prediction algorithm is proposed in this paper. It utilizes SAD value to decide whether 4×4 intra prediction can be skipped, and uses the block's SATD to filter out some less probable modes in 4×4 intra prediction. Experimental results show that intra predicting time of the proposed algorithm is 69.3% less than that of H.264, with negligible loss of PSNR and increase of bite-rate.

【Key words】 video coding; intra prediction; fast algorithm

1 概述

H.264^[1]是由联合视频组JVT于2003年推出的新一代视频编码标准。它采用了许多新技术,与以往的视频编码标准相比,在相同恢复图像质量下,编码效率提高了一倍以上。由于具有高效的编码效率和良好的网络适配性,因此H.264有着非常广泛的市场应用前景。

基于空域的多模式帧内预测技术是H.264获得高编码效率的关键技术之一。H.264充分利用图像的空间相关性,用周围的像素以多种模式对待编码宏块进行帧内预测,并利用RDO技术选择最优帧内预测模式,从而降低了空域冗余,提高了视频编码的压缩效率。然而高效的帧内预测算法也大大增加了编码器的运算复杂度。在对帧内预测的运动估计使用快速算法后,帧内预测算法已经成为影响H.264编码速率的主要瓶颈之一。因此必须对H.264帧内预测过程进行优化,降低其运算复杂度,以满足实时应用的需求。

本文在充分挖掘Intra_16×16和Intra_4×4 2种预测方式之间的相关性及帧内预测模式与SATD相关性的基础上,提出了一种快速帧内预测算法,对Intra_4×4方式进行预判,并缩小Intra_4×4预测模式的选择范围。

2 H.264 帧内预测算法

2.1 H.264 帧内预测模式

H.264对亮度信息提供了2种帧内预测方式:基于4×4块的预测方式(Intra_4×4)和基于16×16宏块的预测方式(Intra_16×16)。色度信息则是基于8×8块进行预测的。

其中,Intra_4×4方式提供了9种预测模式,分别是垂直(Vertical)、水平(Horizontal)、均值(DC)、左对角(Digonal_Down_Left)、右对角(Digonal_Down_Right)、垂直偏右(Vertical_Right)、水平偏下(Horizontal_Down)、垂直偏左

(Vertical_Left)、水平偏上(Horizontal_Up)模式。

H.264还支持4种Intra_16×16预测模式,分别是垂直(Vertical)、水平(Horizontal)、直流(DC)、平面(Plane)模式。8×8色度块的预测模式与Intra_16×16的相似,只是模式的顺序不同。

2.2 H.264 帧内预测模式确定过程

H.264帧内编码中,采用亮度和色度相结合的率失真优化(Rate-Distortion Optimization, RDO)判优准则,RDO判决准则,通过遍历所有预测模式,寻找率失真代价RDCost最小的作为最优预测模式。率失真代价公式为

$$RDCost = SSD + \lambda \times Rate$$

其中,SSD表示当前块与重构块的平方差之和; λ 为量化参数QP的函数;Rate表示熵编码后的码率。

H.264中确定最优帧内预测模式的过程如下:

(1)确定Intra_16×16方式下的最优预测模式:

1)用4种预测模式对待编码宏块进行预测;
2)对预测残差宏块进行哈达马变换,并计算编码代价;
3)选择编码代价最小的模式作为Intra_16×16方式下的最优预测模式。

(2)结合色度编码,计算待编码宏块在Intra_16×16方式下的率失真代价RDCost₁₆。

(3)确定Intra_4×4方式下的最优预测模式。

将待编码宏块划分为16个4×4块,分别确定每个4×4块的最优预测模式:

1)用9种预测模式对4×4块进行预测;

作者简介: 谢翠兰(1973—),女,讲师,主研方向:计算机多媒体技术;郑艺玲,讲师

收稿日期: 2008-01-11 **E-mail:** xcl911@sina.com

2)对预测残差块进行 DCT 变换/量化、反 DCT 变换/反量化, 计算编码代价 $RDCost$;

