

H. 264 快速帧内预测模式选择算法

张江鑫 冯明

(浙江工业大学浙江省光纤重点实验室,杭州 310014)

摘要 H. 264 引入帧内预测技术提高了 I 帧的编码效率,但也大大增加了编码的计算复杂度。为了降低计算复杂度,根据预测模式算法的特征提出了一种基于像素边缘矢量方差及子块宏块映射关系的帧内预测模式快速选择算法。实验结果表明,该算法在保证图像质量和比特率大致不变的前提下,极大地降低了帧内预测编码的计算复杂度。

关键词 视频编码 H. 264/AVC 帧内预测

中图分类号:TN919.81 文献标识码:A 文章编号:1006-8961(2008)10-2023-04

Fast Intra-prediction Mode Selection Algorithm for H. 264

ZHANG Jiang-xin, FENG Ming

(Zhejiang Provincial Key Laboratory of Fiber Optic Communication Technology, Zhejiang University of Technology, Hangzhou 310014)

Abstract H. 264 introduces intra-prediction coding technique to improve the intra-coding efficiency. This enormously increases the computational complexity. In order to decrease the computational complexity of the algorithm, a fast algorithm for intra-prediction mode decision based on pixel's edge vector's variance, the relation between sub-block's mode and macro-block's mode. Experimental results show that compared with the original intra-prediction coding scheme, the proposed algorithm reduce the computational complexity of intra-prediction coding dramatically with only trivial loss of PSNR and increment of bite-rate.

Keywords video coding, H. 264/AVC, intra-prediction

1 引言

H. 264 是当今世界上最先进的视频压缩编码之一,能够在较低宽带(2M 以内)情况下提供高质量的图像传输。据数据分析,在同等的画质下,H. 264 比上一代编码标准 MPEG2 节约平均 64% 的传输码流,而比 MPEG4 要节约平均 39% 的传输码流。之所以 H. 264 具有如此优越的压缩效率,是因为它采用了多种新的技术,帧内预测便是其中之一,它根据多数相邻像素具有类似性质的特点,充分利用图像的空间相关性进一步提高了压缩效率^[1-3]。

2 H. 264 帧内预测模式选择算法

在 H. 264 的帧内预测中,对于亮度信息的预测有两种类型,分别是 4×4 块和 16×16 块,其中 4×4 块有 9 种预测模式如图 1 所示,而 16×16 只有 4 种预测模式如图 2 所示。通常, 4×4 亮度块适合于纹理比较丰富的区域,而 16×16 块适合于较为平坦的区域^[4]。H. 264 编码标准中的 RDO (rate-distortion optimization) 技术是提高编码性能的一项重要技术,如果采用了 RDO,那么帧内编码就要对帧内预测模式的每一种组合进行计算。根据 H. 264 帧内预测的结构,每一个宏块都要进行 592 次 RDO 计算,这个计算量是巨大的,因此快速帧内预测模式选择算法的提出是十分必要的。

