一种面向 H.264/AVC 的快速帧内预测选择算法

孟庆磊 姚春莲 宋建斌 李 炜

(北京航空航天大学 计算机学院，北京 100083)

摘 要：为提高 I 帧的编码效率，最新的视频压缩编码标准 H.264/AVC，在空间域上从多个方向进行多种模式的帧内预测，并使用率失真优化算法选择最佳模式，算法复杂度很高.为此，提出了一种快速的帧内预测算法.针对亮度块的模式选择过程，提出了基于图像熵值和滑动窗口机制自适应地调整熵值阈值的方法对宏块模式进行预判；针对亮度模式中多个预测方向的选择，充分利用中间结果，预先排除一些可能性小的预测方向；在亮度块和色度块相结合的率失真统计中，改变算法结构，预先计算色度块最佳模式.实验结果表明，算法在不降低图像质量的基础上，编码速度平均提高 76%，码率平均增加 2.6%.

关 键 词：H.264/AVC；帧内预测；图像熵；预测模式；相关性

中图分类号：TP 391

文献标识码：A **文 章 编 号**：1001-5965(2007)02-0219-05

Fast selective algorithm of Intra prediction for H.264/AVC

Meng Qinglei Yao Chunlian Song Jianbin Li Wei

(School of Computer Science and Technology, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: The newest video coding standard H.264/AVC provides several predicting modes and directions for Intra prediction in the spatial region, and the optimal mode was selected by rate-distortion optimization(RDO) with high coding complexity. In order to solve this problem, the fast selection algorithm was proposed from different aspects: Firstly, for luminance mode selection, the method was proposed that utilizes the image entropy value to predict the mode of the macroblock, and updates the threshold adaptively by a sliding window. Secondly, intermediate computing results were reused to reduce the predicting directions. Finally, in the RDO performance, the arithmetic structure of Intra prediction was changed with calculating the best chromance mode separately. Experimental results show that with similar distortion performance, the algorithm that adopt the above three strategy can efficiently reduce the computational complexity up to 76% with 2.6% increase of bits rate averagely.

Key words: H.264/AVC; Intra prediction; image entropy; prediction mode; correlation

ITU 与 ISO 在 2001 年底成立联合专家组 JVT(Joint Video Team)，于 2003 年 7 月完成了目前压缩效率最高的视频编码标准 H.264/AVC[1].它采用传统的基于变换和运动补偿的混合编码框架，在延用以前标准中优秀技术的同时，在编码关键模块中使用了一些先进的技术：多尺寸块的帧内和帧间编码、4×4 整数正交变换、环内滤波器，更精细的运动估计与补偿和新的熵编码等.与H.263和 MPEG-4(简单类)标准相比，在相同恢复图像质量下一般可以节约 40%～50%的码率，但运算复杂性增加了十几倍[2].

H.264/AVC 的帧内预测是基于图像子块在空间域的相关性进行预测的，预测精度较 H.263+[3]与 MPEG-4[4]所采用的频率域预测算

收稿日期：2006-03-21

基金项目：国家自然科学基金资助项目(60505007)；新世纪优秀人才支持计划资助项目；国防基础科研资助项目

作者简介：孟庆磊(1980－)，男，辽宁沈阳人，博士生，mql198029@gmail.com.

法有显著提高.在标准中,H.264/AVC在空间域上从多个方向进行多种模式帧内预测,并使用率失真优化算法,扫描所有模式以及该模式下的方向,因此其计算复杂度非常大,限制了该技术在实时视频领域中的应用,必须研究相应的快速算法.文献[5－6]利用空间域或变换域的某种特征,预先排除一些可能性小的预测模式,每次只计算3～4个方向.文献[7]根据邻块间预测选项的相关性、预测选项和SAD(Sum of Absolute Difference)相关性的概念,利用"半步停"技术自适应地中止预测,计算复杂度仅为原有算法的7%,但是图像质量下降比较多.文献[8]利用二维直方图在预测前直接判断宏块模式,但算法只是根据固定的阈值进行预判,没有合理的更新机制,其预判宏块模式的正确率不高.

本文在详细分析帧内预测的基础上,从3个方面对算法提出了改进策略:针对亮度块的模式选择,提出了一种基于图像熵值和滑动窗口机制自适应地调整熵值阈值的方法对宏块模式进行预判;针对亮度块的多个预测方向,充分利用中间结果,预先排除一些可能性小的方向;在亮度块和色度块相结合的率失真统计中,改变算法结构,预先计算色度块最佳模式.最终通过三者的结合,提出了一种新的适用于H.264/AVC的快速帧内预测算法,在有效地降低帧内预测的运算复杂度的同时,仍然能够维持较好的图像质量和压缩性能.

