

# 基于平滑滤波的 H. 264-I 帧模式决定策略

余冰 张太镒

(西安交通大学, 电子与信息工程学院, 西安 710049)

**摘要** 为了降低 H. 264 中 I 帧编码模式选择时的计算量, 提出了一种快速编码算法。该方法首先选择适当的滤波算子对待编码宏块进行平滑滤波, 同时使用滤波后的结果与原始宏块数据进行比较来得到滤波残差量; 然后将该滤波残差量作为依据, 提前判断出当前宏块编码所属的基本模式集, 这样由于减小了候选模式的范围, 因而提高了模式决定的效率。在判定过程中还使用了双阈值方法, 该方法可确保在作出合理判决时, 亦不影响图像质量。采用该算法不仅一定程度上降低了计算复杂度, 并且对图像质量的影响可以忽略不计。实验表明, 这种新的算法可以降低约 16% 的计算量, 而编码图像峰值信噪比的下降只有 0.1dB, 且对比特率的影响不到 1%。

**关键词** I 帧 帧内预测 低通滤波

中图法分类号: TN919.81 文献标识码: A 文章编号: 1006-8961(2008)04-0629-05

## Mode Decision Strategy for I-frames in H. 264 Based on Low-pass Filter

YU Bing, ZHANG Tai-Yi

(Department of Communication & Information Engineering, Xi'an Jiaotong University, Xi'an 710049)

**Abstract** To reduce the complexity of I-frame mode decision process, a fast mode decision algorithm is proposed. A low-pass filter is implemented for each macroblock. After the macroblock is smoothed by this filter we compare the result with the original data and generate the difference between them. Then the basic candidate-mode class is determined based on the difference, so that the number of candidate modes is reduced. Therefore this new method can reduce the complexity of mode decision process in H. 264 coding at a certain degree with almost no influence on the performance and the bitrate. Further experiments indicate that this new method can save about 16% time in computation with 0.1dB decline of performance, and less than 1% increase on bitrate.

**Keywords** I-frame, intra prediction, low-pass filter

## 1 引言

H. 264<sup>[1]</sup> 是联合视频小组 (joint video team, JVT) 和 ITU-T VCEG 合作发布的新一代视频压缩编码标准。该标准使用了许多先进的技术, 其中包括可变块大小的运动补偿、帧内预测、1/4pixels 精度的使用、多参考帧、解块滤波等。这些新技术的应用使得 H. 264 标准在保证图像质量与 MPEG4 标准相同的前提下, 可以降低 50%<sup>[2]</sup> 的码率。然而在性能提高的同时, 由于伴随着计算复杂度的增加, 致使 H. 264 的计算量要比 MPEG4

(simple profile) 高出 16 倍以上<sup>[3]</sup>。巨大的计算量使得 H. 264 在应用于实际压缩编码时困难重重, 因而针对 H. 264 各复杂模块的优化也呼之欲出, 其中帧内预测的模式决定过程, 便是有待优化的重要模块之一。

H. 264 中的 I 帧采用帧内 (Intra) 预测方式进行编码。帧内预测的基本思想是利用相邻像素的空间相关性, 同时使用当前块左侧和上侧的可用像素, 通过预测表达式, 首先得到当前块像素的预测值, 再对预测值和真实值之间的残差进行编码, 以达到使用较少比特进行编码传输的目的。H. 264 的帧内预测模式可以分为 Intra 4 × 4 和 Intra 16 × 16 两个基本模

式集。进一步划分,在 Intra  $4 \times 4$  基本模式集中, H. 264 又规定了 9 种预测方向;而 Intra  $16 \times 16$  基本模式集中也相应地划分了 4 种预测方向,用来进行不同方向性的预测(如表 1 所示)。模式决定的过程就是要选择出两个基本模式集涵盖的所有预测方向中最优的一个方向,用于对当前宏块进行编码。图 1 为一个使用 Intra  $4 \times 4$  基本模式集进行预测的示意图。图中当前块的像素值  $a \cdots p$  是通过周围像素  $A \cdots P$  根据具体的方向预测来得到的。例如若选取了表 1 Intra  $4 \times 4$  模式集中的模式 1,即垂直预测时,则  $a \cdots p$  的值是由  $A$ 、 $B$ 、 $C$ 、 $D$  预测得到的。对于

