

一种新的 H.264/AVC 快速帧内预测模式选择判决算法

刘 佳 董金明

(北京航空航天大学电子信息工程学院 北京 100083)

摘 要: 针对 H.264 视频编码标准帧内预测模式选择部分计算量大的问题, 本文提出了一种快速帧内预测模式选择判决算法。该算法首先提出了一个新的算子, 用来描述图像灰度分布信息, 作为判断图像复杂程度的依据, 进而提出两个对量化系数(QP)自适应的阈值, 将图像分为平滑、变化剧烈和特征模糊 3 种。平滑图像宏块使用 Intra_16×16 模式预测, 变化剧烈的宏块使用 Intra_4×4 模式预测, 特征模糊的宏块使用标准算法预测。实验结果表明, 该优化算法能在保证很好的图像质量的同时, 使帧内模式选择计算次数减少 30%以上, 而且对传输码率基本没有影响。

关键词: H.264; 帧内预测; 模式选择; 自适应阈值

中图分类号: TN919.8

文献标识码: A

文章编号: 1009-5896(2009)05-1072-05

An Improved Intra-frame Prediction Mode Selection Algorithm for H.264/AVC

Liu Jia Dong Jin-ming

(School of Electronic and Information Engineering, Beihang University, Beijing 100083, China)

Abstract: In this paper, to simplify the computational complexity of intra-frame prediction mode selection for H.264, an improved Intra-frame prediction mode selection algorithm is proposed. Firstly, a new operator is proposed to describe the gray-level info of an image which could be the image complexity degree criterion. Then two QP (Quality Parameter)-based adaptive thresholds are proposed and images can be divided into three kinds: smooth, variability and in fuzzy feature. Intra_16×16 mode is used when the macroblock is smooth, Intra_4×4 mode is used when the macroblock is variability and the original algorithm is used when the macroblock is in fuzzy feature. Experimental results show that the improved algorithm almost has 30% decrement on intra-frame mode selection times, meanwhile the image quality is still good enough with a neglectable rate increase.

Key words: H.264; Intra-frame prediction; Mode selection; Adaptive threshold

1 引言

H.264/AVC 视频编解码标准是由 VCEG(视频编码专家组)和 MPEG(运动图像专家组)组成的 JVT(联合视频组)共同推出的^[1]。其编解码复杂度极高, 在不改变码流结构和其优秀的压缩性能的前提下, 寻找简单易行的优化算法来替代已有标准算法成为当前研究的热点。文献[2]提出了一种新的帧内预测模式选择算法——FIPAMP; 文献[3]针对帧内预测模式选择不当而在帧内编码过程中产生闪动的现象, 提出一种帧内预测模式选择方法, 通过使用修正的失真方法, 对率失真优化模式选择部分不仅考虑输入和重建宏块间的失真, 考虑两个连续帧间的失真来减少帧内编码帧的闪动; 文献[4]用直方图中频数最大值来判断图像特征继而进行帧内选择的优化。他们在优化中并没有考虑图像宏块本身的特征。文献[5]通过求出图像的直流与交流分量的比值的方法来判断图像复杂程度; 文献[6]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。文献[5,6]所提出的描述图像复杂程度的方法都有一定的缺陷, 本文第3节将详细讨论。针对以上这些问题, 本文提出了一种能准确描述图像宏块复杂程度的算子对帧内预测过程进行了优化。

本文第2节对帧内预测技术进行了深入分析; 第3节给出了新算法的具体思路和实现方法; 第4节列出实验数据并给予了详细分析, 证明了新算法的有效性; 最后为结束语。

2 H.264 标准的帧内编码技术

视频编码的目的是去除视频序列的时间相关性和空间相关性。帧内预测技术用于去除图像空间相关性。H.264 标准中帧内亮度 16×16 块模式(下称 Intra_16×16 模式)包括 4 种预测模式, 帧内亮度 4×4 块模式(下称 Intra_4×4 模式)包括 9 种预测模式。色度块也有 4 种预测模式, 但是, 因为色度块不存在模式选择的问题, 所以本文的模式选择只针对亮度块进行, 即讨论当前宏块的亮度分量是使用一个 Intra_16×16 模式进行预测还是用 16 个 Intra_4×4 模式进行预