1 H.264/AVC帧内预测算法

图像的纹理特性,按其方向性可以划分为水平、垂直、对角线和平面方向等.将图像分割成块后,相邻子块间往往具有一致的纹理,各像素之间具有相似性.帧内预测就是基于这种子块间的纹理特性和空间域像素间的相关性进行预测.

1.1 宏块预测模式

H.264/AVC亮度编码块有2种划分方式,分为16×16模式和4×4模式,色度分量为8×8模式.针对16×16模式设定了水平(V)、垂直(H)、均值(DC)和平面(Plane)4种预测方向,如图1所示.对于细节较多、变化剧烈的宏块,各像素点间的差别较大,需要将宏块分为16个4×4大小的子块,以减少预测误差,该模式共有9种(0,1,2,…,8)预测方向,见图2,其中方向2是均值预测.块内各像素点预测值的计算公式见文献[1].

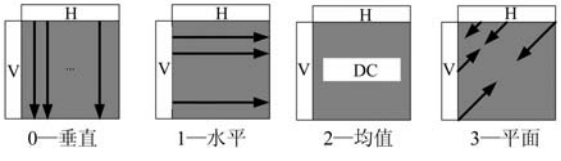
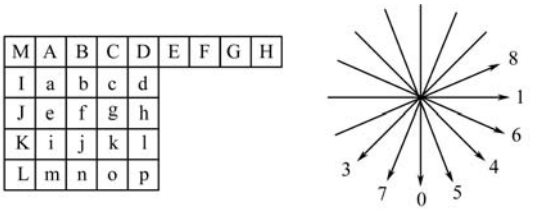


图1 16×16模式的预测方向

H.264/AVC对色度分量单独进行预测,考虑到人眼对色度分量的敏感度不如亮度分量,H.264/AVC对色度分量采用一种精度较低的方法,仅有8×8模式,预测方向为4种,与16×16模式的4种预测方向一致.



a 4×4块像素分布 b 9种预测方向模式示意图

图2 4×4模式示意图

1.2 宏块模式的选择算法

H.264/AVC采用亮度和色度相结合的RDO(Rate-Distortion Optimization)判决准则,对所有的预测模式以及该模式下的方向进行循环,利用拉格朗日代价函数选出最优的预测模式,见式(1):

$$J_{\mathrm{MODE}} = D_{\mathrm{MODE}} + \lambda \times R_{\mathrm{MODE}} \qquad (1)$$

式中,D_{MODE}和R_{MODE}表示MODE模式的失真和码率;λ为拉格朗日参数.选择使编码后的失真和码率总和J_{MODE}最小的模式MODE.

校验模型Jm94实现了这种全搜索的算法,它以色度模式为外循环,然后依次扫描亮度的所有模式,搜索的总数为(9×16+4)×4=592次.这意味着为了得到一个宏块的最佳模式,需要计算592个不同的J_{MODE},包括每次预测所必需的生成参考块、匹配准则的计算和残差块编码等.

2 基于图像熵值的亮度块模式选择

在编码过程中,依据宏块的某种特征,对于特征明显宏块提前决定最佳预测模式,避免对亮度16×16和4×4两种模式的全搜索,减少不必要的计算;特征不明显的宏块,用标准的算法获得最佳编码模式,获得最好的效果.本文选择图像熵值表征宏块的特征.

2.1 图像熵值的定义及其在统计特性的表现

图像的熵值是一种特征的统计形式,它反映了图像中灰度分布的聚集特征所包含的信息量,令P_i表示图像中灰度值i的像素所占的比例,则熵定义为

$$\left.\begin{aligned} H &= -\sum_{i=0}^{255} P_i \ln P_i \\ \sum_{i=0}^{255} P_i &= 1 \end{aligned}\right\} \tag{2}$$

图像的熵值也能表现出图像块内信息量的变化,即图像细节的丰富程度.如果熵值比较大,可认为该图像的细节很丰富,图像中的灰度变化比较剧烈,反之亦然.

表 1 是 QCIF 格式(176×144)的 Paris(纹理复杂)、Claire(平坦)和 Container(中等)序列中采用全搜索算法编码第一帧图像的熵值均值,括号中是采用该模式的宏块数,帧中 16×16 模式的宏块如图 3 中粗实方框所示.对比发现,采用 16×16 模式的宏块大部分为平坦区域块.通过表 1 的统计可以看出,不同模式宏块的熵值有较明显的差别,由此可见熵值在一定程度上确实反映了宏块的特征.本算法就是利用熵值对宏块变化的反映进行预判,也是本文提出的快速模式选择算法的基础.