表 1 Intra  $4 \times 4$  及 Intra  $16 \times 16$  预测模式分类

Tab. 1 Prediction class of Intra  $4 \times 4$  and Intra  $16 \times 16$

模式序号	不同预测模式分类								
	0	1	2	3	4	5	6	7	8
Intra $4 \times 4$ 模式集	DC 预测	垂直预测	水平预测	右对角预测	左对角预测	垂直左对角	垂直右对角	水平左对角	水平右对角
Intra $16 \times 16$ 模式集	垂直预测	水平预测	均值预测	平面预测					

$M$	$A$	$B$	$C$	$D$	$E$	$F$	$G$	$H$
$I$	$a$	$b$	$c$	$d$				
$J$	$e$	$f$	$g$	$h$				
$K$	$i$	$j$	$k$	$l$				
$L$	$m$	$n$	$o$	$p$				

图 1  $4 \times 4$  块及其相邻像素

Fig. 1  $4 \times 4$  block and neighbor pixel

如今已有越来越多的人开始关注帧内预测算法的优化问题,并已经有一些行之有效的方法被提出来,如 Jiang Gang-yi 等人提出了一种与改进的边缘检测方法相结合的快速算法,用来确定最佳的  $4 \times 4$  模式的选择<sup>[4]</sup>;Chen Chao-chung 等人提出了三步快速算法<sup>[5]</sup>,首先检测出现概率最高的某一模式,再使用与该模式方向上相邻几个模式进行预测来选择最佳的预测结果,以减少检测范围;Meng Bo-jun 等人提出对待编码的  $4 \times 4$  块进行采样,并先将其划分为 4 组,然后以某组代替整个块的方法来降低计算的复杂度<sup>[6]</sup>;朱光熹等人提出分级的预测方法<sup>[7]</sup>,并通过引入复杂度变量来进行计算量与复杂度之间的权衡。纵观上述方法,由于它们都是针对某一个基本模式集内部预测方

使用 Intra  $16 \times 16$  模式集进行预测的宏块来说,帧内预测过程也基本相同。H. 264 所使用的全搜索策略是针对当前待编码宏块,首先要进行 Intra  $16 \times 16$  模式集中 4 个方向的预测,以便从中选择出最佳的预测方向;然后再进行 Intra  $4 \times 4$  模式集中 9 个方向的预测,以选出最佳的预测方向;接着比较两个基本模式集中的最佳预测方向,最终得出当前宏块的最佳预测模式。这种方法无疑可以获得最优的预测结果,但是其计算复杂度也是最大的。要应用于实际编码中,就必须针对复杂度大这一弊端进行优化,以改善压缩效率。

向的选择提出的,因而不能预先判断出编码宏块属于哪一个基本模式集,其待编码宏块必须同时检测这两个基本模式集才能得出最佳模式。在这一背景下,本文提出使用低通滤波的方法,即先对滤波前后的图像亮度数据进行比较,然后利用比较结果提前判断出待编码宏块所属的基本模式集,从而弥补了当前快速算法的不足。

## 2 提出的算法

模式决定的首要任务是先判断出当前宏块所属的基本模式集,然后在基本模式集中选取最佳模式。如前所述,全搜索算法和目前的快速算法都不能预先判断出宏块所属的基本模式集,而要同时对两个基本模式集进行检测,目前所提出的优化方法都是在某一个基本模式集内部使用的。针对这一现象,本文提出的算法则是从宏观上对宏块所属的基本模式集进行快速的预判,以使得后续的判决仅限于某一个基本模式集内部,而不需要同时检测两个基本模式集。从这个角度可以认为,本文提出的算法是其他快速算法的基础。

### 2.1 算法的主要思想

研究表明,帧内预测时,基本模式集的选取和图像本身的属性有很大关系。当图像亮度分布均匀、

质地平滑、低频分量较多时,通常会选择 Intra  $16 \times 16$  模式集中的某一方向对整个宏块进行预测;反之,当图像处于细节区域、高频分量丰富的时候,一般会采取 Intra  $4 \times 4$  中的模式进行预测。为了验证上述理论,本文研究了5组不同的序列(如表2所示),其中具有最广泛平滑区域的序列 Akiyo 的 Intra  $16 \times 16$  模式占了 40.8%,而具有最多细节的序列 Mobile 的 Intra  $16 \times 16$  模式占了 2.4%。可见平滑区域较多的序列将会更多地选用 Intra  $16 \times 16$  模式进行编码。基于图像属性和基本模式集选取之间的对应关系,本文指出,如果能够提前对图像是否平滑作出判断,就可以预先判断出预测模式所属的基本模式集,而不需要对两个模式集都进行测试。这既在一定程度上缩小了候选模式的范围,也降低了模式决定过程复杂度。

