

一种基于 AVS 视频编码的快速帧间 预测模式选择算法

王蕾睿¹⁾ 沈礼权²⁾ 张兆扬²⁾

¹⁾(上海大学通信与信息工程学院, 上海 200072) ²⁾(上海大学新型显示技术及系统应用教育部重点实验室, 上海 200072)

摘 要 由于在 AVS1-P2 中采用了多种预测模式的帧间预测技术,但为了使编码器性能最优,需要对所有帧间候选模式进行遍历,以便从中选择最适合的模式,这无疑会大大增加编码器的计算复杂度。为了快速的进行帧间预测,根据 AVS1-P2 帧间预测模式的特点提出了一种快速的帧间预测模式选择算法,其特点是将基于全零块的提前中止准则与基于纹理特性的预测模式尺寸预选准则相融合来加快编码速度。实验结果表明,与 AVS1-P2 参考软件 rm52h 相比较,该方法在保持比特率和图像质量基本不变的同时,可显著减少编码器的计算复杂性。

关键词 AVS1-P2 全零块检测 整数变换 帧间模式选择

中图分类号: TN919.81 **文献标识码**: A **文章编号**: 1006-8961(2007)10-1794-04

A Fast Inter Mode Decision Algorithm Based on AVS Video Coding

WANG Lei-rui¹⁾, SHEN Li-quan²⁾, ZHANG Zhao-yang²⁾

¹⁾(College of Communication & Information Processing, Shanghai University, Shanghai 200072)

²⁾(Key Laboratory of Advanced Display and System Application, Ministry of Education, Shanghai University, Shanghai 200072)

Abstract AVS1-P2 introduces several prediction modes for inter coding. In order to achieve the best coding efficiency, the rate distortion optimization is exploited to choose the most suitable predicting mode, but it is very computationally expensive. In this paper, according to the characteristics of AVS1-P2 integer transform, we propose a new fast inter mode decision algorithm to improve the encoder speed based on two criterions, early termination based on all zero block detection and the prediction of inter-mode size using texture characteristic of macroblock. Compared with the reference code, simulation results show that this method will reduce the computing complexity of encoder and improve the encoding efficiency of AVS1-P2 software encoder while bit-rate and video quality almost remain unchanged.

Keywords AVS1-P2, all-zero block detection, integer transform, inter prediction

1 引 言

AVS1-P2 是指由国家信息产业部科学技术司于 2006 年 2 月批准的我国第 1 个先进音视频编解码标准中的视频标准。对于帧间预测,AVS1-P2 支持 P 帧和 B 帧两种帧间预测模式。P 帧至多采用 2 个前向参考帧;B 帧采用前后各一个参考帧。与 H.264/AVC 的多参考帧相比,AVS1-P2 在不增加存储和数据带宽

等资源的情况下,在尽可能避开 H.264/AVC 核心专利的前提下,可提高压缩性能^[1]。为了获得最佳的率失真性能,AVS1-P2 参考软件 rm52h 先对所有候选模式进行遍历,然后从中选择最适合的模式,但由于这会大幅增加编码器的计算复杂度,不利于实时应用,因而本文提出了一种将全零块预先判决准则和纹理复杂度预选块尺寸准则两者相结合的快速帧间预测模式选择算法,实验表明,本文算法能在保持编码质量的同时,显著地提高编码速度。

基金项目: 电子信息产业部电子发展专项基金资助和上海市重点学科项目资助(T0102)

收稿日期: 2007-06-01; **改回日期**: 2007-07-05

第一作者简介: 王蕾睿(1983 ~),女。2006 年获长沙理工大学电子信息科学与技术专业学士学位,现为上海大学通信与信息工程学院硕士研究生。主要研究方向包括数字图像处理和数字视频编码和传输。E-mail: rui8585@163.com

