

基于块参考像素特征的快速帧内预测模式选择新算法

蒋刚毅^{①②} 李世平^{①②} 叶锡恩^① 郁梅^{①②}

^①(宁波大学电路与系统研究所 宁波 315211)

^②(中国科学院计算技术研究所 北京 100080)

摘要 帧内预测是 H.264 采用的一种编码新技术。与以前的编码标准相比, H.264 编码性能有了很大的提高, 但同时编码复杂度和计算量也明显增加。该文研究了提高 H.264 编码速度的途径, 着重分析了帧内预测模式快速选择的问题。经过分析发现, 4×4 亮度块的参考像素亮度值通常具有相似性, 对应地, 4×4 亮度块的 9 种预测模式所得的预测结果也是相近的。根据 4×4 亮度块参考像素的这一特征, 该文提出了一种基于 4×4 亮度块参考像素特征的快速帧内预测模式选择算法。实验结果说明了该文算法的有效性, 与 H.264 编码器 JM92 相比, 该算法的计算量减少了 46%, 而编码质量、PSNR 和输出码率保持基本不变。

关键词 H.264 标准, 帧内预测模式选择, 编码复杂度, 块参考像素特征

中图分类号: TN919.8

文献标识码: A

文章编号: 1009-5896(2006)10-1874-05

A Novel Fast Mode Selection Algorithm for Intra Prediction Based on Characteristics of Block's Reference Pixels

Jiang Gang-yi^{①②} Li Shi-ping^{①②} Ye Xi-en^① Yu Mei^{①②}

^①(Institute of Circuits and Systems, Ningbo University, Ningbo 315211, China)

^②(Institute of Computing Technology, Chinese Academy of Science, Beijing 100080, China)

Abstract Intra-prediction is one of the new techniques adopted by H.264. Compared with the previous video coding standards, H.264 improves coding performance greatly. However, the computation load of H.264 video coding increases drastically. This paper investigates improvement of encoding speed for H.264 with special focus on fast intra prediction mode selection. Some experimental analyses indicate that the luminance values of the reference pixels for each 4×4 luminance block are commonly similar, and in that case, the prediction blocks of the nine prediction modes are also similar. According to this feature of 4×4 luminance block, a fast mode selection algorithm for intra prediction is proposed. Experimental results show the proposed scheme is very effective. Compared with JM92 in H.264 encoder, it reduces 46% encoding time while maintaining the PSNR and bit rate.

Key words H.264, Intra-prediction mode selection, Encoder complexity; Characteristics of block's reference pixels

1 引言

传统的帧内编码技术是将宏块的亮度或色度数据直接进行变换和量化,它消除的只是宏块块内冗余,这使得帧内编码输出的码率比帧间编码输出的码率高很多。为了进一步提高帧内编码效率, H.264 引入了帧内预测技术^[1,2]。帧内预测是通过相邻块的像素来预测当前块的像素,得到预测块,预测块和原始块的差值数据再进行变换量化,使得输出的码率大大减少,进一步消除了空间冗余信息。H.264 的亮度信息预测,有 4×4 块和 16×16 块两种块类型,其中 4×4 块有 9 种预测模式,而 16×16 只有 4 种预测模式。对于色度信息的预测只有 4 种 8×8 块预测模式,这 4 种模式和 16×16 亮度

块的 4 种预测模式是相同的^[3]。H.264 要通过率失真优化(Rate Distortion Optimization, RDO)选择最佳的预测模式,这个过程是非常耗时的。所以有必要提出快速的帧内预测模式选择算法。目前,在帧内预测方面的快速算法还比较少。文献[4]利用图像的边缘方向直方图对帧内预测模式进行预判,从而提高了帧内预测的速度。文献[5]提出了基于优化的率失真模型的两步选择快速帧内预测算法。文献[6]提出了一种针对 4×4 亮度块的快速预测模式选择算法,它先对下采样后的块进行模式选择,粗选出候选预测模式,再进一步精确选择。以上 3 种方法都是针对块内部的各种特征提出的。