3)选择 $RDCost$ 最小的模式为当前 4×4 块的最优预测模式。

(4)结合色度编码, 统计待编码宏块在 $Intra_4 \times 4$ 方式下的率失真代价 $RDCost_{t4}$ 。

(5)确定最优的帧内预测模式。

比较待编码宏块的 $RDCost_{t16}$ 和 $RDCost_{t4}$, 选择 $RDCost$ 小的作为待编码宏块最终的亮度编码方式, 该方式对应的预测模式即最优帧内预测模式。

由上述可知: 每一个宏块都要按照 2 种方式进行编码, 共需要进行 3 328 次预测、144 次解码/重构以及 68 次 4×4 哈达马变换, 可见其复杂性是相当高的。

3 基于 SAD 和 SATD 的快速帧内预测算法

近年来出现了许多降低帧内编码复杂度的快速算法。如文献[2]提出了基于方向场筛选预测模式及块方式提前中止策略; 文献[3]则根据各预测模式之间及与相邻块预测模式间的相关性缩小预测模式范围。这些算法在减少编码时间、保持编码性能方面各有优劣。

在吸取相关快速算法优点的基础上, 本文提出了一种基于绝对残差和(Sum of Absolute Differences, SAD)和变换域中的绝对残差和(Sum of Absolute Transformed Differences, SATD)的快速帧内预测算法来优化 H.264 帧内预测过程, 主要思路是:

(1)充分挖掘 $Intra_16 \times 16$ 和 $Intra_4 \times 4$ 2 种预测方式之间的相关性, 用待编码宏块在 $Intra_16 \times 16$ 方式下的 SAD 值来衡量宏块的平滑程度, 从而判定是否可以不执行 $Intra_4 \times 4$ 方式的预测过程。

文献[2]在对 $Intra_4 \times 4$ 方式的预判中使用了双阈值。该方法在统计值处于双阈值之间时, 采用正常的流程搜索最优预测模式, 并不会减少计算量。统计结果表明: 使用上阈值所引起的计算量的减少是很小的, 因此在本文提出的算法中只设置了下阈值, 并且为提高预测精度, 对不同范围的量化参数设置不同阈值。

(2)根据 SATD 特征对 $Intra_4 \times 4$ 方式的预测模式进行筛选, 选择出可能性最大的几种模式, 排除掉那些可能性小的模式, 避免不必要的 $RDCost$ 的计算, 从而降低计算复杂度。

3.1 快速算法基本原理

3.1.1 对 $Intra_4 \times 4$ 方式的预判

在 H.264 帧内编码过程中, 对亮度信息首先以 $Intra_16 \times 16$ 方式进行预测, 再以 $Intra_4 \times 4$ 方式进行预测, 两者之间是并行处理的, 完全忽略了这 2 种方式间的联系。

理论上, 当待编码宏块内部细节信息很多时, 采用 $Intra_4 \times 4$ 方式进行预测的效果会很好; 当待编码宏块充分平滑时, $Intra_16 \times 16$ 方式的编码效率要比 $Intra_4 \times 4$ 方式的好很多, 此时就没有必要再以 $Intra_4 \times 4$ 方式进行预测了。

可以用待编码宏块在 $Intra_16 \times 16$ 最优模式下的预测残差绝对值之和(最优预测残差的 SAD)衡量宏块是否平滑。SAD 定义为

$$SAD = \sum_{x,y \in MB} |PE(x,y) - Orig(x,y) - Pred(x,y)|$$

其中, $Orig(x,y)$ 为待编码宏块像素的原始值; $Pred(x,y)$ 为待编码宏块像素的预测值; $PE(x,y)$ 为待编码宏块的预测残差。

经统计可知, 当最优预测残差的 SAD 小于某值 $T1$ 时,

表示待编码宏块平滑, 可以直接选择 $Intra_16 \times 16$ 方式作为待编码宏块最终的亮度编码方式, 不必再以 $Intra_4 \times 4$ 方式进行预测, 从而提前中止模式选择过程。