测^[7]。

帧内预测时, 预测块 P 的像素 a, b, c, \dots, p 可以由宏块左边和上边的 13 个已经编码重建的像素点 ($A \sim H, I \sim L, M$) 的不同加权和来预测, 如图 1 所示。

Intra_4×4 模式的 9 种模式预测方向如图 2(不包括 DC 模式), Intra_16×16 模式的 4 种预测模式参见文献[8], 本文不再赘述。

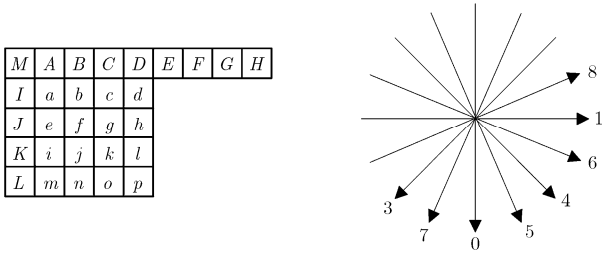


图 1 预测样点的标注 图 2 Intra_4×4 模式的 9 种预测方向

在 H.264 标准中, 帧内预测模式的选择通过计算拉格朗日开销来确定, 拉格朗日开销计算式如式(1)所示:

$$J = SAD + \lambda_{MODE} \cdot R \tag{1}$$

其中 J 是拉格朗日开销, SAD 是预测后的残差的平方和, R 是指在某一种 MODE 下的编码码率, λ_{MODE} 是一个跟量化系数(QP)相关的系数^[9]。拉格朗日开销最小的模式就作为预测的最佳模式。所以对于最佳模式, 要求 SAD 和 R 都尽可能小。

而实际上, SAD 和 R 是此消彼长的关系, 如果选择 Intra_4×4 模式, 则预测所得到的残差 SAD 就较小, 但是对于每一个 4×4 块, 都需要有一定的码长来表征运动矢量等特征, 这样, R 就较大; 如果选择 Intra_16×16 模式, 则整个宏块只有一个运动矢量需要表征, R 较小, 但是所得到的残差 SAD 却相对较大, 所以整个宏块的拉格朗日开销最小就是要在这两个值之间有一个权衡^[10]。在 H.264 标准中明确指出, Intra_16×16 模式适用于图像内容变化平滑的区域, 言下之意, Intra_4×4 模式适用于图像内容变化剧烈的区域。本文利用这个特点, 首先根据所提出的算子计算判断出宏块是平滑还是变化剧烈, 对于平滑的宏块, 选择 Intra_16×16 模式; 对于变化剧烈的宏块, 选择 Intra_4×4 模式; 对于特征模糊的宏块, 则需要对两种模式都计算拉格朗日开销, 选择开销最小的那个模式作为最佳预测模式。

3 算法基本思路和具体方法

基于上面的分析, 可以看到, 选择 Intra_4×4 模式还是 Intra_16×16 模式理论上讲是通过计算拉格朗日开销来判断, 但是如果图像特征(平滑或者变化剧烈)明显, 则可以根据图像特征直接判断出选用 Intra_4×4 模式还是 Intra_16×16 模式预测, 计算量将大大减少。于是, 计算拉格朗日开销的问题转化为判断图像宏块复杂度的问题。

3.1 对已有表征图像复杂度方法的分析

当前在表征图像宏块复杂度方面主要有通过提出计算交流系数 AC 与直流系数 DC 的比值来判断该宏块的复杂度^[5]和用图像的信息熵来判断该宏块的复杂度^[6]两种方法。AC, DC 系数和信息熵的计算公式如式(2)所示, 其中 AC, DC, $H(X)$ 分别为 AC 系数, DC 系数和信息熵。

$$DC = \frac{1}{256} \left(\sum_{m=0}^{15} \sum_{n=0}^{15} p_{mn} \right)^2, \quad AC = \left(\sum_{m=0}^{15} \sum_{n=0}^{15} p_{mn}^2 \right) - DC,$$

$$H(X) = - \sum_{i=1}^q p(x_i) \log p(x_i) \tag{2}$$

分析式(2)可知, AC 与 DC 的比值表征图像的细节信息, 而信息熵只能反映图像各灰度值所提供的信息量, 某些情况下它们不能准确表征图像复杂程度。例如图 3 所示的 3 个灰度宏块, 宏块图 3(a)上面 8 行的灰度值均为 255, 下面 8 行的灰度值均为 0; 宏块图 3(b)上面 8 行和下面 8 行的灰度值分别为 127 和 128; 宏块图 3(c)上面 8 行和下面 8 行的灰度值分别为 1 和 2。