本文针对 4×4 亮度块的参考像素的特征,提出了一种新的帧内预测模式选择的快速算法。新算法首先判断 4×4 亮度块参考像素亮度值的方差,如果方差比较小,则直接采用最有可能预测模式。最有可能预测模式是通过当前 4×4 块的左边块和上边块的预测模式得到的,不需要复杂的 RDO 计算。与 9 种预测模式完全遍历的方法相比,新算法大

2005-01-31 收到, 2005-10-08 改回

国家自然科学基金(60472100), 浙江省自然科学基金青年人才基金(RC04057), 浙江省科技攻关项目(2004C31105)和宁波市科技局项目(2003A61001, 2004A61001, 2004A630002)资助课题

大减少了 RDO 计算的次数, 从而明显降低了帧内预测的计算复杂度。

2 快速帧内预测模式选择算法

图 1(a)中, $a-p$ 表示 4×4 亮度块的 16 个像素, $A-L$ 和 X 表示预测时可能用到的 13 个参考像素。图 1(b)是 4×4 亮度块 9 种预测模式中的 8 种方向预测模式, 按顺序分别是垂直预测(VP), 水平预测(HP), 对角线左下预测(DDL), 对角线右下预测(DDR), 垂直偏右预测(VRP), 水平偏下预测(HDP), 垂直偏左预测(VLP), 水平偏上预测(HUP)。余下一种为直流系数预测模式(DP)。假设 4×4 亮度块选择了水平预测模式, 那么 $a-d$ 像素的亮度值就用 I 像素的亮度值进行预测, $e-p$ 像素的预测过程和 $a-d$ 像素的预测过程类似。

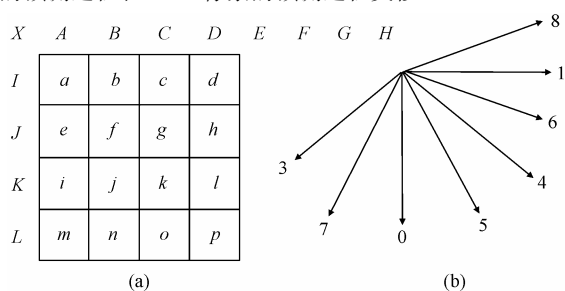


图 1 4×4 亮度块的 8 种方向预测模式

Fig.1 4×4 luma Intra prediction modes

在 H.264 标准中, 帧内预测块的预测模式不是直接进行熵编码的, 而是将其与最有可能模式相减后再进行熵编码。最有可能模式的符号表示为 `most_probable_mode`, 它是由当前块的左边块和上边块的预测模式得到的。当前块的预测模式为 `most_probable_mode` 时, 只需要用一个 bit 来表示预测模式, 相对于选用其它模式, 其所用的比特数最少。H.264 编码标准中的 RDO 是提高编码性能的一个重要技术, 但它同时也大大提高了编码复杂度和计算量。如果编码器采用了 RDO, 那么帧内编码就要对帧内预测模式的每一个组合进行计算, 最后选出一个率失真最优的模式。根据 H.264 中帧内预测的结构, 每一个宏块都要进行 $M8 \times (M4 \times 16 + M16)$ 次 RDO 计算, 其中 $M8$ 、 $M4$ 、 $M16$ 分别表示 8×8 色度块、 4×4 亮度块、 16×16 亮度块的预测模式数, 这也就意味着每个宏块要进行 592 次 RDO 计算。这个计算量是巨大的, 所以, 一个快速的帧内预测模式选择算法的提出是很有必要的。

2.1 4×4 亮度块的参考像素亮度值的相似特征分析

图 1 显示, 4×4 亮度块的参考像素一共有 13 个。这 13 个像素根据位置的不同可以分成 4 个部分, 分别是左边 4 个像素(I, J, K, L), 上边 4 个像素(A, B, C, D), 左上 1 个像素(X), 右上 4 个像素(E, F, G, H), 本文分别用 `available_left`, `available_up`, `available_up_left`, `available_up_right` 来表示它们的有效性, 其值等于 1 表示有效, 等于 0 表示无效。通过对 4×4 亮度块 9 种预测模式的分析, 发现这 13 个参考像素不是在每种预测模式中都要用到的。表 1 给出了 4 个部分的参考像素和 9 种预测模式之间的关系。从表 1 可以看出, 左上 1 个像素和右