阈值 $T1$ 的选取很关键, $T1$ 设置越大, 预测复杂度越小, 但会引起解码图像质量较大的下降和比特率的上升。阈值的选取应该在保证图像质量的前提下尽可能多地降低计算复杂度。同时 $T1$ 依赖于量化参数 QP , 随其增加而增加, 必须在不同 QP 范围下对 $T1$ 进行调整。

根据实验数据, $T1$ 采用 1 000 时, 在 QP 较大的情况下综合效果较好, 但随着 QP 的下降, 解码图像质量下降较大并且比特率普遍上升较多; $T1$ 采用 500 时, 可以改善较小 QP 时解码图像质量下降及比特率上升的问题。因此可设置 2 个 $T1$ 值, 当 $QP \leq 20$ 时, $T1=500$; 否则 $T1=1000$ 。

3.1.2 $Intra_4 \times 4$ 预测模式的筛选

H.264 帧内编码的压缩率与绝对残差和 SAD 是有很大关系的。如果用某预测模式进行预测时, 预测准确, 则其 SAD 应该很小。根据帕斯维尔定理, 预测残差在空域上的总能量与经过正交变换后在频域上的总能量是一致的。因此作为最优的预测模式, 其变换域中的绝对残差和 SATD 应该也是很小的。SATD 定义为

$$SATD = \sum_{x,y \in block} |T\{PE(x,y)\}|$$

其中, T 表示二维正交变换, 在这里采用哈达马变换。

与 SAD 相比, SATD 与 RDO 性能之间具有更强的相关性。由于前面已经采取了对 $Intra_4 \times 4$ 方式的预判方法, 对解码图像质量及比特率会略有影响, 因此为了提高预测准确度, 在筛选 $Intra_4 \times 4$ 预测模式时, 不再采用 SAD 而是采用 SATD 作为判断准则。

SATD 虽然可以作为选择模式的准则, 但是并不是 SATD 最小的模式就是性能最好的模式。这里以所有预测模式的 SATD 的平均值做为阈值 $T2$ 。选择 $SATD \leq T2$ 的预测模式为候选模式, 只对候选模式计算 $RDCost$, $RDCost$ 最小的模式即为 $Intra_4 \times 4$ 方式下的最优预测模式。

3.2 快速算法流程

该算法流程如下:

(1)确定 $Intra_16 \times 16$ 方式下的最优预测模式, 步骤同 H.264 原算法。

(2)结合色度编码, 计算待编码宏块在 $Intra_16 \times 16$ 方式下的率失真代价 $RDCost_{t16}$ 。

(3)计算最优 $Intra_16 \times 16$ 预测模式下待编码宏块的 SAD, 如果 $SAD < T1$ ($QP \leq 20$ 时, $T1=500$; 其余, $T1=1000$), 则跳到(6)。

(4)确定 $Intra_4 \times 4$ 方式下的最优预测模式。

将待编码宏块划分为 16 个 4×4 块, 分别确定每个 4×4 块的最优预测模式:

1)用 9 种预测模式对 4×4 块进行预测;

2)分别对每种模式对应的预测残差块进行哈达马变换, 并计算变换后的绝对残差和 SATD;

3)计算 SATD 的平均值 $SATD_{average}$;

4)选择满足 $SATD \leq SATD_{average}$ 条件的预测模式为候选模式;

5)分别对每种候选模式对应的预测残差块进行 DCT 变换/量化、反 DCT 变换/反量化, 并计算其编码代价 $RDCost$;

