H.264/AVC 快速帧内预测模式选择算法 *

张 敏[1,2], 耿国华[1]
(1. 西北大学 信息科学与技术学院, 西安 710069; 2. 西安翻译学院 计算机教研室, 西安 710105)

摘 要: 提出一种基于边缘方向预测模式度量的快速预测模式选择算法,通过对当前 4×4 块帧内预测方向的度量,根据估计的块边缘方向来选择候选预测模式。实验结果表明,与全搜索算法相比,该算法在保证图像质量和比特率基本不变的前提下,编码时间平均减少了 40%,提高了编码效率。

关键词: 视频编码; 帧内预测; 率失真优化; 模式选择
中图分类号: TP301.6 **文献标志码**: A **文章编号**: 1001-3695(2011)02-0782-03
doi:10.3969/j.issn.1001-3695.2011.02.107

Fast intra-prediction mode selection algorithm for H.264/AVC

ZHANG Min[1,2], GENG Guo-hua[1]
(1. *College of Information Science & Technology, Northwest University, Xi' an* 710069, *China*; 2. *Computer Faculty Working Office, Xi' an Fanyi University, Xi' an* 710105, *China*)

Abstract: This paper proposed a fast prediction mode selection algorithm based on measurement of direction prediction mode, by measuring the intra-prediction direction of 4×4 block, it used the correlation between neighboring blocks and neighboring direction prediction mode to predict the intra-frame mode. Experimental results show that compared with the full searching algorithm, this algorithm can save the encoding time up to 40% averagely when guarantee the capability of the rate distortion and the bits rate, and it improves the encoding efficiency tremendously.

Key words: video coding; intra-prediction; rate distortion optimization(RDO); mode selection

H.264/AV 是由 ITU-T VCEG 和 ISO/IEC MPEG 共同建立的联合视频工作组(Joint Video Team, JVT)制定的新一代视频编码标准,这一标准在保持了传统的基于块的混合编码方法外,又采用了大量的新技术[1]。例如:统一的 VLC 符号编码、1/4 像素运动估计、基于 4×4 块的整数变换、分层的编码语法等,特别是 H.264 引入帧内预测,并通过率失真优化(RDO)技术对帧内预测模式的每一种组合进行计算,以取得最优的编码模式,提高 I 帧的编码效率,但也大大增加了编码的计算复杂度。因此,帧内预测快速算法成为 H.264/AVC 研究的热点之一。

目前,对帧内预测的研究已取得很多成果。文献[2]最早提出对帧内预测模式的选择进行优化,根据物体边缘连续的特性以及相似的特性,提出了基于相同边缘方向的相邻像素进行预测,采用 Soble 算子计算当前块的边缘信息,在提高码率的同时也增加了 Soble 算子的计算梯度值。文献[3]利用最佳预测模式和各模式下像素差值和的关系,通过计算各预测模式残差来得出像素块的纹理方向,从而获得最佳预测模式。文献[4]提出自适应阈值快速算法,利用邻块相关性为当前块的率失真设置一个阈值,提前判断当前块是否为帧内块,但视频图像的动态变化性导致其阈值很难预测。文献[5]利用编码块内像素方向的相关性减少了 4×4 亮度块候选预测模式的数目和 RD 值的计算次数,从而降低帧内预测的计算复杂度。

本文首先根据模式选择的提前终止原则预先排除可能性较小的预测模式,缩小模式选择的范围,在 4×4 帧内模式预测中结合基于相邻块的空间相关性方法计算当前块的 MPM (most probable mode,最可能模式),采用基于方向特性的掩模算法[6],对文献[5]作了改进,使参与 RDO 运算的预测模式大大减少,有效降低了计算复杂度。

1 H.264/AVC 帧内预测算法分析

图 1 给出了一个 4×4 亮度块的帧内预测示意图,其中大写字母 A~M 是已经编码并重建后可以作为预测的像素,小写字母 a ~ p 是当前正在编码的亮度块像素。对于 4×4 亮度编码块,共有九种不同的帧内预测模式(模式 0~模式 8)。图 2 中箭头指明了八种具有方向的预测模式,另外一种预测模式为 DC 预测(模式 2)。