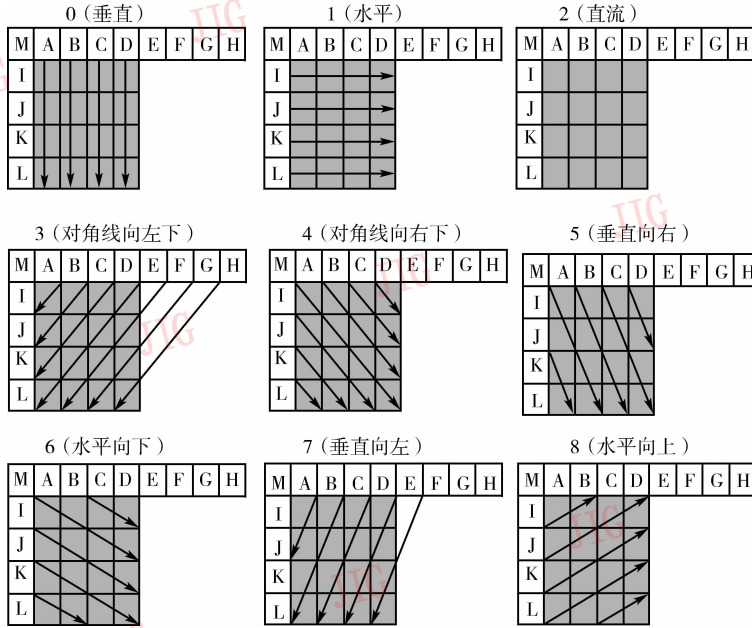


图 1 Intra_{4 × 4} 的 9 种预测模式

Fig. 1 9 prediction modes of Intra_{4 × 4}

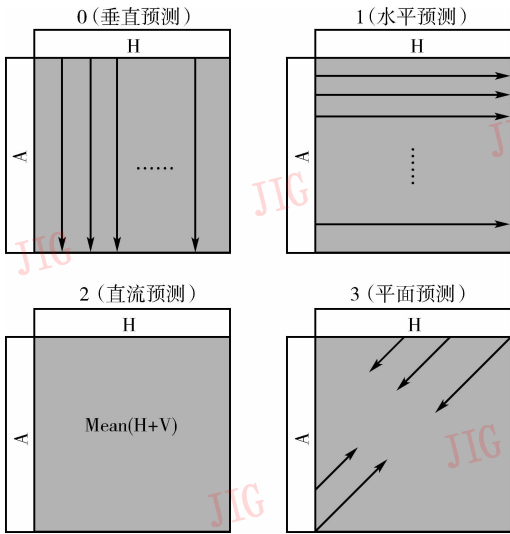


图 2 Intra_{16 × 16} 的 4 种预测模式

Fig. 2 4 prediction modes of Intra_{16 × 16}

3 快速帧内预测模式选择算法

3.1 4 × 4 帧内预测算法的改进

在 H. 264 的帧内预测模式中,大部分的预测模式都是基于方向的预测模式,特别是 4 × 4 亮度块的 9 种预测模式,其中 8 种便是方向预测。在编码时,若选择了某一方向的预测模式,也就表示该块的纹理具有所选预测模式对应的方向^[5],而当块的方向性不明显时,DC(直流)模式则显得比较有效。基于

这样的事实,对于 4 × 4 子块预测模式的选择过程如下:

首先,对于一个 4 × 4 子块,计算其中每个像素的边缘矢量,由于边缘矢量的方向与预测模式的方向未必保持一致,为了提高准确性,计算每个边缘矢量在与它方向最接近的两个预测模式上的投影。记像素 $D_{i,j}$ ($i = 0, 1, 2, 3; j = 0, 1, 2, 3$), $dx_{i,j}, dy_{i,j}$ 分别为利用 Sobel 算子得到的该点的水平和垂直边缘信息,边缘矢量的幅值为 $Amp(D_{i,j})$,边缘矢量的方向为 $Ang(D_{i,j})$, $Ang(D_{i,j})$ 与预测模式 n ($n = 0, 1, \dots, 8$) 的夹角为 θ_n , $D_{i,j}$ 在预测模式 n 上的投影值为 $AvAmp(n)_{i,j}$, 则各个值的计算公式如下:

$$Amp(D_{i,j}) = |dx_{i,j}| + |dy_{i,j}| \quad (1)$$

$$Ang(D_{i,j}) = \begin{cases} \arctan\left(\frac{dy_{i,j}}{dx_{i,j}}\right) & dx_{i,j} \neq 0 \\ -\frac{\pi}{2} & dx_{i,j} = 0 \end{cases} \quad (2)$$

$$AvAmp(n)_{i,j} = |Amp(D_{i,j}) \sin(\theta_n - Ang(D_{i,j}))| \quad (3)$$

由于帧内预测模式的方向集中在 $[-\pi/2, \pi/2]$, 其中垂直向左和对角线向下的方向分别为 $133\pi/900$ 和 $\pi/4$, 因此求边缘矢量的方向时,可以直接使用式(2)。

其次,计算出 16 个像素的 32 个 $AvAmp(n)_{ij}$,记 $Sum_AvAmp(n)$ 为该子块中像素边缘矢量到预测模式 n 上的投影值总和,则

$$Sum_AvAmp(n) = \sum_{i=0}^{i=3} \sum_{j=0}^{j=3} AvAmp(n)_{i,j} \quad (4)$$

对于子块的方向性是否明显,可以利用 $Sum_AvAmp(n)$ 的波动大小即方差来反映。考虑以下两种事实:

(1) 假定对于每个 n , $Sum_AvAmp(n)$ 都相等,即子块的方向最不明显时, $Sum_AvAmp(n)$ 的方差为 0。

(2) 假定对于每个 n , 只有一个 $Sum_AvAmp(n)$ 不等于 0, 记这个值为 $M (M > 0)$, 即子块的方向最明显时, $Sum_AvAmp(n)$ 的方差为 $\sqrt{14}/4M$ 。

