

文章编号:1001 - 9081(2005)08 - 1808 - 03

H. 264/AVC 帧内预测模式选择算法研究

裴世保, 李厚强, 俞能海

(中国科学技术大学 电子工程与信息科学系, 安徽 合肥 230026)

(psb@ustc.edu)

摘要: H. 264/AVC 采用空间域上的帧内预测技术, 进一步提高了编码效率, 但由于 H. 264/AVC 支持的帧内预测模式数较多, 使预测的复杂度大幅度增加。详细分析了帧内预测模式的选择过程, 提出一种快速的率失真优化(RDO)模式下的快速 Intra_4 × 4 模式选择算法, 该算法根据 SATD (Sum of Absolute Transformed Difference) 以及相邻块的预测模式之间的相关性等特征, 预先排除了超过 65% 可能性小的 Intra_4 × 4 模式, 避免了不必要的计算, 从而大幅度降低帧内预测的复杂度, 同时基本保持了 H. 264/AVC 的编码性能。

关键词: 视频编码; 帧内预测; 率失真优化

中图分类号: TP317.4; TN919.8 **文献标识码:** A

Research of H. 264/AVC intra-prediction mode selection algorithm

PEI Shi-bao, LI Hou-qiang, YU Neng-hai

(Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei Anhui 230026, China)

Abstract: H. 264/AVC employed intra-prediction technique in space domain to enhance coding efficiency, but it increased prediction complexity drastically because of quite a few intra-prediction modes. A new Intra_4 × 4 mode selection algorithm based on rate distortion optimization (RDO) was proposed. It employed SATD (Sum of Absolute Transform Difference) of each block as a judgment and used correlation between prediction modes of adjacent blocks to filter out more than 65% less probable modes beforehand so as to avoid unnecessary computations. The new algorithm reduces the intra-prediction complexity greatly while maintaining coding performance very well.

Key words: video coding; intra-prediction; RDO(Rate Distortion Optimization)

0 引言

H. 264/AVC 是 2003 年 5 月 ITU-T 正式批准的新一代视频编码标准, 与以前的视频编码标准相比, 该标准的显著优点是压缩效率高、网络性能好, 适用于交互和非交互应用环境, 现已受到越来越广泛的关注。H. 264/AVC 性能的提高与它采用的一些新技术密不可分, 如新的帧内预测方法、整数变换、4 × 4 块的运动估计/运动补偿、环路滤波、新的熵编码方法、率失真优化技术(RDO)等。H. 264/AVC 在性能提高的同时, 编码器的复杂度大幅度增加, 很难应用于实时性要求较强的场合, 因而研究有关快速算法具有重要意义。

基于空间域的帧内预测是 H. 264/AVC 采用的一项重要新技术, 它充分利用了图像的空间相关性, 根据已解码的相邻块的信息来预测当前块的信息, 对进一步提高编码的性能具有重要作用。H. 264/AVC 采用 RDO 技术进行帧内预测模式选择, 用最少的比特率获得了最好的编码效果, 但由于 H. 264/AVC 支持的帧内预测模式很多, 为了确定一个宏块 (Macroblock, MB) 的帧内预测模式, 需要计算 592 种组合模式的 RD 代价 (RD_Cost), 所以编码器的复杂度很高^[1]。从帧内预测的重要性和预测的复杂度可以看出, 研究 RDO 模式下的快速帧内预测模式选择算法具有重要的应用价值。我们在文献[2]中提出了一种简单的快速帧内预测模式选择算法, 本文在此基础上对帧内预测进行了深入研究, 并对算法进行

了改进和完善, 取得了显著效果。新算法能预先排除超过 65% 可能性小的 Intra_4 × 4 模式, 大幅度降低了帧内预测的复杂度, 并且基本保持了 H. 264/AVC 的编码性能, 具有更好的适用性。