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 4×4块像素点分布

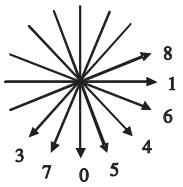

图2 4×4块八个帧内预测模式

为了在编码质量和码率花费之间获得平衡,H.264/AVC 采用了率失真优化(rate distortion optimization, RDO)作为模式选择的精确代价函数。RDO 代价函数为[7]

$$RDcost = SSD + \lambda\ mode \cdot rate \tag{1}$$

其中:SSD(sum of square difference)为原始块与重建块像素值的平方差和;λ mode 为 Q_P 的指数函数,Q_P 为量化参数;rate 为

收稿日期: 2010-07-22; **修回日期**: 2010-08-31 **基金项目**: 国家博士点基金资助项目(200806970014)

作者简介:张敏(1980-),女,内蒙古赤峰人,讲师,主要研究方向为图形图像处理和视频编码(zhm521@hotmail.com);耿国华(1955-),女,教授,博导,主要研究方向为信息系统、人工智能.

采用该预测模式需要的比特数。在 H.264 的帧内预测编码模式下,采用遍历所有可能的模式,通过寻找编码后失真和码率总和最小的 RDcost 来确定最优编码模式,其算法描述如下:

a)分别计算 4×4 亮度块中九种模式的 RDcost,选择最小代价模式。

b)重复步骤 a),把宏块中 16 个 4×4 亮度块对应的最小 RDcost 相加得到该宏块帧内 4×4 下的 RDcost。

c)分别计算四种模式下宏块的绝对误差和 SAD(sum of absolute difference),选择最小 SAD 的模式,然后计算该模式下的 RDcost 得到该宏块帧内 16×16 方式下的 RDcost。

d)比较 16×16 方式和 4×4 方式下的 RDcost,选择最小 RDcost 的方式作为该宏块的帧内预测模式。

如果采用全搜索方法选取 RDcost 最小的模式为最佳模式,则确定一个宏块的预测模式必须进行 M8×(M4×16+M16)次 RDO 计算,其中,M8、M4、M16 分别表示 8×8 色度块、4×4 亮度块、16×16 亮度块的预测模式数,这就意味着每个宏块要进行 592 次 RDO 计算。这个计算量是非常庞大的。

2　快速帧内预测模式选择算法

在 H.264/AVC 编码标准中,帧内预测应用于 I 帧和 P 帧。利用图像相邻块的空间相关性,以及当前宏块的左边和上边宏块的像素值对当前宏块进行预测。在编码时,对宏块的实际值和预测值的差值进行编码,这样可以节省图像帧的比特数。

在 H.264/AVC 帧内预测编码中,对于色度分量,包含四种 I8 预测模式;对于亮度分量,包含四种 I16 预测模式和九种 I4 预测模式。

2.1　空间相邻块相关的帧内预测

根据一般自然图像的内容均具有不同程度的相关性可知,除了极个别变化特别剧烈的图像,各个块的最优模式具有一定的相关性,即对于即将预测的块而言,其最优模式值与其上边及左边两个块的模式值有较强的相关性,如图 3 所示。

块 C 为当前要预测的 4×4 块,当已知其相邻 4×4 块 A 和 B 的预测模式时,就可以把 A 或 B 的预测模式作为 C 的候选预测模式。如果 A 和 B 的预测模式一致,则候选模式记为一种,如果不一致,则候选模式记 A、B 预测模式的较小者。

2.2　宏块帧内纹理复杂度的预先判断

H.264/AVC 帧内预测编码中,I16 亮度块适合对纹理简单、平坦的图像块进行预测,I4 亮度块适合对纹理复杂、细节丰富的图像块进行预测。若能在宏块帧内预测进入模式搜索前对纹理复杂度加以判断,从而选择相应尺寸的预测模式编码,结合提前退出原则,就能一定程度地减少模式搜索带来的计算开销,从而降低计算复杂度。判断图像的细节是否丰富,即是计算图像中像素灰度值与灰度均值的偏离程度,本文取均方差来表示这个度量值。

$$V=\frac{1}{16\times16}\sum_{i=1}^{16}\sum_{j=1}^{16}(x_{i,j}-\bar{x})^2 \tag{2}$$