因为不同的图像之间, $Sum_AvAmp(n)$ 的大小可能会有较大差异, 为了方便后面的阈值设定, 计算 $Sum_AvAmp(n)$ 的方差和 $Sum_AvAmp(n)$ 的平均值的商, 记为 T , 以消去这个差异, 这样一来, 上述两种情况下的 T 分别为 0 跟 $2/\sqrt{14}$ 。

设定一个阈值, 如果该子块 T 小于等于阈值, 认为子块的方向性不明确, 那就用 DC 预测模式; 如果大于阈值, 那么就选 $Sum_AvAmp(n)$ 最大的 3 个模式, 再通过计算代价函数从中选出最优的预测模式。

3.2 16 × 16 帧内预测算法的改进

同样, 在 16 × 16 宏块的预测模式中, 宏块纹理的方向性对预测模式的选择也起到了重要作用。对比 16 × 16 宏块的 4 种预测模式和 4 × 4 子块的 8 种预测模式的方向, 建立了表 1, 它显示了 4 × 4 子块的预测模式在 16 × 16 宏块里对应的模式, 并做出如下规定: 当子块的最优模式 0, 5, 7 的和大于 1, 6, 8 的和时选用垂直、DC、平面预测, 当 0, 5, 7 的和等于 1, 6, 8 的和时, 4 种模式都选用, 不然则选择水平、DC、平面预测, 再通过计算代价函数值确定最优模式; 特别地当最优模式 0 的和大于 9 时, 只选用垂直模式, 当最优模式 1 的和大于 9 时, 只选用水平方式, 当最优模式 3, 4 的和大于 9 时, 只选用 DC 和平面预测。对于表 1 和以上规定, 采用纹理不同的标准序列进行测试, 每种序列以 1 帧压缩 100 帧, 得到的正确率情况如表 2 所示。整个快速算法的主要流程如图 3 所示。

表 1 4 × 4 预测模式与 16 × 16 预测模式关系

Tab. 1 The relation between 4 × 4 and 16 × 16

关系编号	子块预测模式	宏块预测模式
1	Mode 1	水平预测
2	Mode6, Mode8	水平、DC、平面
3	Mode3, Mode4	DC、平面预测
4	Mode5, Mode7	DC、平面、垂直
5	Mode0	垂直预测

表 2 预测模式关系正确率

Tab. 2 Exactness rate of prediction mode's relation

单位: %

	测试序列			
	mother	pairs	flower	mobile
1	82.6	84.3	79.3	80.7
2	90.2	92.7	87.1	90.4
3	83.5	80.8	84.7	85.3
4	91.1	89.5	90.7	86.4
5	83.7	85.6	80.5	77.1

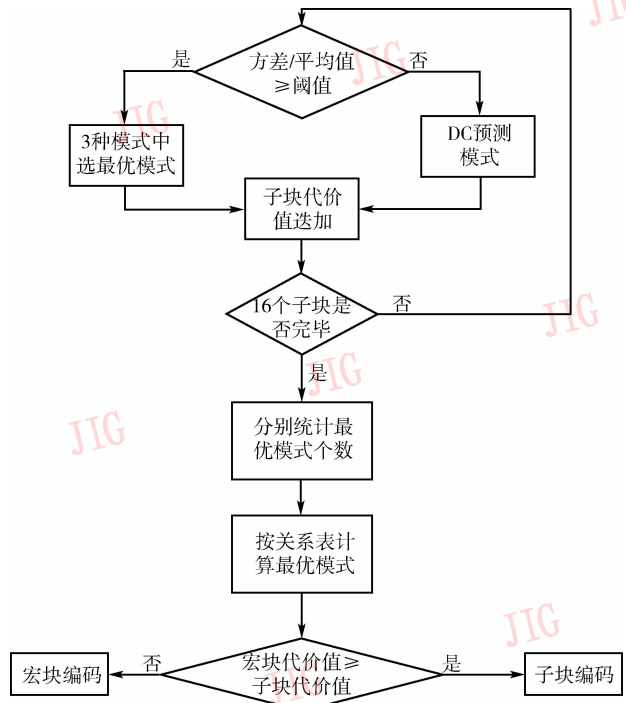


图 3 快速算法流程图

Fig. 3 Flow chat of the fast selection algorithm

4 实验结果与分析

实验平台为 Intel Pentium Dual CPU, EZ160 1.80GHz, 2.00GB 内存, 测试软件为 JVT 提供的 JM8.6^[6], 使用了哈达玛 (Hadamard) 变换, RDO 技

