

基于图像特征的 H.264 快速帧内模式预测算法

New Fast Intra Mode Decision Algorithm Based on Image Feature for H.264

李志娟 戴瑜兴 刘宏立 (湖南大学 电气与信息工程学院 湖南 长沙 410082)

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。在深入分析帧内预测原理以及预测模式选择过程的基础上,提出了一种快速 H.264 帧内预测模式选择算法。该算法通过预测模式尺寸预先选择原则、帧内预测模式快速选择原则和提前中止原则来加快编码速度。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,与 H.264 原有的全搜索算法相比,编码速度能够平均提高 41.3% ~ 53.4%;与快速三步法相比,编码速度能够平均提高 21.4% ~ 32.4%。

关键词: H.264 帧内预测 模式选择

H.264 是由 ISO/IEC 与 ITU - T 组成的联合视频组 JVT 制定的新一代视频压缩编码标准。H.264 采用和以前的标准一样,其编码算法的基本构成就是将输入图像预测和预测误差图像相关性的编码组合的混合编码模式,通过宏块来描述各个编码图像。但与已有的视频标准相比,在相同失真率条件下 H.264 的编码效率提高了 50% 左右,高效的压缩效率与 H.264 采用的一些新技术密不可分,其中基于空间域的帧内预测就是其中一个亮点。然而随着大量新技术的引入, H.264 编码器的计算复杂度增加,在资源受限或高实时性要求的环境下很难达到预定的要求,因此为了满足实时应用的要求,从帧内预测的重要性和预测的复杂度可以看出,研究帧内预测的快速算法具有重要的实用价值。

与大量的帧间预测快速算法相比,快速帧内预测的研究还不多,比较有代表性如 FIPMS 算法^[1]利用 4×4 块内像素的相关性,选择部分像素进行预测。文献^[2]提出的算法通过整数 DCT 变换提供的方向信息以及最小均方误差来减少需要进行 RDO 运算的次数。文献^[3]通过分析块的边缘信息,而文献^[4]提出的快速三步预测模式选择算法只是针对帧内 4×4 预测模式。这些方法都各有其优缺点。本文首先根据宏块的直方

图特性,利用自相关函数判断宏块采用帧内 4×4 预测还是帧内 16×16 预测,然后改进了快速三步法,提出一种高效的快速帧内预测模式选择算法。

1 H.264 帧内预测原理与模式选择

在以往的视频标准中,除 MPEG-4 外都是直接对帧内编码模式的宏块进行变换,量化和熵编码。为了进一步提高帧内编码的压缩效率, H.264 第 1 次采用了基于空间像素预测技术的帧内编码技术,对当前宏块内的像素亮度采用基于 4×4 像素分块预测和 16×16 像素分块预测算法;对像素色度块采用了基于 8×8 分块预测算法。H.264 标准中规定了 9 种 4×4 的预测模式、4 种 16×16 预测模式和 4 种 8×8 预测模式^[5],具体预测模式如图 1 所示,图中箭头表明每种模式的预测方向。

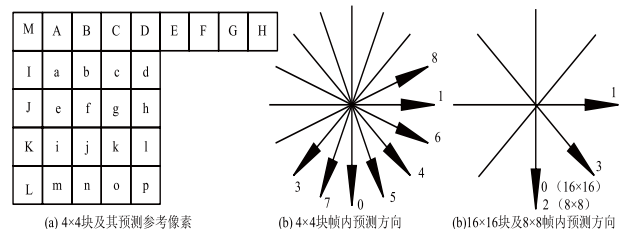


图 1 帧内预测模式

基金项目:国家高技术研究发展计划(863)(2007AA1A121)

收稿时间:2009-01-20

在 H.264 帧内编码中,每个宏块亮度信号都要完成 9 种 Intra_4 × 4 预测模式和 4 种 Intra_16 × 16 预测模式,计算率失真开销进行预测模式选择,得出一种最优预测模式,对每一种 8 × 8 色度宏块的帧内预测模式计算相应率失真开销,获得相应的宏块组合率失真开销,选择最小的宏块率失真开销作为该宏块组合的最佳帧内预测模式,使得编码后的图像在码流和图像质量两者权衡之后获得一个最佳。在标准推荐的算法中,对帧内预测采用全搜索算法,可知完成一个宏块组合共有 $M8 \times (M4 \times 9 + M16) = 592$ (其中 $M8 = 4, M4 = 16, M16 = 4$) 种模式计算才能获得最优模式,因此这一过程的计算复杂度很高。