表 1　图像宏块的熵值数据

序列	熵值均值		宏块数
	16×16	4×4	
Container	2.51(37)	3.82(62)	熵值低于 2.51 时为 22.2% 熵值超过 3.82 时为 40.4%



图 3　宏块模式示意图

2.2　基于双阈值的宏块模式判定

表 1 的结果表明,熵值相对较小时,宏块很可能主要由一种灰度组成,有可能为平坦区域;细节较多、变化剧烈的区域,其灰度值比较分散,熵值相对大些.基于这种特点,本文采用双阈值法来判断,具体过程如下:

1) 预先设定 2 个阈值:上限 α 和下限 β.

2) 计算当前宏块的熵值 T_c,与 α,β 进行比较.大于 α,则认为该宏块细节较多,采用 Intra 4×4 模式;小于 β,则认为该宏块较平坦,采用 Intra 16×16 模式;介于 α 与 β 之间,采用 H.264/AVC 的全搜索算法,获得最佳的模式.

3) 保存宏块的熵值 T_c,对该模式的阈值 α 或 β 进行更新.

2.3　基于滑动窗口机制的阈值更新策略

为了使阈值能够准确地反映视频内容的变化,阈值的更新成了关键,其过程包括 3 个步骤:

步骤 1(采集数据)

最新的数据总是在数据集当中,在得到当前宏块的熵值 T_c 后,根据宏块的模式将其赋值给 α_0 或 β_0,同时计算出临近宏块的均值 T_{avg}.

步骤 2(数据点选择)

所谓的数据点:就是被用作更新阈值的数据集,该数据集的质量与数量对于阈值的精度非常重要.本文给出了一个基于滑动窗口的数据选择机制,该机制能够自适应地为阈值更新过程选择具有代表性的数据点.

本文使用熵值度量宏块内容的变化,如果相邻的宏块有较大的运动场景(即临近宏块的相关性较差),只有最近的少量数据点被选择;反之,较多的数据点被选中.因此滑动窗口的大小为

$$n = \begin{cases} \dfrac{T_c}{T_{avg}} \cdot W_{max} & T_{avg} > T_c \\ \dfrac{T_{avg}}{T_c} \cdot W_{max} & T_{avg} \leqslant T_c \end{cases} \tag{3}$$

式中 W_{max} 在实验中设置为常数 20.

步骤 3(参数 α 和 β 的计算)

根据被选数据点的数目和实际值,使用式(4)可以得到 2 个新参数 β 和 α.

$$\left.\begin{aligned} \beta &= \left(\varepsilon \sum_{i=1}^{n} \beta_i + (1-\varepsilon) \times \beta_0\right) \times 0.9 \\ \alpha &= \left(\varepsilon \sum_{i=1}^{n} \alpha_i + (1-\varepsilon) \times \alpha_0\right) \times 1.05 \end{aligned}\right\} \tag{4}$$

式中,n 是被选中的宏块数;α_i 或 β_i 分别是各宏块实际的熵值,宏块模式分别对应着 16×16 和4×4 模式,ε 是比例系数,实验中取为 0.9.

2.4 实验结果和分析

为了验证快速算法的实用性和鲁棒性，本文选取标准组织推荐的9个QCIF格式的序列进行测试. 其中Claire和Table纹理简单，运动相对静止；Mthr_dot和Foreman纹理中等，运动相对静止；Container，Coastguard和Salesman纹理复杂度和运动程度居中；Paris纹理复杂，运动相对静止；Football纹理复杂，运动剧烈. 编码100帧，全部采用I帧方式编码，量化参数为28. CPU是AMD3200+，内存512 MB的PC机为实验平台，采用H.264的参考模型JM94，结果见表2，本文以后所有的测试条件与之一致.

从表2中可以看出，自适应快速亮度块模式选择算法不受视频运动类型的影响，保证重构图像质量PSNR(Peak Signal Noise Ratio)值基本不变，具有较强的自适应性. 但是由于图像内容的复杂性和多变性，速度增加从10.5%～19.5%不等，平均提高为14.2%. 表2还给出了实验中各个序列经过本文算法进行预判的宏块数、预判错误的宏块数和正确率. 预判错误的宏块数的表格中，前一个数据表示将4×4模式误判为16×16模式的宏块个数，后一个数据表示另一种误判情况. 由于本文采用了一种新的基于滑动窗口机制的阈值更新策略，阈值设定比较合理，算法的模式预判正确率较高，平均的预判正确率为95.3%.