其中:x_{ij}为 16×16 块中的每一个像素灰度值,$\bar{x}$为该块中所有像素的灰度均值。V 值越大,表示图像中像素的灰度值与均值的偏离程度越高,即该图像块的细节越丰富,适用于 4×4 帧内模式预测;V 值越小,表示这种偏离程度越小,该图像细节越少,则采用 16×16 模式的帧内预测更有效。本文算法设定两个阈值 Tv_1、Tv_2,$Tv_1<Tv_2$。在算法的初始阶段,计算 V 值。如果 $V\leqslant Tv_1$,则宏块采用 16×16 的模式进行帧内预测,不需要进入 4×4 的模式选择阶段,算法提前终止;如果 $V\geqslant Tv_2$,则宏块采用 4×4 块模式进行预测,不需要与 16×16 模式的预测结果进行比较,4×4 预测得到的最优模式即为该宏块帧内预测编码的最优模式;如果 $Tv_1<V<Tv_2$,则两种模式选择阶段都要进行,并比较结果得到最优预测模式。

2.3　快速 4×4 块帧内预测模式选择

一般来说,每一个块的最佳预测方向与块内是一致的,因此,如果能够快速、合理地估计块边缘方向,就可以根据估计的块边缘方向来选择候选预测模式,降低计算复杂度。

2.3.1　DC 模式的阈值判断

因为 DC 模式不具有方向性,所以在选择候选模式时还应考虑到 DC 模式。如果当前块的纹理变化较为平缓,即灰度均匀,则最佳预测模式为 DC 模式的可能性就很大;反之,如果当前块纹理变化较剧烈,即灰度不均,那么最佳预测模式为 DC 模式的可能性就很小。为了判断纹理变化的快慢,这里同样可以使用均方差作为度量值,因此引入变量 V_{dc}:

$$V_{dc}=\frac{1}{4\times4}\sum_{i=1}^{4}\sum_{j=1}^{4}(x_{i,j}-\bar{x})^2 \tag{3}$$

其中:x_{ij}为 4×4 块中的每一个像素灰度值,$\bar{x}$为该块中所有像素的灰度均值。如果 V_{dc}的值小于等于某一阈值 T_{dc},则说明当前 4×4 块变化缓慢,只需选用 DC 模式,不必进行方向判断;如果 V_{dc}的值大于阈值 T_{dc},则表明当前 4×4 块变化剧烈,不必将 DC 模式列为候选预测。

2.3.2　4×4 块边缘方向帧内预测

首先根据文献[5]对 4×4 亮度块进行分割(图 4),得到 A、B、C、D 四个 2×2 块,如图 4 所示,并根据式(4)对这四个小块求和,把相应的和也记为 A、B、C、D。其中 $p(x,y)$($0\leqslant x\leqslant3$,$0\leqslant y\leqslant3$)表示当前 4×4 亮度中点(x,y)的像素值。

$$\begin{cases}A=\sum_{y=0}^{1}\sum_{x=0}^{1}p(x,y),B=\sum_{y=0}^{1}\sum_{x=2}^{3}p(x,y)\\ C=\sum_{y=2}^{3}\sum_{x=0}^{1}p(x,y),D=\sum_{y=2}^{3}\sum_{x=2}^{3}p(x,y)\end{cases} \tag{4}$$

图3　4×4块及其相邻块　　图4　4×4块分割

为了估计当前的块边缘方向,引入两个边缘特征数,见式(5):

$$\begin{cases}F_1=\left|\dfrac{(A+C)-(B+D)}{S}\right|\\ F_0=\left|\dfrac{(A+B)-(C+D)}{S}\right|\end{cases} \tag{5}$$