表2 Intra  $4 \times 4$  及 Intra  $16 \times 16$  分布Tab.2 Distribution of Intra  $4 \times 4$  and Intra  $16 \times 16$ 

序列	Intra $4 \times 4$ 模式 占的比例 (%)	Intra $16 \times 16$ 模式 占的比例 (%)
Akiyo	59.2	40.8
Bus	95.5	4.5
Foreman	87.4	12.6
Mobile	97.6	2.4
Stefan	80.6	19.4

为了对图像平滑特性作出判断,就必须对图像本身的属性进行研究。本文考察了大量图像处理基本方法之后,提出使用平滑滤波器进行判决的方法。平滑滤波是一种空域滤波增强技术,由于它可以对图像的低频分量进行增强,同时还可以削弱图像的高频分量,从而起到平滑图像的作用。根据平滑滤波器本身的特点,笔者认为,如果编码宏块本身是平滑的,且低频分量多,而高频分量少,则滤波之后的效果将不甚明显,滤波前后宏块亮度的差异将很小;反之,如果宏块本身高频分量丰富,滤波之后的宏块由于增强了低频分量和削弱了高频分量,因而宏块的变化较大,即滤波前后宏块的亮度差异将比较大。同理,如果滤波前后的宏块亮度差值小于某个数值,则可认为当前宏块是平滑的,适宜用 Intra  $16 \times 16$  模式集中的模式进行预测;反之如果滤波前后宏块的亮度差值超过某个特定的界限,则认为当前宏块是细节较丰富的,更趋向于使用 Intra  $4 \times 4$  模式集中的模式进行预测。

为了将上述思想通过表达式加以说明,本文定义

变量  $D$  表示滤波前后宏块亮度的差值(见式(1))。

$$D = \sum_{x=0}^{16} \sum_{y=0}^{16} \text{abs}(I_0(x,y) - I(x,y)) \quad (1)$$

式中,  $I_0(x,y)$  是原始图像数据,  $I(x,y)$  是滤波后的图像数据,  $D$  称为滤波残差量,它等于滤波前后宏块亮度的差值,如上所述,  $D$  值从一定程度上是可以反映出当前宏块的平滑程度的。

## 2.2 滤波残差量分布研究以及双阈值判断

在进行模式集选取时,如果希望根据亮度差  $D$  值的预判来得出宏块所属的基本模式集,就需要分别对 Intra  $16 \times 16$  和 Intra  $4 \times 4$  两个模式集合的滤波前后宏块亮度差  $D$  的分布进行研究,以找到合适判决点。

根据对上述 Akiyo 等5组序列两个基本模式集的亮度差的研究发现,这两个基本模式集的亮度差值的分布差异是明显的。如后续实验中图2所示 Intra  $4 \times 4$  模式集的亮度差值大部分都大于 60,而 Intra  $16 \times 16$  模式集的亮度差值基本上集中于 15 附近。在 15~60 之间的区域,两个基本模式集都可能占用。传统的单一门限的方法虽然只需要一步就可以判断出来,但是由于这种非此即彼的做法对于那些平滑特性并不明显的宏块来说,无疑会降低图像的质量,因此本文采用双阈值的方法,即通过设定两个门限值  $T_{\max}$ 、 $T_{\min}$  来进行判别,如果当前宏块的亮度差值大于  $T_{\max}$ ,则判断它为使用 Intra  $4 \times 4$  模式集中的模式预测的宏块;如果当前宏块的滤波前后亮度差值小于  $T_{\min}$ ,则判断它为使用 Intra  $16 \times 16$  模式集中的模式预测的宏块,否则认为当前宏块的特征不明显,要使用全搜索方法同时对两个集合进行预