表 1 参考像素和预测模式的关系

参考像素	涉及的预测模式	模式数
左边 4 个像素	HP, DDRP, VRP, HDP, HUP, DP	6
上边 4 个像素	VP, DDL, DDRP, VRP, HDP, VLP, DP	7
左上 1 个像素	DDRP, VRP, HDP	3
右上 4 个像素	DDL, VLP	2

上 4 个像素涉及的预测模式都比较少, 这也就表示它们对 4×4 亮度块帧内预测模式选择的影响程度比较小, 所以本文不将这两部分像素作为相似性分析的对象。

本文相似性分析的对象是左边 4 个像素和上边 4 个像素。衡量相似性的标准是 8 个参考像素亮度值的方差, 方差越小, 表示这 8 个像素越相似。表 2 给出了对于一系列标准测试序列进行相似性分析的实验结果。

表 2 参考像素亮度值相似性分析实验结果

Tab.2 Similarity analysis results of luma values of the reference pixels

序列	尺寸	各方差域值下所含的 4×4 亮度块的比例(%)				
		<8	<16	<32	<48	<64
Container	QCIF	46.9	53.5	59.3	61.3	63.5
Foreman	QCIF	29.0	36.7	46.3	50.6	54.9
Claire	QCIF	65.0	70.8	74.4	76.1	77.1
Missa	QCIF	67.9	74.0	79.3	82.5	84.1
Foreman	CIF	42.4	52.7	62.0	67.0	70.6
Paris	CIF	26.3	33.5	41.9	46.0	49.2
Mthr_dotr	CIF	55.8	63.3	72.4	77.0	80.3
最大值		67.9	74	79.3	82.5	84.1
最小值		26.3	33.5	41.9	46	49.2
平均值		47.6	54.9	62.2	65.8	68.5

8 个参考像素亮度值方差在理论上的最大值为 $128 \times 128 = 16384$, 从表 2 可以看出, 亮度值方差的取值在整个区间 $[0, 16384]$ 上并不是均匀分布的, 而是在方差值比较小的区间比较集中。当 4×4 亮度块的参考像素亮度值相近时, 即其方差比较小时, 4×4 亮度块的 9 种预测模式的预测结果是非常相近的。因此, 可以提出一种很好的快速帧内预测模式选择算法, 即在 4×4 亮度块的参考像素亮度值非常相近的时候, 对该块只采用一种预测模式。假设每个 4×4 亮度块都只采用一次预测, 那么 RDO 计算可以从原来的 592 次降为 80 次, 这是本文快速算法能达到的最好结果。表 2 中给出, Missa 序列在方差域值 <16 时, 4×4 亮度块比例达 74%。此时平均一个宏块进行的 RDO 计算次数可降为 $592 \times 26\% + 80 \times 74\% \approx 213$ 次, 与原来的 592 次相比, 计算量大大地降低了。

图 2 给出了 Claire 序列首帧在各个方差值下的 4×4 块个数的直方图, 它显示了和表 2 相同的结论, 即大部分 4×4 亮度块都集中于方差值比较小的区间。对于其它测试序列也可以得到类似结果。

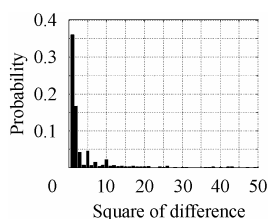


图2 Claire序列首帧在各个方差值下的 4×4 块数统计

Fig.2 Histogram of the number of 4×4 luma blocks in Claire

2.2 基于参考像素亮度值方差的快速帧内预测模式选择算法

在H.264的帧内预测中,亮度块预测模式和色度块预测模式的组合有592种,完全遍历的方法是对每一种组合进行非常耗时的RDO计算,最后得到最佳的一种组合,这是帧内预测计算复杂度高的最主要原因。本文所提出的快速帧内预测模式选择算法将根据 4×4 亮度块的参考像素亮度值方差,提前确定 4×4 亮度块的预测模式,从而减少了RDO计算的次数,降低了帧内预测的计算量。本文算法的步骤如下:

(1)根据当前块的左边块和上边块的预测模式,求取most_probable_mode。

(2)求当前块的available_left和available_up。

(3)当available_left=1且available_up=1时,转向(4);当available_left=1且available_up=0时,转向(5);当available_left=0且available_up=1时,转向(7);当available_left=0且available_up=0时,转向(9)。

(4)计算左边4个像素和上边4个像素总的方差,如果方差值小于所设定的域值,那么当前块的预测模式就采用most_probable_mode;否则转向(9)。

(5)计算左边4个像素的方差,如果方差值小于所设定的域值,则转向(6);否则转向(9)。

(6)如果most_Froable_mode为HP或HUP预测模式,那么当前块的预测模式就采用most_probable_mode,否则采用DP模式。

(7)计算上边4个像素的方差,如果方差值小于所设定的域值,则转向(8);否则转向(9)。

(8)如果most_probable_mode为VP,DDL或VLP预测模式,那么当前块的预测模式就采用most_probable_mode,否则采用DP模式。

(9)对当前块进行9种预测模式完全遍历的RDO计算,最后选择率失真代价最小的预测模式作为当前块的最佳预测模式。

3 实验结果及分析

为了验证本文算法的有效性,对比JM92^[7]在编码时间、输出码率、亮度值峰值信噪比等3个方面进行了实验比较。实验采用的参数设置文件(即JM中的CFG文件)是以JVT的主框架(main profile)设置文件为蓝本,仅将参考帧数目改为1, GOP选用IPPP的方式。实验中,方差域值设定为16。实验程序所运行的环境如下,CPU为Pentium4 1.7G,内存为DDR 256M,操作系为Windows2000 Professional。实验采用

了7个常见的测试序列,其中4个是QCIF格式,3个是CIF格式,每个测试序列帧数都是150帧。

表3-表6给出了本文算法和JM92相比较的实验结果,表中正值表示增加,负值表示减少。表中一些符号的定义如下:(1)CHG_T_ALL:本文算法较JM92在总编码时间上的变化;(2)CHG_T_I:本文算法较JM92在所有I帧编码时间上的变化;(3)CHG_PSNR:本文算法较JM92在亮度峰值信噪比上的变化;(4)CHG_BITS_ALL:本文算法较JM92在总输出码率上的变化;(5)CHG_BITS_I:本文算法较JM92在所有I帧输出码率上的变化;

从表3-表6可以看出,本文算法较JM92,在编码时间上有明显的降低,而峰值信噪比下降很微小,输出码率增加也很微小。在高码率(QP=28)情况下,本文算法较JM92,总编码时间最多减少了30%,所有I帧编码时间最多减少了46%,而亮度峰值信噪比平均下降只有0.01dB,总输出码率和所有I帧输出码率的增加均在1%左右。在低码率(QP=40)情况下,本文算法较JM92,总编码时间最多减少了34%,所有I帧编码时间最多减少了52%,而亮度峰值信噪比基本上没有变化,总输出码率平均增加0.32%,所有I帧输出码率平均增加0.51%。从上面的统计结果可以得出,本文算法在低码率时比在高码率时性能更优越,其原因在于低码率时,重建图像质量较差,从而使参考像素亮度值方差较小的 4×4 亮度块的比例更大。表6实验结果还显示,在QP=40时,对于container序列,本文算法在提高编码速度的同时,还略微提高了编码性能。这是因为JM92采用的RDO策略中,衡量失真的公式是SAD,和计算PSNR中的MSE略有差别。另外,从表3-表6可以看到,实验选用了同一测试序列Foreman的QCIF和CIF两种格式,本文算法对于高分辨率(CIF)的Foreman序列的实验结果优于低分辨率(QCIF)的Foreman序列的实验结果,这是因为对于高分辨率的图像,其所包含的方差值较小的 4×4 亮度块的比例较大。