2 适用于 AVS 的全零块检测技术

2.1 AVS 变换与量化原理

AVS1-P2 采用 8×8 的 2 维整数余弦变换,可以

$$Y = HXH^T = \begin{bmatrix} 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 \\ 10 & 9 & 6 & 2 & -2 & -6 & -9 & -10 \\ 10 & 4 & -4 & -10 & -10 & -4 & 4 & 10 \\ 9 & -2 & -10 & -6 & 6 & 10 & 2 & -9 \\ 8 & -8 & -8 & 8 & 8 & -8 & -8 & 8 \\ 6 & -10 & 2 & 9 & -9 & -2 & 10 & -6 \\ 4 & -10 & 10 & -4 & -4 & 10 & -10 & 4 \\ 2 & -6 & 9 & -10 & 10 & -9 & 6 & -2 \end{bmatrix} [X] \quad (1)$$

在量化级数选取方面,AVS1-P2 不同于 H.264/AVC,其对 52 种步长用一个量化参数(quantization parameter, QP)值作为索引^[2],而采用 64 级量化,每当 QP 值增加 8,则量化步长 Q_{step} 增加一倍^[3]。这就增强了适应不同的应用和业务对码率和质量有不同要求的能力。为了减少变换和量化过程中取整带来的误差,AVS1-P2 中将缩放和量化结合在一起操作(如式(2)所示)。

$$F_q(u, v) = \text{sgn}[F(u, v)] [((|F(u, v)|M(u \& 3, v \& 3) + 2^{18}) > 19) \times QC + f \times 2^L > L] \quad (2)$$

其中, F 为变换系数矩阵; u, v 分别为行和列索引; $L = 15 + QP/8$; M 为缩放矩阵, QC 是以 QP 为索引的由若干正整数组成的量化参数表,其取值可以参见文献[4]。 f 用来舍入取整,在实际应用中,有时也可不考虑。

2.2 基于 AVS 的全零块预先判决方法

一些学者在进行 H.264/AVC 的编码块模式的选择中,对全零块的判别曾进行了研究,而本文的全零块检测则是基于 AVS1-P2 的特点之上的。对于前述的 8×8 整数变换,就是先对图像差值做 1 维行整数变换,再做 1 维列整数变换,式(1)的变换矩阵也可表示为

$$F(u, v) = \sum_{i=0}^7 \sum_{j=0}^7 c_{u,i} x_{i,j} c_{j,v}, \quad 0 \leq u, v \leq 7 \quad (3)$$

对上式进行不等式缩放可得到

$$F_{\max}(u, v) \leq \sum_{i=0}^7 \sum_{j=0}^7 |c_{u,i} x_{i,j} c_{j,v}| \leq \sum_{i=0}^7 \sum_{j=0}^7 |c_{u,i} c_{j,v}| |x_{i,j}| \quad (4)$$

根据式(1)中给出的变换矩阵可知,

$$F_{\max}(u, v) \leq 100 \sum_{i=0}^7 \sum_{j=0}^7 |x_{i,j}| \quad (5)$$

消除采用 DCT 变换存在失配的固有问题。 8×8 整数变换如式(1)所示,式中, X 表示 8×8 的图像残差矩阵, Y 表示 8×8 的变换系数矩阵, H 表示 8×8 的整数正变换矩阵, H^T 表示 8×8 的整数正变换矩阵的转置矩阵。

根据 AVS1-P2 中给出的量化公式(式(2)),结合式(5),可得如下全零块判别准则:

$$\sum_{i=0}^7 \sum_{j=0}^7 |x_{i,j}| < \frac{2^L \times 2^{19}}{100 \times M_{\max} \times QC} \quad (6)$$

其中, $M_{\max}$ 为缩放矩阵 M 中的最大值,由于 AVS1-P2 采用了 64 级量化,每当 QP 值增加 8 时,量化步长 Q_{step} 大致增加一倍,即量化参数表 QC 中的值可以表示为 $2^{15-QP/8}$ 。因此,式(6)可表示为

$$\sum_{i=0}^7 \sum_{j=0}^7 |x_{i,j}| < \frac{2^{QP/4} \times 2^{19}}{100 \times M_{\max}} \quad (7)$$