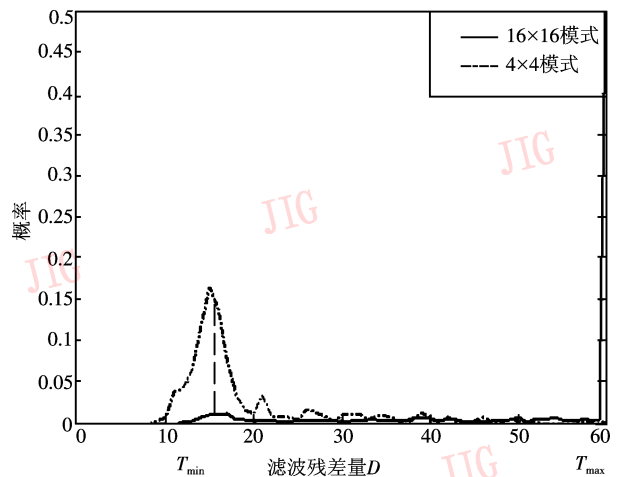


图2 滤波残差量分布图

Fig.2 Distribution of  $D$

测模式的检测。这样做既能保证能够对特征鲜明宏块进行预判,同时又能保证图像的质量。

相比于全搜索方法,使用上述算法对一个帧内宏快进行模式决定时,如果该宏块适合 Intra  $4 \times 4$  模式编码,则可以少计算 Intra  $16 \times 16$  模式集中 4 个方向的预测;反之,如果该宏块适合 Intra  $16 \times 16$  模式编码,则可以少计算  $16 \times 9 = 144$  次的 Intra  $4 \times 4$  模式预测,而且使用简单的均值平滑滤波并不会带来过多的计算量,因此对编码效率影响不大。从总体上看,复杂度被降低了。

### 3 实验结果及分析

#### 3.1 滤波前后宏块亮度差分布实验

为了获取 Intra  $4 \times 4$  以及 Intra  $16 \times 16$  两个基本模式集中亮度差值的分布情况,本文选取前述 5 组 CIF ( $352 \times 288$ ) 格式的序列,分别针对两个基本模式集进行亮度差值的计算,并将得到的结果绘制成分布图(图 2)。实验结果表明,由于两组数据的分布确实有很大差异,因此可以通过该分布对宏块所属的基本模式集进行预判。分别考查两个基本模式集宏块亮度差的分布情况设置出两个阈值(如图 2 中所示的  $T_{\max}$ 、 $T_{\min}$  两个阈值)。由图 2 可以看出,当亮度差  $D$  值大于  $T_{\max}$  时,几乎已经没有 Intra  $16 \times 16$  模式出现,而大部分的 Intra  $4 \times 4$  模式都集中于此。

#### 3.2 基于低通滤波的模式决定实验

为验证本文算法的效果,使用 JM86 软件参考

模型对上述的 5 组视频序列进行了实验,实验采用了双阈值判决的方法。实验是在 PC 下完成,使用 Intel Petium4 处理器,3.0GHz,内存 512MB。针对 5 组 CIF 格式的视频序列的前 300 帧进行测试,QP 值分别选取 20、30 和 40,使用 CAVLC 编码。表 3 所示为实验结果。这里采用  $t_s$  表示新的算法比全搜索算法所节省的时间,它的计算公式定义如下:

$$t_s = \frac{t_r - t_o}{t_r} \times 100\% \quad (2)$$

式中, $t_r$  为采用新的算法后所需要的执行时间, $t_o$  表示原执行时间。峰值信噪比(PSNR)表示客观图像质量。本文为了计算简便,还采用  $3 \times 3$  的均值滤波器对宏块进行平滑滤波。需要注意的是,由于本文是只针对 I 帧进行优化,而 I 帧的数量是有限的,因此算法对整个序列比特率的改变是微乎其微的,为了说明优化对视频序列产生的影响,这里使用比特率变化  $\Delta\text{BitCnt}$  来表示新的算法对编码之后 I 帧大小的影响;由于对一个宏块进行 16 次 Intra  $4 \times 4$  模式的分析要比进行一次 Intra  $16 \times 16$  分析复杂得多,因此理论上讲,如果能够尽量少地避免 Intra  $4 \times 4$  模式分析,则会尽可能多地提高效率。也就是说,该方法对于使用较多 Intra  $16 \times 16$  模式进行编码的序列来说,其效率会提高得更多。实验结果也基本符合理论的分析,该方法对 Akiyo 序列效率提高最多,而对 Mobile 序列则改善较少。表 3 是针对 5 组 CIF 格式序列,在不同量化步长下所得到的结果。

表 3 基于平滑滤波的实验结果

Tab. 3 Experiment result based on Low-pass filter

测试序列	$\Delta\text{PSNR}(\text{dB})$	$\Delta\text{BitCnt}(\%)$	$t_s(\%)$	$\Delta\text{PSNR}(\text{dB})$	$\Delta\text{BitCnt}(\%)$	$t_s(\%)$	$\Delta\text{PSNR}(\text{dB})$	$\Delta\text{BitCnt}(\%)$	$t_s(\%)$
Akiyo	0.02	1.7	33.37	0.03	1.9	34.45	0.05	2.4	39.35
Bus	0.01	0.0	8.50	0.01	0.1	7.85	0.03	0.8	10.10
Foreman	0.01	2.5	11.40	0.00	2.7	11.00	0.01	4.0	15.00
Mobile	0.00	0.1	7.70	0.00	0.1	6.08	0.02	0.5	11.13
Stefan	0.01	0.3	16.55	0.01	0.5	15.88	0.03	1.0	19.45
QP		QP = 20			QP = 30			QP = 40	