表2 自适应快速亮度块模式选择算法与原算法的比较

序列	PSNR值	码率/%	速度/%	预判	错误宏块	正确率/%
Claire	+0.01	0.120	18.2	3213	122+22	94.9
Table	0.00	0.003	10.5	2786	72+17	96.8
Mthr_dot	0.00	0.150	19.5	4587	58+15	98.4
Foreman	0.00	0.150	11.6	4141	100+12	97.3
Container	+0.01	0.840	12.3	3857	430+145	85.1
Coastguard	0.00	0.040	16.1	3372	245+7	92.5
Salesman	0.00	0.004	14.3	3317	129+30	95.2
Paris	0.00	0.001	12.5	3261	19+0	99.4
Football	0.00	0.001	13.0	3307	46+27	97.8

3 结合相关性的预测方向选择策略

在实验统计中，超过70%的编码块为4×4模式，帧内预测算法的复杂度主要体现在这个环节上. 为此，本文在4×4模式多个预测方向的选择过程中，首先计算平面、水平和垂直方向，根据中间结果预先排除一些可能性小的预测方向，只是计算潜在的可能对角线方向，在速度和质量间取得较好的平衡. 算法的选择策略如下：

① 首先计算0(垂直)、1(水平)、2(均值)预测方向，计算相应失真值(RDcos t). 如果均值模式的RDcos t 小于阈值 R 时，直接选择DC模式作为4×4的帧内预测模式，转到步骤④；否则到步骤②；

② 计算3和4预测方向，获得相应的失真值；

③ 选取0和1，3和4中较小的两个方向，再计算两者中的夹角方向：若选择的是0和3，则进一步计算方向7；若选择的是0和4，则进一步计算方向5；若选择的是1和3，则进一步计算方向8；若选择的是1和4，则进一步计算方向6；

④ 保留最近的20个DC模式下的RDcos t 值，使用均值来更新阈值 R.

测试表明，亮度块预测方向的改进算法使重构图像PSNR值略有下降，根据图像序列的不同从0.04～0.08 dB不等，说明改进的预测算法具有较高的预测精度，总的编码码率有所增加，运算速度平均提高了22.5%.

4 基于亮度和色度RDO的改进策略

H.264/AVC的RDO判决方式，以色度模式为外循环，然后依次扫描亮度的所有模式，这种方式运算较复杂. 为此，本文将色度部分从循环中提出，单独计算色度块的编码失真 D，选择失真最小的方向作为色度分量的最佳，减少不必要的循环次数. 结合此预测方向，在亮度循环中计算最佳的亮度和色度相结合的最小RDcos t，每个宏块只需计算9×16+4+4=152个不同的 J_{MODE}，相比于原来的592次RDO计算，明显可以降低帧内预测的计算复杂度.

色度块预测方向选择策略如下：

① 对于色度 U 和 V 两个8×8块在各种色度块的预测方向下生成残差块；

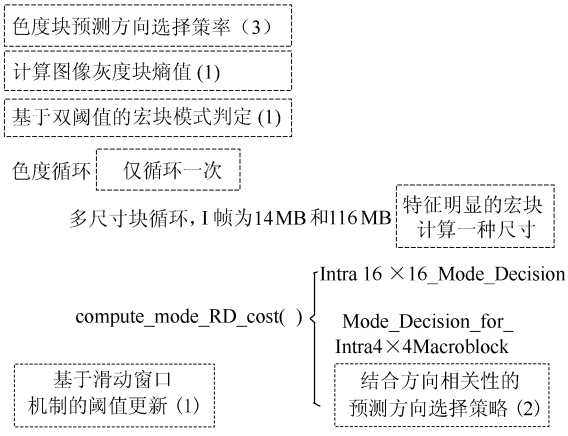
② 以色度块的4种预测方向为外层循环，分别计算 U 和 V 分量的编码块失真 D；

③ 选择编码块失真 D 最小时对应的色度预测方向为色度最佳的结果.

在2.4节描述的试验平台和测试条件下进行测试，与全搜索帧内预测算法相比，9个序列的PSNR_Y、PSNR_U和PSNR_V最大的变化仅为0.04 dB，0.04 dB，0.05 dB，码率增加平均为0.17%，而帧内预测速度平均提高68.9%. 表明基于亮度和色度RDO计算的改进策略能够很好地保持图像质量，并极大提升了编码器速度.

5　帧内预测算法描述和实验结果

本文结合两部分改进策略的帧内预测快速算法,如图 4 所示,给出了 3 个改进部分的作用域,具体细节参见各部分改进的描述.需要指出的是,由于本文算法是对帧内预测的 3 个不同方面分别提出了改进,三者是相互独立的,可以单独拿出一部分和其它方法结合或改进.