图3给出了Container序列的码率-PSNR的三阶多项式拟合曲线,显然,本文算法和JM92所得的两条曲线是基本重合的。这也就说明,本文算法和JM92相比,在编码时间明显降低的同时,图像质量和输出码率基本上没有发生变化,从而证明了本文算法的有效性。

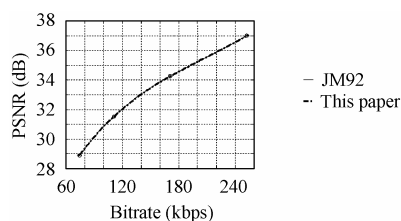


图3 Container序列码率-PSNR拟合曲线

Fig.3 RD performance curves of container sequence

4 结束语

本文对H.264的帧内预测技术进行了深入的分析,实验分析发现, 4×4 亮度块的参考像素的亮度值方差通常比较小,而当方差值比较小时,采用9种预测模式所得的预测结果非

表 3 QP=28 实验结果(%)

Tab.3 QP=28, comparison results(%)

序列	尺寸	CHG_T_ALL	CHG_T_I	CHG_PSNR (dB)	CHG_BITS_ALL	CHG_BITS_I
Container	QCIF	- 20.54	- 29.03	- 0.01	0.16	- 0.89
Foreman	QCIF	- 13.02	- 18.28	0	0.88	1.39
Claire	QCIF	- 28.25	- 41.77	- 0.04	0.95	1.57
Missa	QCIF	- 29.97	- 45.95	0.02	0.93	1.69
Foreman	CIF	- 21.08	- 30.67	0	1.26	1.20
Paris	CIF	- 12.96	- 17.02	- 0.01	0.34	0.37
Mthr_dotr	CIF	- 27.25	- 39.94	- 0.02	2.20	3.05
最大值		- 12.96	- 17.02	0.02	2.20	3.05
最小值		- 29.97	- 45.95	- 0.04	0.16	- 0.89
平均值		- 21.87	- 31.81	- 0.01	0.96	1.20

表 4 QP=32 实验结果(%)

Tab.4 QP=32, comparison results(%)

序列	尺寸	CHG_T_ALL	CHG_T_I	CHG_PSNR(dB)	CHG_BITS_ALL	CHG_BITS_I
Container	QCIF	- 20.90	- 30.36	- 0.02	- 0.11	- 0.86
Foreman	QCIF	- 13.55	- 19.64	- 0.01	0.48	2.15
Claire	QCIF	- 28.46	- 42.28	0	0.82	0.94
Missa	QCIF	- 30.80	- 47.52	0.01	0.66	0.00
Foreman	CIF	- 21.91	- 32.45	0.01	0.52	0.61
Paris	CIF	- 13.77	- 18.69	0	0.37	0.71
Mthr_dotr	CIF	- 28.23	- 41.76	0.01	1.37	1.78
最大值		- 13.55	- 18.69	0.01	1.37	2.15
最小值		- 30.80	- 47.52	- 0.02	- 0.11	- 0.86
平均值		- 22.52	- 33.24	0	0.59	0.76

表 5 QP=36 实验结果(%)

Tab.5 QP=36, comparison results(%)

序列	尺寸	CHG_T_ALL	CHG_T_I	CHG_PSNR(dB)	CHG_BITS_ALL	CHG_BITS_I
Container	QCIF	- 21.98	- 32.26	0.01	0.51	0.59
Foreman	QCIF	- 15.12	- 21.43	0.03	0.65	0.40
Claire	QCIF	- 29.83	- 44.68	0.05	1.37	0.14
Missa	QCIF	- 32.24	- 48.89	- 0.07	0.74	0.17
Foreman	CIF	- 23.10	- 34.16	0	0.11	- 0.51
Paris	CIF	- 15.28	- 20.32	0	0.17	- 0.09
Mthr_dotr	CIF	- 30.28	- 44.70	- 0.01	1.13	- 1.20
最大值		- 15.12	- 20.32	0.03	1.37	1.20
最小值		- 32.24	- 48.89	- 0.07	0.11	- 0.51
平均值		- 23.98	- 35.20	- 0.01	0.67	0.27