其中 S 的取值为

$$S=\begin{cases}8 & Q_p<20\\ 16 & 20\leqslant Q_p<30\\ 32 & 30\leqslant Q_p<40\\ 64 & Q_p\leqslant40\end{cases} \tag{6}$$

它由量化步长决定。因为 DC 模式在前面作了预判,所以只在边缘方向不明显时被选择,最后根据 F_1 和 F_0 来估计该亮度块的边缘方向,选择候选预测模式如表 1 所示。

2.4　快速帧内预测模式选择算法描述

a)对当前 16×16 亮度块的 V 值:若 $V\leqslant Tv_1$,则选择 16×

16 模式预测的最佳模式即为最终帧内预测模式,算法结束;若 $V \geqslant Tv_2$ 转向步骤 b);若 $Tv_1 < V < Tv_2$,则比较 16×16 和 4×4 预测的最佳模式的 RDcost,较小者为最终帧内预测模式,算法结束。

b)计算当前 4×4 块 V_{dc} 值,若 $V_{dc} \leqslant T_{dc}$ 则 DC 模式即为最终帧内预测模式,即为 $Mode_{dir}$,算法结束;若 $V_{dc} \geqslant T_{dc}$,则转向步骤 c)。

c)选取当前 4×4 块划分为四个 2×2,利用式(4)求出相应的和记为 A、B、C、D。

d)利用公式(5)计算其边缘特征值 F_1 和 F_0,利用表 1 选取候选预测模式。

表 1　预测模式判断表

case	F_1和 F_0关系	边缘方向	预测模式
1	$\lvert F_1\rvert = \lvert F_0\rvert = 0$	没有明显边缘	DC
2	$\lvert F_0\rvert = 0$ 且 $\lvert F_1\rvert > 0$	垂直边缘	0
3	$\lvert F_1\rvert = 0$ 且 $\lvert F_0\rvert > 0$	水平边缘	1
4	$\lvert F_1\rvert = \lvert F_0\rvert > 0$ 且 $F_1/F_0 = 1$	左下角对角线	3
5	$\lvert F_1\rvert = \lvert F_0\rvert > 0$ 且 $F_1/F_0 = -1$	右下角对角线	4
6	$\lvert F_1\rvert > \lvert F_0\rvert > 0$ 且 ($F_1/F_0 > 1$ 或 $F_1/F_0 < -1$)	垂直区域边缘	0,5,7
7	$\lvert F_1\rvert > \lvert F_0\rvert > 0$ 且 ($F_1/F_0 < 1$ 或 $F_1/F_0 > -1$)	水平区域边缘	1,6,8

e)用拉格朗日率失真函数计算候选预测模式的代价 RDcost。

f)重复步骤 d)e),直到所有候选预测模式都计算完,选取最小 RDcost 的模式为最佳帧内预测模式。

g)对当前宏块中 16 个 4×4 亮度块重复步骤 c)~f),获得每一个 4×4 亮度块的最佳帧内预测模式和相应的代价 RDcost,从而获得该宏块在 4×4 亮度块帧内预测模式下的率失真开销 RDcostMB。

本文所提快速算法中,16×16 亮度块和 8×8 色度块的帧内预测编码方法与 H.264 中全搜索帧内模式选择算法相同。4×4 亮度块原算法帧内预测模式数为 9,本章所提算法 4×4 块帧内预测模式数为 1~4,对于大部分块为 2。

实验结果表明(表 2 和 3):提出的快速算法在保持 PSNR 和比特率基本不变的前提下,有效地减少了 4×4 亮度块候选预测模式数,从而减少了 RD 值的计算次数,使得计算量大大减少,提高了编码速度。

3　实验结果与分析

本文在 H.264/AVC 的官方平台 JM 测试模型 8.6version[8] 上,分别对多个不同类型的 QCIF 和 CIF 视频图像序列进行测试,并与 JM 的全搜索算法在编码时间、峰值信噪比、码率这三个方面作了比较。