2 快速帧内预测模式选择算法

H.264 标准将处理的图像块分为两种情况:4 × 4, 16 × 16。对原图像某区域而言,若图像变化平缓则利用 16 × 16 的块为处理对象;而当图像变化剧烈或要求较高的压缩比(在同样的图像质量情况下)就采用 4 × 4 的块进行处理。图像块大小的不同使其可以根据图像特征在预测前就判断宏块是采用 4 × 4 预测模式还是采用 16 × 16 预测模式,从而减少运算量。本文在此基础上结合预测模式选择的统计特性提出一种高效简单的快速帧内预测算法。

2.1 图像特征描述

对于细节较少的图像区域,编码器通常会选择大尺寸的预测模式,而对于细节丰富的图像区域,编码器选择小尺寸的预测模式。如果能预先检测当前宏块是否包含大量细节,从而选择相应尺寸的预测模式,就能有效地节省编码时间。为了更清楚地说明这一事实,选取标准 QCIF 图像序列 silent、foreman、coastguard、news、akiyo、hall、container 进行测试。每个测试序列的大小都为 300 帧,即每个测试序列含有的宏块数为 $(176 \times 144 \times 300) / (16 \times 16) = 29700$ 。帧内编码不同模式宏块数及百分比如表 1 所示。

实验表明,对于细节较少的图像区域,编码器通常会选择大尺寸的预测模式,而对于细节丰富的图像区域,编码器选择小尺寸预测模式的几率会相增加。

图像直方图是图像处理中一种十分重要的图像分析工具,它描述了一幅图像的灰度级内容,任何一幅图像的直方图都包含了丰富的信息。图像直方图是图像各灰度值统计特性与图像灰度值的函数,它统计一

表 1 帧内编码不同模式宏块数及百分比

测试序列	模式	QP		
		24	28	32
silent	Intra 4x4	29283(98.60%)	28676(96.55%)	27039(91.04%)
	Intra 16x16	417(1.40%)	1024(3.45%)	2661(8.96%)
foreman	Intra 4x4	27575(92.85%)	26837(90.36%)	25618(86.26%)
	Intra 16x16	2125(7.15%)	2863(9.64%)	4082(13.74%)
coastguard	Intra 4x4	28600(96.30%)	26800(90.24%)	22704(76.44%)
	Intra 16x16	1100(3.70%)	2900(9.76%)	6996(23.56%)
news	Intra 4x4	25251(85.02%)	24547(82.65%)	22550(75.93%)
	Intra 16x16	4449(14.98%)	5153(17.35%)	7150(24.07%)
akiyo	Intra 4x4	26078(87.80%)	24033(80.92%)	21690(73.03%)
	Intra 16x16	3622(12.20%)	5667(19.08%)	8010(26.97%)
hall	Intra 4x4	21745(73.22%)	21115(71.09%)	20100(67.68%)
	Intra 16x16	7955(26.78%)	8585(28.91%)	9600(32.32%)
container	Intra 4x4	18650(62.79%)	17728(59.69%)	16981(57.18%)
	Intra 16x16	11050(37.21%)	11972(40.31%)	12719(42.82%)

幅图像中各个灰度级出现的次数或概率,恰好可以反映图像细节的丰富程度。如果一个灰度级出现的次数很多,就说明这个灰度级是图像中主要的灰度分量,即该图像主要是由一种灰度构成。相反,如果每个灰度级出现的概率几乎相等,可以判断该图像的细节丰富。如图 2 所示,选择了标准视频序列 foreman 中典型的采用 Intra16 × 16 模式的宏块和采用 Intra4 × 4 模式的宏块分别得到亮度直方图。采用 Intra16 × 16 模式的宏块能量集中进行平移后直方图和原直方图几乎正交,而采用 Intra4 × 4 模式的宏块能量分散平移之后得直方图与原直方图还具有强相关性。因此可以通过求直方图的自相关函数来判定模块是否平坦。