6)比较所有候选模式的 $RDCost$, 选择具有最小 $RDCost$

的预测模式为当前 4×4 块的最优预测模式。

(5)结合色度编码,统计待编码宏块在Intra_4×4 方式下的率失真代价 $RDCost_{I4}$ 。

(6)确定最优的帧内预测模式,步骤同 H.264 原算法。

4 实验结果

将快速算法应用于参考软件 JM86 上,在全 I 帧、RDO 模型下,分别选用 6 个量化参数 $QP(10, 16, 24, 28, 32, 36)$,对 QCIF, CIF 共 6 个序列各 30 帧进行测试。

表 1~表 3 列出 6 个视频序列采用快速算法时与原算法的比较数据。其中, $\Delta PSNR$ 表示峰值信噪比的变化; $\Delta Bits$ 表示比特率的相对变化; $\Delta Time$ 表示编码时间的相对变化。

表 1 QCIF 序列快速算法与原算法的比较数据

| 序列名 | 参数 | QP | | | | | |
|----------|-------------------|-------|-------|-------|-------|-------|-------|
| | | 10 | 16 | 24 | 28 | 32 | 36 |
| ackiyv | $\Delta PSNR/dB$ | -0.18 | -0.09 | -0.03 | -0.07 | +0.06 | +0.02 |
| | $\Delta Bits/(%)$ | +1.6 | +1.5 | +2.5 | +2.1 | +2.1 | +2.1 |
| | $\Delta Time/(%)$ | -69.6 | -68.8 | -70.2 | -69.0 | -67.4 | -64.9 |
| carphone | $\Delta PSNR/dB$ | -0.12 | -0.06 | -0.03 | -0.04 | 0.00 | +0.01 |
| | $\Delta Bits/(%)$ | +0.9 | +1.3 | +2.2 | +2.0 | +2.4 | +2.2 |
| | $\Delta Time/(%)$ | -70.5 | -70.3 | -71.4 | -70.1 | -69.1 | -67.3 |
| foreman | $\Delta PSNR/dB$ | -0.12 | -0.05 | -0.01 | -0.01 | 0.00 | +0.02 |
| | $\Delta Bits/(%)$ | +0.2 | +0.6 | +1.1 | +1.3 | +1.8 | +2.3 |
| | $\Delta Time/(%)$ | -69.9 | -70.0 | -69.4 | -69.0 | -68.0 | -66.6 |

表 2 CIF 序列快速算法与原算法的比较数据

| 序列名 | 参数 | QP | | | | | |
|----------|-------------------|-------|-------|-------|-------|-------|-------|
| | | 10 | 16 | 24 | 28 | 32 | 36 |
| football | $\Delta PSNR/dB$ | -0.11 | -0.05 | -0.04 | -0.01 | -0.01 | +0.02 |
| | $\Delta Bits/(%)$ | +0.4 | +0.8 | +1.5 | +1.7 | +1.9 | +2.3 |
| | $\Delta Time/(%)$ | -69.6 | -69.4 | -70.0 | -68.9 | -67.8 | -66.2 |
| news | $\Delta PSNR/dB$ | -0.26 | -0.08 | -0.04 | -0.02 | +0.01 | +0.02 |
| | $\Delta Bits/(%)$ | +1.2 | +2.3 | +3.4 | +3.1 | +2.7 | +2.6 |
| | $\Delta Time/(%)$ | -71.5 | -69.1 | -71.8 | -71.3 | -67.9 | -68.7 |
| stefan | $\Delta PSNR/dB$ | -0.25 | -0.11 | -0.07 | -0.05 | -0.04 | -0.03 |
| | $\Delta Bits/(%)$ | +1.1 | +1.4 | +1.6 | +1.6 | +1.5 | +1.6 |
| | $\Delta Time/(%)$ | -70.6 | -70.7 | -70.7 | -70.2 | -69.5 | -68.9 |