由于在上面的分析过程中采用了不严格的比例缩放,式(5)右端所得到的值大于 $F_{\max}(u, v)$,所以式(7)是全零块判别的充分条件,虽无误判,但存在漏判。在本文提出的算法中,为了减少漏判,所选择的阈值适当地有所增大。

3 AVS 帧间预测模式选择算法

AVS1-P2 采用图 1P 帧的 4 种宏块划分方式,而 H.264/AVC 中的 8×4 、 4×8 、 4×4 3 种划分方式,虽利于运动的精度估计,但由于需要传送更多的块模式信息,因此未被 AVS1-P2 采用,实验证明,在高清测试中,性能没有明显下降^[3]。

3.1 基于 AVS 的快速帧间预测模式选择算法

为篇幅所限,本文提出的快速帧间预测模式选择算法只考虑不含 B 帧的基本类,即只有 P 帧的 4 种帧间预测模式: 16×16 、 16×8 、 8×16 、 8×8 。选择帧间预测模式时,首先将 16×16 和 8×8 预测模式作为候选模式对当前宏块进行预测,如果宏块残差满足全零块判决条件,就判当前预测模式为宏块的

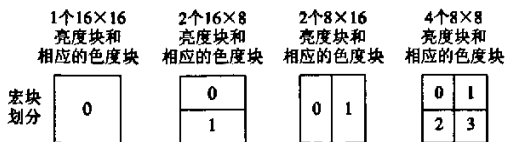


图 1 宏块划分方式

Fig. 1 Segmentation of MB

最佳预测模式,并退出帧间预测模式循环;如果不满足全零块判决条件,再比较 16×16 和 8×8 这两种预测模式的率失真值,如果 16×16 的率失真值小于 8×8 的率失真值,则选择 16×16 预测模式,并退出帧间预测模式循环,否则基于预测模式尺寸预选原则,计算宏块纹理复杂度,再由宏块纹理复杂度确定下一候选的预测模式是采用 16×8 ,还是 8×16 ,最后选择率失真值最小的预测模式。该算法总流程如图 2 所示。下面阐述流程中的主要部分。

3.2 基于宏块纹理复杂度预测模式尺寸预先选择

对于大小为 16×16 的宏块而言,如果待编码宏块的水平方向属于同一个物体,那么在帧间模式循环中选择 16×8 帧间预测模式的几率较大。这是因为这样的宏块在水平方向上纹理比较平滑,而在垂直方向上纹理复杂的缘故;如果待编码宏块的垂直方向属于同一个物体,那么在帧间模式循环中选择 8×16 帧间预测模式的几率较大^[5],这是由于这样的宏块在垂直方向上纹理比较平滑,而在水平方向上纹理复杂的缘故。通过上面的分析发现,可以通过对宏块纹理特征的分析来选择相应的帧间预测模式。本文用平均绝对差值 $MAD(MAD_v, MAD_h)$ 来分析图像纹理复杂度,其可用式(8)表示。

$$MAD = \frac{1}{16 \times 16} \sum_{y=0}^{15} \sum_{x=0}^{15} |p(x, y) - m| \quad (8)$$

式中, $p(x, y)$ 为宏块中的各个像素值, m 为宏块中所有像素的均值。

同理,可以求出垂直平均绝对差和水平平均绝对差,其分别可用式(9)和式(10)表示。

$$MAD_v = \frac{1}{16 \times 16} \sum_{x=0}^{15} \sum_{y=0}^{15} |p(x, y) - m_x| \quad (9)$$

$$MAD_h = \frac{1}{16 \times 16} \sum_{y=0}^{15} \sum_{x=0}^{15} |p(x, y) - m_y| \quad (10)$$