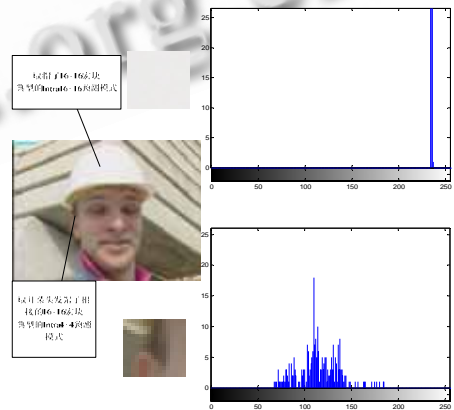


图 2 典型的 Intra16 × 16 和 Intra4 × 4 模式模式的宏块

2.2 帧内预测算法改进

在 Intra_4 × 4 帧内模式预测的研究中,本文引用文献[6]中的统计数据选择 7 个具有不同运动特征的视

频序列,对 9 种 Intra_4 × 4 模式出现的百分比进行统计,如表 2 所示。

表 2 Intra4×4 模式统计百分比

测试序列	Mode0	Mode1	Mode2	Mode3	Mode4	Mode5	Mode6	Mode7	Mode8
silent	14.25	16.37	10.06	10.6	7.96	7.21	7.48	9.87	15.84
foreman	11.14	9.98	15.71	5.32	19.20	6.70	10.97	6.42	14.56
coastguard	4.37	50.79	8.36	1.98	4.23	1.74	7.42	3.24	17.87
news	30.92	30.51	13.35	3.53	3.4	4.77	3.92	4.84	4.76
akiyo	41.53	23.68	9.80	4.19	3.12	5.77	3.06	5.21	3.64
hall	33.1	22.6	25.0	4.2	3.6	2.5	3.1	2.3	3.6
container	13.2	33.8	27.3	3.5	4.3	2.1	7.7	1.8	6.4

由表 2 可以看出,对于不同纹理复杂度的序列,MO、MI 和 M2 对应的百分比一般排在前 3 位;对于纹理较平滑的序列,如 akiyo、news 和 hall monitor(HM),M3~M8 这 6 种模式各自所占的百分比均不超过 6%;对于纹理比较丰富的序列,如 foreman 和 silent,M3~M8 这 6 种模式所占的百分比则相对较高.然而,根据表中的统计数据,M3~M8 这 6 种模式所占的百分比之和平均仅为 30%左右,因此它们与 MO、MI 和 M2 相比还是处于次要地位。

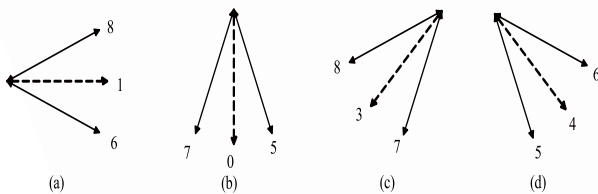


图 3 最优模式和次优模式组合

结合模式的纹理趋势及统计结果可见,当 MO 为 4 × 4 块 RDO 模式时,与 MO 预测方向相邻的 M7 和 M5 成为次优模式的可能性较大,当 M1 成为 4 × 4 块 RDO 模式时,与 M1 预测方向相邻的 M8 和 M6 成为次优模式的可能性较大。由此我们得到最小和次最小模式的组合如图 3,虚线表示当前最优模式,实线为次优模式。因此,我们提出一个 4 × 4 帧内模式快速选择算法,在进行帧内模式预测时,首先根据上述统计特性所得到的优先级次序依次计算模式 M1 率失真开销 RDCost1,模式 M0 率失真开销 RDCost0,模式 M2 率失真开销 RDCost2,计算时比较如果残差为零或阈值 TH 则判断为最优模式停止预测;否则比较 RDCost0, RDCost1, RDCost2,取最小值再按图最