表 6 QP=40 实验结果(%)
Tab.6 QP=40, comparison results(%)

序列	尺寸	CHG_T_ALL	CHG_T_I	CHG_PSNR(dB)	CHG_BITS_ALL	CHG_BITS_I
Container	QCIF	- 23.24	- 34.25	0.02	- 0.41	- 0.66
Foreman	QCIF	- 16.00	- 23.61	0.02	0.12	- 2.79
Claire	QCIF	- 29.59	- 45.26	- 0.06	0.91	1.41
Missa	QCIF	- 33.76	- 51.52	0.03	1.11	2.72
Foreman	CIF	- 24.91	- 37.09	0	- 0.43	0.98
Paris	CIF	- 17.50	- 23.76	0	0.15	0.59
Mthr_dotr	CIF	- 31.98	- 46.69	- 0.01	0.79	1.30
最大值		- 16.00	- 23.61	0.03	1.11	2.72
最小值		- 33.76	- 51.52	- 0.06	- 0.43	- 2.79
平均值		- 25.28	- 37.45	0	0.32	0.51

常相似, 根据这一特性, 本文提出了一种基于参考像素亮度值方差的帧内预测模式快速选择算法。实验结果表明, 本文算法和 JM92 相比, 帧内预测的计算复杂度明显降低, 同时编码质量和输出码率保持基本不变。现有的帧内预测模式选择算法是基于块的内部特征的, 而本文算法是针对预测的参考像素特征提出的, 两类方法具有互补性。下一步的工作就是将两类方法进行结合, 进一步提高帧内预测的速度。

参 考 文 献

- [1] ITU-T Rec. H.264 | ISO/IEC 11496-10 AVC. Document JVT-G050. 7th Meeting: Pattaya, Thailand. March 2003.
- [2] Topiwala P, Sullivan G, Joch A, *et al.*. Performance evaluation of H.26L, TML 8 vs. H.263++ and MPEG-4. Santa Barbara, CA; USA, September 2001.
- [3] Wiegand T, Sullivan G, Bjontegaard G, Luthra A. Overview of the H.264/AVC video coding standard. *IEEE Trans. on Circuits Syst. Video Technol.*, 2003, 13: 560 – 576.
- [4] Pan F, Lin X, *et al.*. Fast mode decision for intra prediction, in ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, JVT 7th

Meeting, Pattaya II, Thailand, Mar. 2003.

- [5] Kim C S, Li Q, Kuo C C J. Fast Intra-prediction model selection for H.264 codec, SPIE International Symposium ITCOM 2003, Orlando, Florida, Sept. 7-11, 2003.
- [6] Meng B, Au O. Fast Intra-prediction mode selection for 4×4 blocks in H.264, IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03), April 2003. vol 3: 389 – 392.
- [7] ITU-T Rec. H.264 | ISO/IEC 11496-10 AVC. JM9.2[S]. Nov. 2004.

- 蒋刚毅: 男, 1964 年生, 教授, 中国科学院计算技术研究所兼职博士生导师, 主要研究方向为多媒体信息传输与信息安全、图像处理与视频信号编码、基于视觉的智能控制。
- 李世平: 男, 1981 年生, 硕士生, 研究方向为面向网络的立体视频编码和传输技术, 多描述编码和传输。
- 叶锡恩: 男, 1959 年生, 副教授, 主要研究方向为多媒体信息与数字系统设计。
- 郁 梅: 女, 1968 年生, 教授, 主要研究方向为包括多媒体信号处理、编码与通信, 计算机视觉及其应用。