注:虚框部分为改进算法,其它为 H.264/AVC 原算法

图 4　快速选择算法的示意图

表 3 给出了本文算法的运行结果,可见,本文的快速帧内预测模式选择算法不受图像序列纹理和运动程度等因素的影响,比 H.264/AVC 原有的算法可提高编码速度,最小提高为75.1%,平均提高为 76%,灰度平均的信噪比只约降低0.05 dB 左右,码率稍有变化,平均增加为 2.6%.

表 3　算法比较结果(Q_P=28)

序列	PSNR_Y	PSNR_U	PSNR_V	码率/%	速度/%
Claire	−0.05	−0.02	0.00	3.5	76.1
Table	−0.04	+0.04	+0.01	2.0	75.1
Mthr_dot	−0.04	0.00	−0.05	3.5	76.5
Foreman	−0.04	−0.02	−0.01	4.3	75.5
Container	−0.03	+0.03	+0.01	2.5	77.4
Coastguard	−0.06	−0.03	−0.04	1.6	76.7
Salesman	−0.08	+0.01	+0.01	2.7	75.0
Paris	−0.05	−0.02	−0.02	1.6	75.8
Football	−0.04	+0.01	+0.02	2.2	75.9

表 4 所示为 *Paris* 和 *Mthr_dot* 序列在取不同量化参数 Q_P 时的测试结果.

表 4　不同 Q_P 的算法性能比较

Q_p	Mthr_dot			Paris		
	PSNR_Y	码率/%	速度/%	PSNR_Y	码率/%	速度/%
20	−0.04	2.0	77.1	−0.06	1.0	75.4
24	−0.04	2.8	77.6	−0.06	1.2	75.3
28	−0.04	3.5	76.5	−0.05	1.6	75.8
32	−0.02	4.7	75.8	−0.03	2.0	75.5
36	−0.02	5.0	75.8	−0.04	2.8	75.5

结果表明快速算法的性能在低、中和高码率下具有一致性,在编码时间明显减少的同时,R-D 性能相差不多.

6　结 束 语

在新一代视频编码标准 H.264/AVC 中,采用帧内预测来提高帧内编码效率.标准算法在空间上采用不同的模式和多个方向以及亮度块和色度块相结合的方式进行预测,然而全搜索策略和率失真优化的计算复杂度相当高.为了满足实时应用的需求,本文在详细分析了帧内预测的基础上,对算法的 3 个部分分别进行了改进.实验结果表明,采用这 3 种策略的快速算法在几乎不牺牲图像质量和压缩效率的基础上,编码速度平均提高 76%,有效地减少了计算复杂度.

参考文献(References)

[1] ISO/IEC FDIS 14496-10. Information technology-coding of audio-visual objects part 10: advanced video coding[S]

[2] Ravasi M, Mattavelli M, Clerc C A. Computational complexity comparison of MPEG4 and JVT codecs[C]// Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG Meeting. Klagenfurt, Austria:[s.n.],2002

[3] ITU-T. Draft ITU-T recommendation H.263 version 2: video coding for low bit rate communication[S]

[4] ISO/IEC 14496-2. Committee draft, information technology coding of audio visual objects: visual[S]

[5] Pan Feng, Lin Xiao, Susanto R, et al. Fast mode decision for Intra prediction[C]// JVT Meeting. Pattaya, Thailand:[s.n.],2003

[6] Kim Changsung, Shih H H, Kuo C J. Multistage mode decision for Intra prediction in H.264 codec[J]. Proc of Visual Communications and Image Processing, 2004, 5308 (18): 355−363

[7] 倪伟,郭宝龙,杨镪.一种适用于 H.264/AVC 的自适应空域帧内预测算法[J],计算机学报, 2005, 28 (11):1853−1860
Ni Wei, Guo Baolong, Yang Liu. An adaptive Intra prediction algorithm in spatial domain space for H.264/AVC[J]. Chinese Journal of Computers, 2005, 28 (11):1853−1860 (in Chinese)

[8] 谢晶,贾克斌.一种基于二维直方图的 H.264/AVC 快速帧内预测判决算法[J].电子与信息学报, 2005, 27(7): 1053−1057
Xie Jing, Jia Kebin. A fast Intra-frame prediction algorithm based on the two-dimensional histogram for H.264/AVC [J]. Journal of Electronics & Information Technology, 2005, 27(7):1053−1057 (in Chinese)

(责任编辑:张　嵘)