1 H. 264/AVC 的帧内预测模式^[3]

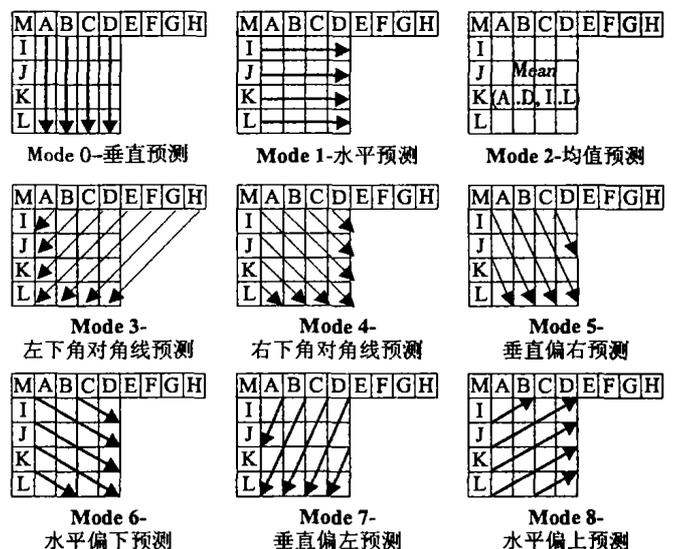


图1 Intra_4 × 4 的 9 种预测模式

收稿日期:2005 - 01 - 16; 修订日期:2005 - 04 - 15 基金项目:国家自然科学基金资助项目(60333020)

作者简介:裴世保(1972 -),男,安徽六安人,硕士研究生,主要研究方向:视频处理与通信; 李厚强(1969 -),男,安徽寿县人,副教授,主要研究方向:图像处理、视频处理与通信; 俞能海(1964 -),男,安徽无为,教授,主要研究方向:视频处理与多媒体通信、信息隐藏与信息安全。

H. 264/AVC 共支持三类帧内预测: 第一类 Intra_4 × 4 是针对 4 × 4 亮度块的预测, 共有 9 种预测模式, 其中包括 1 种均值预测模式和 8 种具有一定方向性的预测模式; 第二类 Intra_16 × 16 是针对 16 × 16 亮度块的预测, 共有 4 种预测模式; 第三类 Intra_8 × 8 是针对 8 × 8 色度块的预测, 共有 4 种预测模式。Intra_4 × 4 模式由于块尺寸较小, 比较适合于编码细节丰富的图像, 且能获得较小的差值, 但同时由于 Intra_4 × 4 的模式数较多, 所以成为增加帧内编码复杂度的一个重要原因。Intra_16 × 16 比较适用于编码图像中的平滑区域。

2 帧内预测的分析

2.1 相邻块之间的相关性

对于解码器来说, 除了要知道差值外, 还要知道所用的预测模式才能重建出当前块。如果对于每个块均用一个值来标记其预测模式, 这样将花费大量的比特来编码预测模式, 结果导致码率的增加。实际上, 相邻的 4 × 4 块的预测模式之间具有很强的相关性, 根据当前 4 × 4 块的左边块和上边块的预测模式可以预测当前块的最可能模式 (Most Probable Mode, MPM)。利用该相关性可以减少编码的比特数。

2.2 采用 RDO 技术的帧内预测模式选择过程

在帧内预测模式下, 为了确定一个宏块的帧内预测模式, 要执行 592 种 RD_Cost 计算。由于 RD_Cost 的计算量相当大, 导致帧内预测的复杂度很大。

宏块的帧内预测模式确定过程^[4-5]如下:

1) 分别计算 9 种 Intra_4 × 4 模式的代价 (RD_Cost), 选择具有最小代价的模式。代价函数 $RD_Cost = SSD + \lambda_{mode} \times Rate$, 其中 SSD (Sum of Squared Difference) 表示当前块与重建块的差值的平方和; $\lambda_{mode} = 0.85 \times 2^{(QP-12)/3}$ (QP 为量化参数); Rate 表示熵编码后的码率。

2) 把 16 个 4 × 4 块最小的 RD_Cost 相加得到当前宏块 Intra_4 × 4 方式下的 RD_Cost。

3) 分别计算 4 种 Intra_16 × 16 模式的 SAD (Sum of Absolute Difference), 选择具有最小 SAD 的模式。计算该模式下的 RD_Cost, 得到当前宏块 Intra_16 × 16 方式下的 RD_Cost。