m_x 和 m_y 分别表示第 x 列和第 y 行所有像素的均值。

属于同一个物体的区域相对于属于不同物体或包含运动物体边界的区域来说,具有较小的 MAD 。

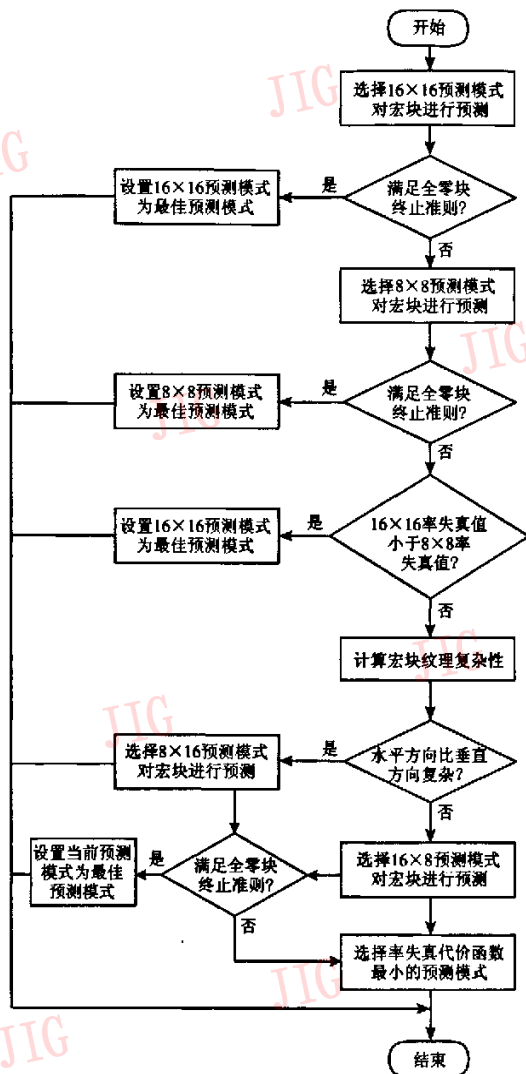


图 2 快速帧间预测模式选择算法流程图

Fig. 2 The flow chart of the fast inter mode decision algorithm

因此如果宏块垂直方向上的平均绝对差值大于水平方向上的平均绝对差值,则表明该宏块的水平方向有可能属于同一个物体。此时选择 16×8 帧间预测模式进行计算,否则,选择 8×16 帧间预测模式进行计算。

4 实验结果分析

为了验证本文提出的快速帧间预测模式选择算法的实用性,采用了不同运动特点的测试序列进行了相应的编码实验,包括 QCIF 格式 (176×144) 的

Akiyo、Miss_American 序列和 CIF 格式(352 × 288) 的 Foreman、Mother_daughter 序列;其中 Miss_American 序列运动比较小,场景基本不变化,而 Foreman 图像序列中的头部则有较大的运动,场景变化剧烈。熵编码采用 2D_VLC,量化参数分别为 28、30、32、34、36,编码 30 帧(IPPPPP...)。本文所提出的算法,与 AVS1-P2 的参考软件 rm52h 相比,不同 QP 下编码图像的峰值信噪比(PSNR)、码率和总编码时间的变化如表 1 所示。在表 1 中用 ΔB 表示比特率的改变, ΔP 表示图像亮度峰值信噪比的改变, ΔT 表示编码 30 帧所用总编码时间的改变,负号表示减少。