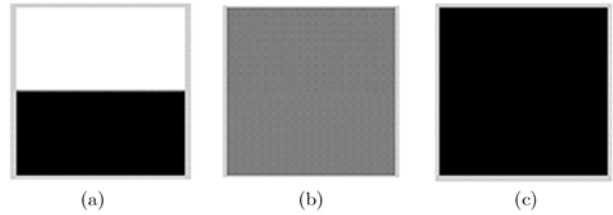


图 3 用作对比的 3 个宏块

计算宏块图 3(a), 3(b), 3(c)的 AC, DC 和熵值如表 1 所示。

表 1 宏块图 3(a), 3(b), 3(c)的 AC, DC 和熵值计算结果

宏块名称	AC	DC	logAC/logDC	$H(X)$
图 3(a)	4161600	4161600	1	log2
图 3(b)	64	4161600	0.2729	log2
图 3(c)	64	576	0.6543	log2

根据以上结果, 如果用信息熵来判断图像宏块复杂度, 则 3 个宏块复杂度一致; 如果用 AC, DC 系数比值来判断图像复杂度, 则宏块图 3(b), 3(c)复杂度不一致。事实上, 宏块图 3(a)的复杂度明显大于图 3(b)和 3(c), 而图 3(b)和 3(c)的复杂度应该是一致的。

所以, 用 AC, DC 系数比值或者信息熵的方法在某些情况下不能准确表征图像的复杂度。

3.2 一种快速帧内预测模式选择判决算法

研究发现, 变化剧烈的宏块的灰度分布分散, 各个灰度值的频数都很小; 而平滑宏块的灰度分布集中, 常常有某一个或者某几个灰度值的频数很大。本文提出一个算子, 如式

(3)所示，通过提取图像宏块的灰度分布信息来判断其复杂度，进而选择采用 Intra_4×4 模式或 Intra_16×16 模式进行帧内预测。

$$Z = \sum_{i=0}^{255} E(i, n) \cdot p_i \quad (3)$$

其中 Z 为本文提出的能表示图像复杂程度信息的算子， $E(i, n)$ 是加权因子，用于对不同灰度值频数加权， i 是当前的灰度值， n 是频数最大值所对应的灰度值， p_i 是灰度值 i 的频数。考虑到 4 个基本图像类型(暗、亮、低对比度和高对比度)的灰度分布情况^[11]，本文选择一个类正态分布的函数作为 $E(i, n)$ ，如图 4(a)所示，通过 $E(i, n)$ 对亮度分量灰度值的加权，使得处于各个灰度值位置的灰度值频数对 Z 的贡献量各不相同，其计算方法如式(4)所示：