表 3 各序列在 6 个 QP 下的平均数据

| 序列名 | $\Delta PSNR/dB$ | $\Delta Bits/(%)$ | $\Delta Time/(%)$ |
|----------|------------------|-------------------|-------------------|
| ackiyv | -0.05 | +2.0 | -68.3 |
| carphone | -0.04 | +1.8 | -69.8 |
| foreman | -0.03 | +1.2 | -68.6 |
| football | -0.03 | +1.4 | -68.7 |
| news | -0.06 | +2.5 | -70.1 |
| stefan | -0.09 | +1.5 | -70.1 |
| 平均 | -0.05 | +1.7 | -69.3 |

从表 1~表 3 中可以看出快速算法所用的编码时间比 H.264 所用的时间少得多,在 QP 变化范围很大(10~36)的情况下,编码时间平均下降了 69.3%。编码时间可以衡量 2 种算法的复杂性。因此与 H.264 帧内预测算法相比,本文提出

的帧内预测快速算法可以较大幅度地降低计算复杂性。

图 1 绘出了 foreman 和 football 序列 2 种算法下的编码性能比较图。

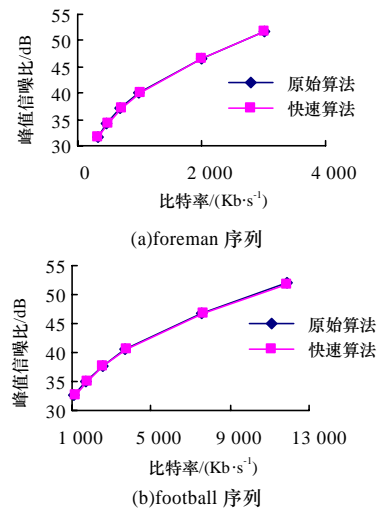


图 1 率失真曲线比较图

与 H.264 相比,快速算法的峰值信噪比平均仅下降了 0.05 dB,总体上保证了解码图像质量。在图 1 中,几乎看不出快速算法编码性能上的降低。由于快速算法在是否提前中止 Intra_4×4 方式的判断中,根据不同的 QP 范围采用了不同的阈值,从而在编码速率与编码性能之间总体上得到了较好的平衡。

采用快速算法,比特率平均上升了 1.7%。这是快速算法为减少 69.3%编码时间所付出的微小代价。

5 结束语

本文在重点分析 H.264 帧内预测算法基础上,提出了一种基于 SAD 和 SATD 的快速帧内预测算法。该算法与原算法相比,在基本保持解码图像质量不变、码率略有增加的情况下,大幅度地降低了运算复杂度,提高了编码速度。

参考文献

- [1] Joint Video Team. JVT-G050 Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification[S]. 2003.
- [2] 黄慧娟,李 挥.基于方向场和中止模式的 H.264 帧内模式选择方法[J].计算机应用,2006,26(2): 167-169.
- [3] 周绪川,蔡利平.一种 H.264/AVC 中帧内快速预测方法研究[J].微计算机信息,2007,23(3/4): 253-255.
- [4] Moravec H P. Towards Automatic Visual Obstacle Avoidance[C]// Proceedings of International Joint Conference on Artificial Intelligence. Cambridge, MA, USA: William Kaufmann,1977.
- [5] Stephen M, Michael J. A New Approach to Low-level Image Processing[J]. International Journal of Computer Vision, 1997, 23(1): 45-78.
- [6] Schmid C, Mohr R, Bauckhage C. Evaluation of Interesting Point Detectors[J]. International Journal of Computer Vision, 2000, 37(2): 151-172.
- [7] 张登荣,刘辅兵,俞 乐,等.基于 Harris 操作数的遥感影像自适应特征提取方法[J].国土资源遥感,2006,6(2): 35-38.

(上接第 214 页)

- [4] Moravec H P. Towards Automatic Visual Obstacle Avoidance[C]// Proceedings of International Joint Conference on Artificial Intelligence. Cambridge, MA, USA: William Kaufmann,1977.
- [5] Stephen M, Michael J. A New Approach to Low-level Image Processing[J]. International Journal of Computer Vision, 1997,