4) 比较前两步获得的 RD_Cost, 选择具有最小 RD_Cost 的模式作为该宏块的帧内预测模式。

2.3 减小帧内预测复杂度的常用方法

一般来说, 减小帧内预测复杂度的方法可分为两类: 一类是简化代价函数; 另一类是缩小预测模式选择的范围。对于第二类方法, 可以利用当前块及其周围像素的某些特征, 预先排除某些可能性很小的预测模式, 或提前终止某些可能性小的模式的代价计算, 从而降低帧内预测的复杂度。为了获取图像的某些特征, 可以直接在空间域上进行分析, 也可以在变换域上进行分析。文献[1]直接利用空间域的特征, 根据边界方向直方图预先排除一些可能性小的预测模式, 从而减小复杂度。该方法的缺点是额外增加了比特开支。文献[6]利用变换域和空间域的联合特征预先排除大部分可能性小的模式, 从而减小复杂度。该算法虽然大幅度降低了复杂度, 但实现起来较麻烦, 且编码性能有一定的损失, 在不同码率的情况下, 复杂度降低程度相差较大。

3 Intra_4 × 4 模式选择的快速算法

编码器对预测的差值和预测模式进行编码, 也就是说预测差值和模式选择直接影响编码的效率和质量。我们对预测

差值进行 Hadamard 变换, 然后计算变换系数的绝对和得到 SATD (Sum of Absolute Transformed Difference)。由于 SATD 与率失真 (RD) 性能具有很强的相关性^[6,7], 所以我们利用 SATD 作为判断准则, 预先排除掉一些可能性很小的 Intra_4 × 4 预测模式, 缩小模式选择的范围, 从而降低复杂度。

SATD 虽然可以作为选择模式的准则, 但是并不是说 SATD 最小的模式就是 RD 性能最好的模式, 如何选择合适的阈值, 是算法的关键。因为相邻的 4 × 4 块的预测模式之间具有很强的相关性, 当前块的最可能预测模式 MPM 可以根据其左边和上边块的模式来预测, 所以我们选择当前块的最可能预测模式 MPM 的 SATD 作为阈值, 预先排除掉 SATD 大于该阈值的模式。大量实验结果显示, 该阈值具有较好的效果, 可以预先排除约 50% 的 Intra_4 × 4 模式。

为了排除掉更多的模式, 进一步减小复杂度, 我们对实验结果进一步分析, 发现主要存在两方面的问题: 1) 在少数情况下, MPM 预测得不准, 只能预先排除约 30% 的模式, 甚至更少。2) 当 QP 值较大时, 算法的性能下降, 其原因是当 QP 值较大时, 大多数预测模式具有相同的 SATD 值, 所以用 MPM 的 SATD 值作为阈值很难排除掉大部分模式。

对于第一种情况, 我们对阈值进行了调整, 用 9 种预测模式的 SATD 的均值作为阈值, 实验结果显示比用 MPM 的 SATD 值作为阈值效果要好得多。对于第二种情况, 通过对大量的 SATD 值进行分析, 我们发现当 MPM 和其他几种模式都具有最小的 SATD 时, 使用 RDO 选择的最优模式就是 MPM。实际上根据代价函数 $RD_Cost = SSD + \lambda_{mode} * Rate$ 也可以进行直观分析, 当 QP 较大时, 如果几种模式的 SATD 相同, 那么它们对应的失真 (SSD) 基本相同, 而如果这几种模式中包括 MPM, 那么该 MPM 对应的 Rate 肯定最小, 因而对应的代价函数值也最小。

实验结果显示, 该算法简单有效, 可以预先排除超过 65% 的 Intra_4 × 4 模式, 大幅度降低帧内预测的复杂度, 同时基本保持了 H. 264/AVC 的编码性能, 并且算法的性能受 QP 值的影响较小。算法的主要步骤如下:

- 1) 计算当前块的最可能预测模式 MPM (Most_Probable_Mode);
- 2) 分别计算 9 种模式的 SATD[k];
- 3) 根据上述的阈值条件预先排除可能性较小的预测模式;
- 4) 在剩下的候选模式中, 分别计算每种模式的 RD 代价 (RD_Cost);
- 5) 选择具有最小 RD 代价的模式, 即为当前 4 × 4 块的帧内预测模式。