$$E(i, n) = \exp\left\{-\frac{(i-n)^2}{1000}\right\} \quad (4)$$

为了便于清楚地表述和方便地判断，将 Z 归一化为 Z_0 ，如式(5)所示：

$$Z_0 = \frac{Z}{256}, \quad \frac{1}{256} < Z_0 \leq 1 \quad (5)$$

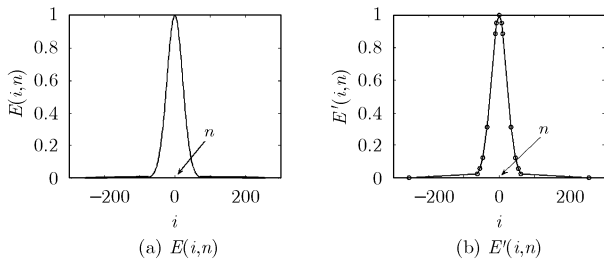


图 4 加权函数 $E(i, n)$ 及其简化计算函数 $E'(i, n)$

归一化以后，复杂度低的宏块的 Z_0 值较大，复杂度高

的宏块 Z_0 值较小。当 Z_0 等于 1 时，图像最平滑，整个宏块只有一个灰度值并且它的频数是 256。

$E(i, n)$ 中存在幂运算，实际计算时计算量很大，为了解决这个问题，本文将该正态分布函数分段线性表示，如图 4(b)所示，采用线性插值求解各点的函数值。

将 i 分为 7 个线性区间，如表 2 所示。

$E(i, n)$ 的简化计算步骤如下：首先判断 i 所在灰度值区间 $[i_m, i_{m+1}]$ ，其中 $(0 \leq m \leq 6)$ ，然后查找表 2 得到 E_m, E_{m+1} ，即可用线性插值的方法得到 $E(i, n)$ 的近似值 $E'(i, n)$ ，如式(6)所示，这样把指数运算转化为加减乘除运算，大大减小了 $E(i, n)$ 的计算量。

$$E'(i, n) = \frac{(E_{m+1} - E_m)}{(i_{m+1} - i_m)} \times (|i - n| - i_m) + E_m \quad (6)$$

为判断图像类型，本文设定两个阈值， T_0 和 T_1 ($T_0 < T_1$)，当所计算得到的 Z_0 大于 T_1 ，认为图像平滑，使用 Intra_16×16 模式；当 Z_0 小于 T_0 时，认为图像变化剧烈，使用 Intra_4×4 模式；当 Z_0 处于 T_0 和 T_1 之间时，认为图像特征模糊，使用标准算法。因为阈值的选择与 QP 大小有关，QP 大，则可以被判决为 Intra_16×16 模式的宏块就比较多，可以被判决为 Intra_4×4 模式的宏块就比较少，反之亦然，即 T_0, T_1 与 QP 成反比。在不考虑率优化失真时， T_0, T_1 与 QP 线性相关。同时，因为 QP 的范围是 $\{0, 51\}$ ^[12]，所以阈值可以用式(7)来表示。

$$T_0 = a\left(1 - \frac{QP}{52}\right), \quad T_1 = b\left(1 - \frac{QP}{52}\right) \quad (7)$$

其中 a, b 为两个线性系数，本文选取 QP 为 20。为了确定 a, b ，将 10 个视频序列的前 100 帧使用帧内编码，得到在标准算法中采用 Intra_16×16 模式和 Intra_4×4 模式的宏块所对应的 Z_0 值的平均值，如表 3 所示。

表 2 i 的 7 个灰度值区间

$ i $	$ i_0 =0$	$ i_1 =7$	$ i_2 =11$	$ i_3 =34$	$ i_4 =46$	$ i_5 =54$	$ i_6 =62$	$ i_7 =255$
$E(i, n)$	E_0	E_1	E_2	E_3	E_4	E_5	E_6	E_7
	1.0000	0.9521	0.8860	0.3147	0.1205	0.0541	0.0214	0.0000

表 3 使用不同块大小进行帧内预测的宏块的 Z_0 平均值

序列名称	Intra_16×16	Intra_4×4
akiyo	0.966138	0.695486
carphone	0.971622	0.648942
claire	0.983297	0.682232
container	0.974149	0.575062
grandma	0.967586	0.696499
mobile	0.950406	0.686517
mother-daughter	0.961886	0.657724
news	0.970738	0.643818
salesman	0.953116	0.740202
silent	0.950717	0.726533

观察两组数据，发现这 10 个视频序列中使用 Intra_16×16 的宏块的 Z_0 的平均值都大于 0.95，使用 Intra_4×4 的宏块的 Z_0 都小于 0.75，为了不至于产生误判，我们给出一定的余量，选择 $T_0 = 0.8, T_1 = 0.93$ ，此时， $a = 1.3, b = 1.5$ 。