测试的实验参数设置为:RDO 有效;Hadamard 变换有效;CAVLC 熵编码;每个序列取 300 帧,均为 I 帧编码;两个阈值 Tv_1、Tv_2 分别设定为 32 和 420,$Q_p = 28,32,48$。比较结果如表 2 和 3 所示。其中:Δtime 表示编码时间的变化;Δbit 示码率的变化;ΔPSNR 表示 Y 分量的峰值信噪比的变化。

可以看到,与全搜索算法相比,改进后帧内预测编码算法时间节省达到 40%。考虑到时间消耗的减少,质量的略微下降和比特率的增加程度仍在可以接受的范围之内。

表 2　QCIF 格式视频序列实验结果

Q_p	序列	Δtime/%	ΔPSNR/dB	Δbit/%
28	foreman	-40.61	-0.016	0.22
	container	-39.65	-0.013	0.16
	mobile	-39.93	-0.018	0.16
32	foreman	-40.87	-0.018	0.17
	container	-41.54	-0.017	0.19
	mobile	-40.05	-0.010	0.12
48	foreman	-41.87	-0.021	0.16
	container	-40.91	-0.020	0.19
	mobile	-41.50	-0.022	0.15

表 3　CIF 格式视频序列实验结果

Q_p	序列	Δtime/%	ΔPSNR/dB	Δbit/%
28	foreman	-41.16	-0.018	0.19
	container	-40.27	-0.015	0.20
	mthr_dotr	-40.00	-0.017	0.18
32	foreman	-39.87	-0.016	0.18
	container	-41.21	-0.017	0.16
	mthr_dotr	-39.95	-0.012	0.10
48	foreman	-40.27	-0.010	0.19
	container	-39.21	-0.013	0.12
	mthr_dotr	-41.95	-0.015	0.15

4　结束语

针对 H.264 /AVC 帧内预测复杂度高的问题,本文提出了一种快速的帧内预测算法。对全搜索算法的两个部分进行了改进:一是针对宏块级的预判,避免了对每一个宏块都作 I4 和 I16 两种预测;二是针对 I4 块的边缘方向预测模式选择算法,减少了参与计算的 I4 块。实验结果表明,该算法在保证编码质量的前提下,大大提高了编码速度,增强了实时视频编/解码的可实现性。该算法满足实时视频通信的要求,可用于实际的视频通信产品。

参考文献:

[1] WIEGAND T, SULLIVAN G J. Overview of the H.264/AVC video coding standard[J]. IEEE Trans on Circuits and Systems for Video Technology, 2003,13(7):560-576.

[2] PAN Feng, LIN Xiao, RAHARDJA S. *et al*. Fast mode decision algorithm for intra prediction in H.264/AVC video coding[J]. IEEE Trans on Circuits and Systems for Video Technology, 2005,15(7):813-822.

[3] 罗亮,刘春生,周芦明. H.264/AVC 快速帧内预测模式选择算法[J]. 计算机应用,2008,28(4):1006-1008.

[4] 杨垣,杨万海,裴继红. 基于小波分解的不同聚焦点图像融合方法[J]. 电子学报,2001,29(6):846-848.

[5] ZENG Huan-qiang, CAI Can-hui, MA Kai-kuang. Fast mode decision for H.264/AVC based on macro block motion activity[J]. IEEE Trans on Circuits and Systems for Video Technology, 2009,19(4):1-11.

[6] KIM C S, LI Qing, Kuo C-C J. Fast intra-prediction model selection for H.264/AVC codec[C]//Proc of SPIE International Symposium on ITCOM. Orlando, Florida: IEEE Press, 2003.

[7] FU Fang-wen, LIN Xing-gang, XU Li-dong. Fast intra prediction algorithm in H.264/AVC [C]//Proc of the 7th International Conference on Signal Processing. 2004:1191-1194.

[8] Joint Video Team (JVT). Reference software JM86 [BP/OL]. (2007). http://iphome.hhi.de/suehring/tml/download/okl_jm86.zip.