4 实验结果

我们将该快速算法应用于 H. 264/AVC JM7.3 参考软件中, 在全 I 帧、RDO 模式下, 分别测试了不同种类的 QCIF 格式的视频序列。表 1 列出了 News, Stefan, Bus, Tempete 四个 QCIF 视频序列在快速算法下的实验结果 (表中 BR 代表比特率, 单位 Kbit/s; PSNR 代表峰值信噪比, 单位 dB; RDC 代表平均每帧中 4 × 4 块 RD_Cost 函数的调用次数; Time 代表整个序列的编码时间, 单位: 秒); 表 2 列出了对应的四个视频序列在 H. 264/AVC 原来的模式选择算法下的实验结果。为了更直观显示实验结果, 图 2 为 Tempete 和 Bus 两个 QCIF 视频序列的性能曲线。图 3 为 RDC、编码时间减少的百分比曲线。实验结果显示, 采用该算法平均可减少序列的编码时间 30%

以上,同时基本保持了 H. 264/AVC 的编码性能,并且没有额外增加比特率。从 4 × 4 块 RD_Cost 函数的调用次数可以更

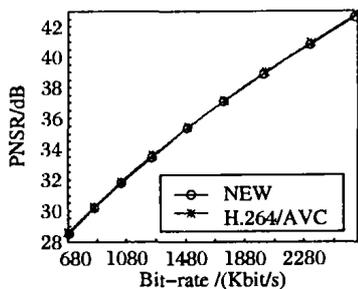
清楚地看出,使用快速算法比 H. 264/AVC 原算法平均减少 65% 以上的函数调用,并且算法的性能受 QP 值的影响较小。

表 1 快速算法的实验结果

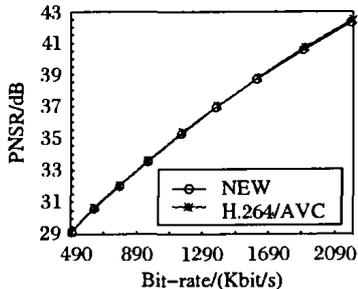
QP	News [300 frames]				Stefan [300 frames]				Bus [150 frames]				Tempete [250 frames]			
	BR	PSNR	RDC	Time	BR	PSNR	RDC	Time	BR	PSNR	RDC	Time	BR	PSNR	RDC	Time
20	1462.96	43.90	15388	85.432	2509.94	43.11	15976	101.380	2185.44	42.35	16219	51.013	2621.14	42.60	16667	89.291
22	1267.84	42.36	15337	82.346	2229.52	41.36	15959	97.371	1902.04	40.57	16211	48.762	2317.60	40.79	16615	85.938
24	1084.37	40.72	15279	78.805	1940.64	39.47	15892	92.613	1618.56	38.69	16226	46.646	2005.91	38.86	16609	82.257
26	926.78	39.19	15424	76.095	1688.25	37.66	15887	88.803	1374.63	36.94	16182	44.425	1733.16	37.06	16587	78.801
28	794.32	37.72	15390	73.554	1460.51	35.93	15851	85.475	1162.77	35.29	16165	42.729	1489.73	35.34	16569	75.770
30	669.33	36.07	15016	69.793	1241.88	33.99	15807	81.610	955.54	33.57	16151	40.790	1254.21	33.47	16563	72.365
32	566.57	34.53	14910	66.962	1038.72	32.24	15822	78.319	781.49	32.02	16072	38.809	1041.75	31.78	16570	69.250
34	477.01	33.15	14872	65.243	865.05	30.56	15777	74.801	636.15	30.61	15943	37.211	863.95	30.20	16608	66.589
36	386.50	31.53	14558	62.185	685.92	28.83	15840	71.472	498.59	29.13	15800	35.564	684.27	28.56	16647	63.492