小和次最小模式的组合进行计算并比较得最小失真开销 RDCost,则当前模式为最优模式。

2.3 算法具体实现流程

基于上面的考虑,本文提出了一种基于图像特征的快速帧内预测算法,总体算法如图 4 所示。首先,我们根据图像自身特征判断每个宏块包含的内容,选择 4 × 4 或 16 × 16 尺寸进行帧内预测。然后我们使用根据模式预测统计特性改进的快速帧内预测算法来选择最优模式。

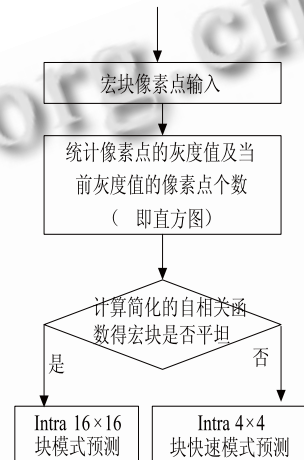


图 4 算法总体流程图

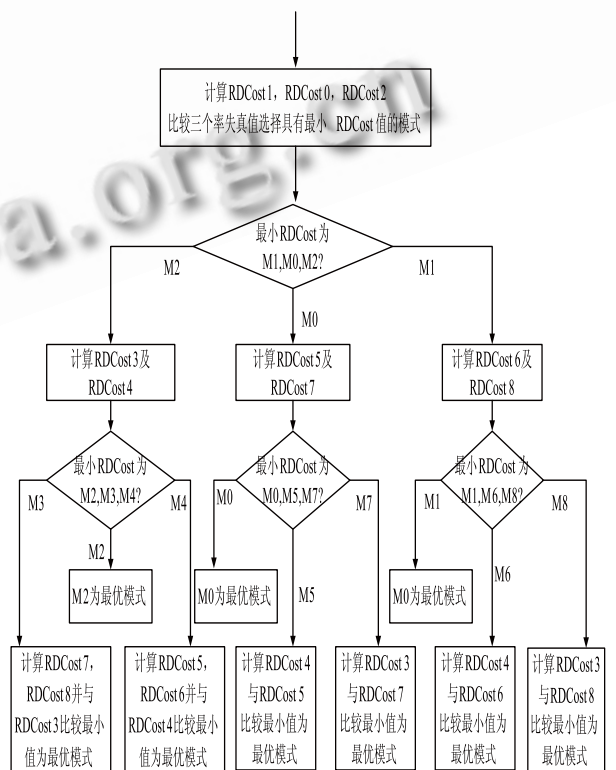


图 5 快速帧内预测算法

图 5 为快速帧内预测算法，率失真开销 RDCost 计算按下面公式进行计算。

$$RDCost(s,c,IMODE|QP,\lambda_{MODE}) = SSD(s,c,IMODE|QP) + \lambda_{MODE} \cdot R(s,c,IMODE|QP)$$

其中，IMODE 表示可选的多种预测模式，SSD(s,c,IMODE|QP) 表示失真，R(s,c,IMODE|QP)表示编码输出比特数，包括帧内预测模式和 DCT 亮度系数所需要比特数。在率失真代价函数模型下，比特率 R 值能精确计算出来从而使得模式选择的准确度大大提高。

3 实验结果与性能分析

系将本文的快速算法应用到 JM10.1 测试模型中进行实验，每个序列取 300 帧，采用全 I 帧编码方式。编码器开启 RDO，熵编码采用 CABAC，Transform8x8Mode 被设置为 0。表 2 和表 3 分别从亮度信噪比(SNR Y)、输出速率(BR)和编码时间(T)3 方面列出了本算法分别与 JM10.1 和快速三步法在不同 QP 下的比较数据。实验数据采用大量测试取平均值的处