#### 3.3 与以往帧内模式决定算法的比较

为了揭示本文提出算法的意义,可将该算法与以往典型快速算法结合起来使用,并将得到的编码结果与以往典型算法的编码结果进行比较。实验是在本论文算法的基础上,选用 Meng Bo-jun 等人提

出的 EIP 算法<sup>[6]</sup>,与单纯使用 EIP 算法得到的编码结果进行比较,并以 CIF 格式的 Stefan 和 Foreman 序列为例来进行比较(见表 4)。结果表明,这两种方法结合使用可以在更大程度上提升效率,而对图像的质量以及比特率却没有明显的影响。

表4 本算法与以往算法结合使用的结果

Tab. 4 Comparing between this new algorithm and the EIP algorithm

QP = 30	原始的 EIP 算法			本论文算法 + EIP 算法		
	$\Delta PSNR$ (dB)	$\Delta BitCnt$ (%)	$t_s$ (%)	$\Delta PSNR$ (dB)	$\Delta BitCnt$ (%)	$t_s$ (%)
Stefan	0.01	1.4	30.25	0.15	2.1	40.52
Foreman	0.01	1.5	26.30	0.11	3.1	34.78

## 4 结论

本文首先通过研究现有的 I 帧模式决定的快速算法,发现了其中的不足之处,然后通过研究帧内编码模式与图像属性之间的关系,提出了一种使用低通滤波器对编码宏块进行处理,再基于该处理结果使用双阈值判断的方法来提前判断出宏块所属的基本模式集合,从而有效地提高了编码效率,并成为进一步使用快速算法在某一基本模式集合中进行模式决定的基础。实验结果表明,由于该方法可以平均节省约 16% 的计算量,PSNR 值平均下降了 0.01dB, I 帧的大小只增加了约 1.1%,因此是一种行之有效的新方法。

由于受实验环境限制,本文的结果是针对 300 帧长的序列所做的,如果应用于实际的视频播放,那么理论上会有更明显的效果。由于本文对宏块进行平滑滤波还是占用了一定的时间,为了降低这一影响,因此可以只针对宏块的某些样本进行滤波,再分析滤波结果。由于同时注意到本文算法对于那些大部分使用 Intra  $4 \times 4$  模式编码的序列,其性能的改善是有限的,因此,今后可以考虑同时使用平滑、锐化滤波器来预处理图像,先进行基本模式集的预判,以消除单独使用平滑滤波方法的局限性。

## 参考文献 (References)

- 1 H. 264 ISO/IEC 14496-10 AVC, Doc. JVT-G0505r1. Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification [S]. Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, 2003.
- 2 Wiegand T, Sullivan G, Bjntegaard G, *et al.* Overview of the H. 264/AVC video coding standard [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 560 ~ 576.
- 3 Zhou M. Evaluation and simplification of H. 26L baseline coding tools [A]. In: Document JVT-B030, 2nd Meeting [C], Geneva, Switzerland, Jan. 29-Feb. 1, 2002.
- 4 Jiang Gang-yi, Li Shi-ping, Yu Mei, *et al.* An efficient fast mode selection for intra prediction [A]. In: Proceedings of IEEE Workshop VLSI Design & Video Technique [C], Suzhou, China, 2005: 357 ~ 360.
- 5 Cheng Chao-chung, Chang Tian-sheuan, Fast three step intra prediction algorithm for  $4 \times 4$  blocks in H. 264 [A]. In: Proceedings of International Symposium on Circuits and Sgstems [C], Kobe, Japan, 2005, 2: 1509 ~ 1512.
- 6 Meng Bojun, Oscar C Au\*, Wong Chi-Wah, *et al.* Efficient intra-prediction algorithm in H. 264 [A]. In: Proceedings of the 2003 IEEE Computer Society, International Conference on Image Processing [C], Warsaw Poland, 2003.
- 7 Zhu Guang-xi, Dai Sheng-kui, Liu Wen-yu, *et al.* A classified method of intra prediction in H. 264 [J]. Micro-Electronics and Computer Science, 2005, 22(3): 211 ~ 214. [朱光喜, 戴声奎, 刘文予等. 一种复杂度可分级的帧内预测方法 [J]. 微电子学与计算机, 2005, 22(3): 211 ~ 214.]