利用本算子对图 3 中的 3 个宏块的复杂度进行判断，计算得到图 3(a), 3(b), 3(c)的 Z_0 分别为：0.129011, 0.996584, 0.996584。从 Z_0 的结果看出，宏块(a)比(b), (c)复杂，宏块(b)和(c)复杂度一致，这与实际相符，可见，本算子在 AC 与 DC 比值和熵值不能准确判断图像宏块复杂度的情况下也是有效的。

4 实验数据和分析

为了验证本文提出的优化算法的有效性, 我们做了大量的实验。

对 10 个视频序列的前 100 帧采用 I 帧编码, 得到的实验结果如表 4 所示, 因为本文选取的视频序列都是 QCIF 类型的, 所以每帧图像包括 99 个宏块, 标准算法中每个宏块都

要计算一次 Intra_16×16 模式的开销和 16 次 Intra_4×4 模式的开销, 故使用原算法所需要的开销计算次数 $N_{\text{bef}} = 100 \times (99 + 99 \times 16) = 168300$ (QP 为 20)。

表 4 中, N_{16} 表示使用本文优化算法判断出使用 Intra_16×16 模式的宏块个数, N_4 表示使用本文优化算法判断出使用 Intra_4×4 模式的宏块个数, N_{now} 表示现在开销计算

表 4 实验结果

序列名	N_{16}	N_4	N_{now}	ΔN	R_{bef}	R_{now}	ΔR	ΔPSNR
akiyo	4786	1600	90124	-0.4645	425247	455662	0.071523	0.008375989
carphone	3604	2417	108219	-0.35699	462660	482631	0.043166	0.004256325
claire	5010	3246	84894	-0.49558	306703	309862	0.0103	0.004256272
container	4273	2838	97094	-0.42309	584850	607536	0.038789	0.008967523
grandma	5899	1019	72897	-0.56686	535880	569003	0.06181	0.013477738
mobile	468	5353	155459	-0.0763	1378188	1382225	0.002929	0.001003261
mother-daughter	4015	1644	102416	-0.39147	409868	438719	0.070391	0.009226402
news	2500	2774	125526	-0.25415	589838	605106	0.025885	0.004051478
salesman	1926	1539	135945	-0.19225	687301	701907	0.021251	0.00687454
silent	2561	1510	125814	-0.25244	659626	686547	0.040813	0.00613196
平均值	—	—	—	-0.34736	—	—	0.038686	0.006662149

次数, ΔN 表示开销计算次数的减少百分比, R_{bef} 表示原来的传输码率, R_{now} 表示现在的传输码率, ΔR 表示传输码率增大的比率, ΔPSNR 表示本算法的亮度分量信噪比相对于标准算法的亮度分量信噪比的增加率, 其中 N_{now} , ΔN , ΔR 和 ΔPSNR 的计算公式如式(8)和式(9)所示。

$$N_{\text{now}} = N_{16} + N_4 \times 16 + 17 \times (9900 - N_{16} - N_4) \quad (8)$$

$$\Delta = \frac{\text{本算法} - \text{标准算法}}{\text{标准算法}} \quad (9)$$

可以看到, 采用新算法后, 图像大部分区域较平滑的视频序列有很多宏块能够被直接判断为 Intra_16×16 模式, 避免了大量的 Intra_4×4 模式开销的计算, 开销计算次数能够有较大的降幅, 例如 claire 序列开销计算次数的下降接近 50%, 而对于大部分区域变化较剧烈的视频序列或者大部分区域特征模糊的视频序列, 开销计算次数也有比较大的下降, 计算次数的平均下降率达到了 30% 以上。同时, 传输码率的增大几乎可以忽略: 对于 mobile 和 claire 这些采用 Intra_4×4 模式较多的序列, 传输码率的增大微乎其微; 对于 grandma 和 akiyo 这些判断后采用 Intra_16×16 模式较多的序列, 传输码率增大率平均值也只有 0.061, 这表明新算法带来的判断错误极少。而这 10 个视频序列的信噪比都有小幅度的上升, 这说明新算法可以提供较好的图像质量。