表 3 实验结果及其对比

本文算法与JM10.1比较

测试序列	QP=24			QP=28			QP=32		
	Δ SNR Y/dB	Δ BR %	Δ T %	Δ SNR Y/dB	Δ BR %	Δ T %	Δ SNR Y/dB	Δ BR %	Δ T %
silent	↓0.016	↑0.98	↓41.3	↓0.022	↑2.49	↓48.9	↓0.01	↑3.17	↓45.4
foreman	↓0.018	↑2.56	↓44.9	↓0.016	↑2.15	↓52.5	↓0.012	↑1.01	↓45.1
coastguard	↓0.018	↑1.38	↓45.2	↓0.018	↑0.88	↓51.2	↓0.016	↑0.59	↓44.6
news	↓0.022	↑1.27	↓46.1	↓0.03	↑1.78	↓50	↓0.032	↑1.48	↓39.6
akiyo	↓0.026	↑1.49	↓53.4	↓0.032	↑1.56	↓50	↓0.024	↑1.97	↓47.3
hall	↓0.03	↑3.22	↓45.4	↓0.024	↑2.16	↓50.3	↓0.01	↑1.34	↓49.1
container	↓0.02	↑0.91	↓44.5	↓0.028	↑1.21	↓50.7	↓0.014	↑0.75	↓49.4

本文算法与快速三步法比较

测试序列	QP=24			QP=28			QP=32		
	Δ SNR Y/dB	Δ BR %	Δ T %	Δ SNR Y/dB	Δ BR %	Δ T %	Δ SNR Y/dB	Δ BR %	Δ T %
silent	↓0.008	↑0.5	↓21.4	↓0.01	↑0.92	↓25.3	↓0.002	↑1.32	↓23.7
foreman	↓0.006	↑0.53	↓21.7	↓0.004	↑1.33	↓25.4	↓0.004	↑0.62	↓21.8
coastguard	↓0.002	↑0.63	↓30.7	↓0.001	↓0.43	↓32.4	↓0.006	↑1.29	↓25.3
news	↓0.02	↑0.9	↓24.3	↓0.01	↑0.78	↓27.5	↓0.007	↑0.8	↓20.4
akiyo	↓0.016	↑1.1	↓21.4	↓0.01	↑0.82	↓23.8	↓0.014	↑1.06	↓25.7
hall	↓0.018	↑1.7	↓22.7	↓0.014	↑1.05	↓24.1	↓0.002	↑0.89	↓23.4
container	↓0.012	↑0.51	↓21.5	↓0.008	↑0.43	↓24.7	↓0.01	↑0.75	↓23.2

理方法，其中，向上的箭头代表增加、向下的箭头代表下降。从表 3 的实验数据可见，本文算法与 JM10.1 相比，编码时间减少了 41.3% ~ 53.4%，PSNR 变化很小，对人的视觉来说几乎不影响图像质量；本文算法与快速三步法相比，编码时间减少达 21.4% ~ 32.4%，PSNR 变化很小。另外，本文算法产生的码率相对于 JM10.1 有些增加，但和三步法相比，码率基本不变。

4 结语

本文从图像自身的特征及预测模式的统计特性，提出了一种高效简单的快速帧内预测算法。从以上的分析中可以看出，本文所提出的快速算法在保证码流结构不变和码率增大极少的情况下，较大提高预测速度，很大程度地减少了所需计算的模式数目，因此是种行之有效的快速算法，具有一定的实际应用价值。

参考文献

- 1 Marpe D, Schwarz H, Wiegand T. Context-Based adaptive binary arithmetic coding in H.264/AVC video compression standard. IEEE Transaction on Circuits and Systems for Video Technology, 2003,13:620 - 636.
- 2 Su R, Liu GH, Zhang TY. Fast mode decision algorithm for intra prediction in H.264/AVC with integer transform and adaptive threshold. Proc. of IEEE International Conference on Acoustics, Speech and Sins Processing. 2006.11 - 27.
- 3 Liu Q, Hu RM, Zhu L. Improved fast intra prediction algorithm of H.264/AVC. Journal of Zhejiang University SCIENCE A, 2006.101 - 105.
- 4 Chao CC. Fast three step intra prediction algorithm for 4x4 blocks in H.264. Proc. of ISCAS 2005.
- 5 冯晓茹,杜建超,周有喜,吴成柯.基于 H.264 的快速帧内预测算法.计算机工程, 2008,34(2):299 - 231.
- 6 Cheng Y, Wu M. A fast mode decision algorithm of 4x4 block intra prediction for H.264/AVC. Proc. of 3rd International Conference on Intelligent System and Knowledge Engineering. 2008.1231 - 1236.