术, cavlc 熵编码, 码流为 IPPP, 帧速为 30fps, 编码长度为 250f, 对 5 个标准测试序列 foreman、silent、claire、salesman、news 分别在不同的量化系数下进行了编码, 实验数据如表 3 所示。

表 3 测试序列
Tab. 3 Test sequences

	量 参	Time (s)		PSNR (dB)		Rate (kbits/s)	
		优化前	优化后	优化前	优化后	优化前	优化后
foreman	20	189.240	72.005	42.86	42.81	1 586.31	1 599.00
	24	166.879	64.548	39.69	39.65	1 096.01	1 104.88
	28	148.075	72.719	36.84	36.82	749.64	758.05
	32	131.607	57.911	33.99	33.97	499.03	505.86
	36	119.358	47.790	31.31	31.27	330.80	335.68
silent	20	198.520	71.923	43.05	43.01	1 739.40	1 754.18
	24	174.425	63.961	39.64	39.60	1 215.43	1 226.01
	28	152.784	57.033	36.49	36.46	812.78	821.65
	32	134.175	50.825	33.57	33.54	530.41	537.94
	36	120.189	46.008	31.07	31.05	338.44	343.53
claire	20	139.862	56.614	46.45	46.39	738.27	744.32
	24	129.880	52.652	43.70	43.67	536.24	540.80
	28	121.450	49.275	40.88	40.84	380.31	383.88
	32	113.703	46.150	37.94	37.93	269.54	272.38
	36	107.027	43.447	35.48	35.46	182.65	184.77
salesman	20	204.579	71.902	43.12	43.05	1 779.10	1 795.47
	24	180.919	64.548	39.62	39.58	1 257.98	1 269.43
	28	158.394	57.181	36.37	36.34	858.85	867.71
	32	137.913	49.729	33.31	33.27	557.59	563.61
	36	121.370	44.213	30.55	30.49	348.70	352.89
news	20	180.595	74.440	44.09	44.06	1 480.03	1 490.37
	24	165.459	68.154	40.88	40.84	1 101.74	1 115.38
	28	151.097	62.310	37.83	37.79	807.46	813.85
	32	137.063	56.564	34.59	34.57	581.25	586.94
	36	124.655	51.495	31.66	31.62	406.29	410.88
	40	114.256	47.194	28.78	28.74	275.90	279.54

从以上实验数据我们可以计算得到, 优化后速度平均提高 61.09% 左右, 信噪比平均增加 0.03dB, 码率平均增加 1.08%, 并且对于不同量化系数, 不同纹理复杂程度的测试序列之间, 优化效果没有大的变化波动, 达到了令人满意的实验结果。

5 结 论

本文分析了 H. 264/AVC 帧内预测技术, 并针对 H. 264 参考代码中帧内预测模式选择算法计算复杂度大的弊端, 提出了一种快速算法, 该算法准确利用了边缘矢量与预测模式方向的关系, 借助于计算矢量方差来确定选择 DC 模式还是 8 种方向不同的模式, 以及借助于子块与宏块的关系减少了 16×16 块的模式选择, 实验结果表明, 该算法在信噪比和码率变化不大的情况下, 大大缩短了预测的时间。

参考文献 (References)

- 1 ITU-T Recommend H. 264/ISO/IEC 11496-10. Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (Advanced Video Coding) [S].
- 2 Thomas Wiegand. Overview of the H. 264 video Coding standard [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 1 ~ 19.
- 3 Iain E G. Richardson. H. 264/MPEG 4 Part 10 White Paper: Intra Prediction [DB/OL]. <http://www.vcodex.com/h264.html>, 2003-04.
- 4 Li Shi-ping Jiang Gang-yi. New fast mode selection for intra prediction [J]. Acta Electronica Sinica, 2006, 34(1): 141 ~ 146. [李世平, 蒋刚毅. 快速帧内预测模式选择新方法 [J]. 电子学报, 2006, 34(1): 141 ~ 146.]
- 5 Pan F, Lin X, Susanto R, et al. Fast mode decision for intra prediction [DB/OL]. http://stamdarots.polycom.com/2003_03_Pattaya/JVT-G013.doc, 2005-01-17.
- 6 Joint Video Team (JVT). Reference Software JM86 [CP/OL]. http://iphome.hhi.de/suehring/tml/download/old_jm/jm86.zip, 2007-08-07.