5 结论

本文深入分析了 H.264 标准中帧内预测算法的技术细

节, 针对模式选择计算量大的问题提出了一种基于图像宏块本身特征的优化算法。本算法提出了一个新的算子, 用来提取图像灰度分布信息并将其作为判断图像复杂度的标准; 提出了两个与 QP 相关的自适应阈值作为判断图像宏块所属类别的依据, 不同类别采用不同预测模式, 并给出 QP 为 20 时两个自适应阈值的经验值; 最后做了大量实验进行验证。实验结果证明这种算法能够在不改变图像质量和传输码率的增大可忽略的前提下, 大大减少开销计算次数, 有效减小帧内模式预测选择算法的复杂度。

参考文献

- [1] Richardson I E G. H.264 and MPEG-4 Video Compression [M]. England: John Wiley & Sons Ltd, 2004: 159-164.
 - [2] Yang Chun-ling, Po Lai-nan, and Lam Wing-hong. A fast H.264 intra prediction algorithm using macroblock properties [C]. IEEE International Conference on Image Processing, Singapore, 2004: 461-464.
 - [3] Chun Seong-Soo, Kim Jung-Rim, and Sull Sang-hoon. Intra prediction mode selection for flicker reduction in H.264/AVC [C]. IEEE Trans. on Consumer Electronics, 2006, 52(4): 1303-1310.
 - [4] 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 and Jia Ke-bin. A fast intra-frame prediction algorithm based on two-dimensional histogram for H.264/

- AVC[J]. *Journal of Electronics & Information Technology*, 2005, 27(7): 1053-1057.
- [5] Lin Yu-Kun and Chang Tian-Sheuan. Fast block type decision algorithm for intra prediction in H.264 FRExt[C]. IEEE International Conference on Image Processing, Genova, 2005: 585-588.
- [6] 孟庆磊, 姚春莲, 宋建斌, 李炜. 一种面向 H.264/AVC 的快速帧内预测选择算法[J]. 北京航空航天大学学报, 2007, 33(2): 219-223.
- Meng Qing-lei, Yao Chun-lian, Song Jian-bin, and Li Wei. Fast selective algorithm of Intra prediction for H.264/AVC[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2007, 33(2): 219-223.
- [7] Wang Zheng-ning, Yang Jun, Peng Qiang, Ma Zheng, and Zhu Chang-qian. A fast transform domain based algorithm for H.264/AVC intra prediction[C]. IEEE International Conference on Multimedia and Expo (ICME2007), Beijing, 2007: 1563-1566.
- [8] 欧阳合, 韩军. H.264 和 MPEG-4 视频编码 [M]. 长沙: 国防科技大学出版社, 2004: 185-190.
- Ouyang He and Han Jun. H.264 and MPEG-4 Video Compression [M]. Chang Sha: National University of Defense Technology Press, 2004: 185-190.
- [9] 韩青, 李莉, 应骏. H.264 帧内 4×4 块预测模式选择快速算法研究[J]. 中国图象图形学报, 2007, 12(10): 1745-1748.
- Han Qing, Li Li, and Ying Jun. On fast intra prediction mode decision algorithm for 4×4 blocks in H.264[J]. *Journal of Image and Graphics*, 2007, 12(10): 1745-1748.
- [10] 余成伟, 陆建华, 郑君里. H.264 编码模式选择快速算法研究[J]. 清华大学学报, 2007, 47(10): 1677-1680.
- Yu Cheng-wei, Lu Jian-hua, and Zheng Jun-li. Research on fast mode decision algorithm for H.264[J]. *Journal of Tsinghua University (Sci & Tech)*, 2007, 47(10): 1677-1680.
- [11] Gonzalez R C. Digital Image Processing[M]. Second Edition. London: Prentice Hall, 2002: 71-75.
- [12] 沈兰荪, 卓力. 小波编码与网络视频传输[M]. 北京: 科学出版社, 2005: 239-243.
- Shen Lan-sun and Zhuo Li. Wavelet Coding and Internet Video Transmission[M]. Beijing: Science Press, 2005: 239-243.
- 刘 佳: 女, 1984 年生, 硕士生, 研究方向为视频编码处理技术.
- 董金明: 男, 1941 年生, 教授, 主要研究领域为微波技术、天线罩设计、图像视频编码处理技术、嵌入式系统和集成电路等领域的研究.