表 2 H. 264/AVC 原算法的实验结果

QP	News [300 frames]				Stefan [300 frames]				Bus [150 frames]				Tempete [250 frames]			
	BR	PSNR	RDC	Time	BR	PSNR	RDC	Time	BR	PSNR	RDC	Time	BR	PSNR	RDC	Time
20	1465.20	43.98	50567	135.383	2520.36	43.30	50567	159.873	2199.03	42.52	50567	80.763	2635.6	42.80	50567	140.569
22	1270.35	42.41	50567	128.792	2239.03	41.54	50567	151.350	1913.19	40.72	50567	76.297	2329.56	40.96	50567	133.690
24	1084.29	40.79	50567	122.251	1949.47	39.64	50567	143.399	1627.68	38.81	50567	71.810	2016.4	39.02	50567	126.360
26	924.95	39.20	50567	116.936	1695.45	37.82	50567	137.051	1381.55	37.03	50567	67.859	1741.35	37.19	50567	119.535
28	791.04	37.72	50567	112.084	1466.10	36.07	50567	130.128	1166.24	35.36	50567	64.050	1495.6	35.45	50567	113.341
30	666.70	36.07	50567	107.306	1246.66	34.12	50567	123.478	957.1	33.63	50567	60.139	1258.69	33.57	50567	106.645
32	563.48	34.53	50567	103.017	1041.87	32.34	50567	117.414	781.77	32.06	50567	56.865	1044.73	31.86	50567	100.668
34	472.67	33.13	50567	99.471	866.36	30.64	50567	111.631	634.78	30.63	50567	54.179	864.72	30.25	50567	95.306
36	382.72	31.52	50567	95.661	686.95	28.90	50567	105.596	495.99	29.15	50567	51.201	682.77	28.58	50567	89.560

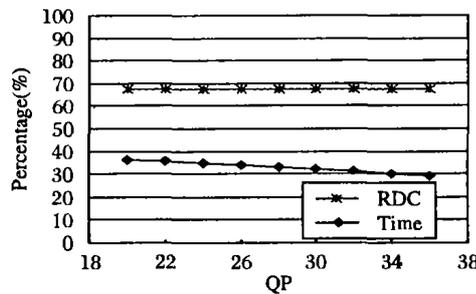


(a) Tempete QCIF [250 Frames]

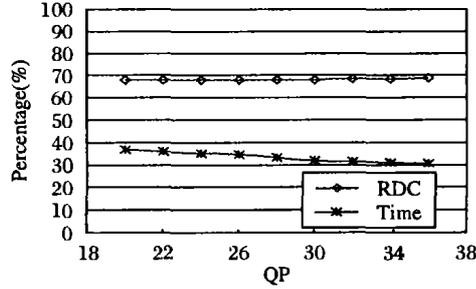


(b) Bus QCIF [150 Frames]

图 2 性能比较曲线



(a) Tempete QCIF [250 Frames]



(b) Bus QCIF [150 Frames]

图 3 RDC 编码时间曲线

NEW 代表快速算法;RDC 代表 RD_Cost 函数调用次数减少的百分比;Time 代表编码时间减少的百分比。

5 结语

本文对 H. 264/AVC 的帧内预测进行了深入研究,利用相邻块的预测模式之间的相关性以及相关预测模式之间编码的特点,提出了一种 RDO 模式下的快速 Intra_4 × 4 模式选择算法,该算法可以预先排除 65% 以上的 Intra_4 × 4 模式,避免了不必要的代价计算,从而大幅度降低帧内预测的复杂度,同时基本保持了 H. 264/AVC 的编码性能。该算法简单有效,具有

重要的实用价值。

参考文献:

- [1] PAN F, LIN X, RAHARDJA S, *et al.* JVT-G013, Fast ModeDecision for Intra Prediction [S]. JVT of ISO/IEC MPEG & ITU-T VCEG, 2003.
- [2] 裴世保, 李厚强, 俞能海. 一种快速的 H. 264/AVC 帧内预测模式选择算法[J]. 计算机工程与应用, 2005, 41(10).
- [3] RICHARDSON IEG. H. 264/MPEG 4 Part 10 White Paper: Prediction of Intra Macroblocks [EB/OL] <http://www.vcodex.com/h264.html>, 2005-01.
- [4] 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).
- [5] JM Reference Software Version 7.3[CP/OL]. <http://bs.hhi.de/~suehring/tml/>, 2005-01.
- [6] KIM C, SHIH H-H, KUO C-CJ. Multistage mode decision for intra prediction in H. 264 codec[A]. 16th Annual Symposium Electric Imaging[C], 2004.
- [7] YANG KH, JACQUIN A, JAYANT NS. A normalized rate-distortion model for H. 263-compatible codecs and its application to quantizer selection[A]. Proceedings of International Conference on Image Processing[C], 1997, 2: 41-44.
- [8] WIEGAND T, SULLIVAN GJ, BJONTEGAARD 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).