表 1 本文算法与参考软件代码 rm52h 的比较

Tab.1 The proposed algorithm compared with the reference code

序列	不同 QP 下与参考软件相比编码性能变化					
		28	30	32	34	36
Akiyo	$\Delta B(\%)$	1.44	1.00	1.32	1.35	0.62
	$\Delta P(\text{dB})$	-0.02	-0.05	-0.05	-0.05	-0.14
	$\Delta T(\%)$	-23.36	-23.46	-24.78	-24.92	-26.87
Miss American	$\Delta B(\%)$	0.61	0.00	1.13	2.62	0.84
	$\Delta P(\text{dB})$	-0.09	-0.09	-0.22	-0.20	-0.13
	$\Delta T(\%)$	-26.01	-27.13	-27.47	-28.70	-31.88
Foreman	$\Delta B(\%)$	0.69	0.91	0.93	1.26	1.45
	$\Delta P(\text{dB})$	-0.02	-0.07	-0.13	-0.13	-0.24
	$\Delta T(\%)$	-21.93	-22.94	-23.75	-24.28	-26.64
Mother_daughter	$\Delta B(\%)$	1.52	1.27	2.06	1.40	1.16
	$\Delta P(\text{dB})$	-0.10	-0.07	-0.15	-0.10	-0.08
	$\Delta T(\%)$	-23.90	-24.68	-25.59	-26.05	-28.77
Average	$\Delta B(\%)$	1.07	0.80	1.36	1.66	1.02
	$\Delta P(\text{dB})$	-0.06	-0.07	-0.14	-0.12	-0.15
	$\Delta T(\%)$	-23.80	-24.55	-25.45	-25.99	-28.54

从表 1 可以看出,虽然由于序列不同,编码的结果略有不同,但是总体来说,本文算法节省的编码时间是可观的,编码效率接近参考软件。由表中的数据可以看出,图像亮度的 PSNR 总体来说,变化不大,平均下降 0.11dB 左右,最大为 0.24dB。绝大多

数情况下,码率增加不超过 2%,平均增加 1.18% 左右。而编码时间则大大地节省,其中节省时间最大为 31.88%,节省时间最小为 21.93%。虽然由于提前退出原则,使得有些序列码率有所上升,但上升的幅度很小,而且随着量化参数的增大,本算法所能节省的编码时间也越多。

5 结论

本文提出了一种基于 AVS1-P2 的帧间预测模式快速选择算法,该算法的特点是将本文提出的基于全零块预先判别准则与基于纹理特性的预测模式尺寸预先选择等方法融合在一起。本文算法的总编码时间与 AVS1-P2 参考软件 rm52h 相比,节省时间最大为 31.88%,节省时间最小为 21.93%。而编码质量(PSNR)和码率几乎没有变化。这既有利于提高 AVS1-P2 在 PC 平台上的实时性,也有利于降低硬件实现的成本。

参考文献 (References)

- GB/T 200090.2-2006. Information Technology Advanced Coding of Audio and Video-Part 2: Video[S]. Feb. 2006. [GB/T 200090.2-2006. AVS1-P2(视频-基准)国家标准. 先进音视频编码 第二部分: 视频[S]. 2006 年 2 月.]
- JVT of ISO/IEC MPEG and ITU-T VCEG. Draft ITU-T Rec. and FDIS of Joint Video Specification[S], H. 264|ISO/IEC 14496-10 AVC, March 2003.
- Huang Hui, Cao Tie-yong, Zhang Xiong-wei. Research on key techniques of AVS1-P2 video standard[J]. Certification & Standard, 2006, (5): 59~61. [黄辉, 曹铁勇, 张雄伟. AVS1-P2 视频标准中的关键技术研究[J]. 认证与标准, 2006, (5): 59~61.]
- Zheng Yu, Liu Yu, Liu Wen-jie, et al. Design and implementation of AVS integer DCT transform & quantization[J]. Journal of Huaihai Institute of Technology (Natural Sciences Edition), 2006, 15(3): 22~25. [郑玉, 刘玉, 刘文杰等. AVS 整数 DCT 变换和量化方法的设计和实现[J]. 淮海工学院学报, 2006, 15(3): 22~25.]
- Shen Li-quan, Zhang Zhao-yang, Liu Zhi, et al. A fast H.264 intra prediction algorithm using all-zero block detection and texture feature[J]. Journal of Sichuan University, 2007, 39(1): 134~139. [沈礼权, 张兆扬, 刘志等. 一种结合全零块检测和纹理特征的快速帧内预测模式选择算法[J]. 四川大学学报, 2007, 39(1): 134~139]