

AVS 帧内模式选择优化

陈云善 苏宛新 王春霞 刘玉生

(中国科学院长春光学精密机械与物理研究所 长春 130033)

摘 要 AVS(Audio Video Coding Standard)采用率失真优化(Rate-Distortion Optimization, RDO)技术,通过计算所有预测组合模式的率失真代价来确定宏块的最优编码模式,使得计算复杂度大大增加,难以适应实时性要求较高的场合。为了降低 AVS 帧内预测模式选择的复杂度,在深入分析帧内预测原理以及预测模式选择过程的基础上,提出了一种基于 SATD(Sum of Absolute Transformed Differences)准则和空间相关性的快速帧内预测算法来优化帧内模式的选择过程。首先,利用 SATD 准则确定候选模式,以大幅度降低帧内预测模式选择的数量;然后,利用相邻宏块预测模式相关性,以进一步减少亮度块候选模式数量,提高模式选择速度。实验结果表明,与 AVS 标准算法相比,在编码效率相近的情况下,其编码速度能够平均提高 53.56%,与经典的基于边缘检测的方法相比,其在质量略有提高、码率还稍有降低的前提下,平均节约 16.39%的编码时间。

关键词 AVS, 帧内预测, 模式选择, SATD, 率失真优化(RDO)

中图分类号 TN919.81

文献标识码 A

Optimization of Intra Mode Decision for AVS

CHEN Yun-shan SU Wan-xin WANG Chun-xia LIU Yu-sheng

(Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China)

Abstract Rate-Distortion Optimization (RDO) technique is employed in AVS (Audio Video Coding Standard) to improve coding efficiency. However, the computational complexity increases drastically since the encoder has to encode the microblock by measuring RD costs of all the possible modes, which makes it difficult to meet the demands of real-time encoding. After the theory of the intra prediction and the process of the intra mode decision were analyzed, the fast intra prediction algorithm based on SATD (Sum of Absolute Transformed Differences) criterion and spatial correlation was proposed to optimize intra mode decision and reduce the complexity. First, some candidate modes were selected based on block's SATD which decreases the number of intra prediction mode. Then, mode correlation of neighbor blocks was used to reduce the complexity of luminance blocks and improve the speed of mode decision. Experimental results indicate that encoding time of the proposed approach is averagely reduced by 53.56% than that of AVS standard algorithm with similar encoding efficiency. Compared with the classical algorithm based on edge detection, the encoding time is saved by 16.39% on average while keeping better image quality and lower bite-rate.

Keywords AVS, Intra prediction, Mode decision, SATD, RDO

1 引言

作为先进音视频编码技术的重要组成部分, AVS 是由我国自主制定\拥有自主知识产权的视频编码标准。它以当前国际上最先进的 H. 264/AVC 框架为基础强调自主知识产权,同时充分考虑了实现的复杂度,其编码效率与 H. 264 相当,而算法复杂度仅为 H. 264 的 70%^[1]。AVS 的帧内预测部分,在传统的帧内编码技术的基础上,充分考虑了图像在局部空间上的高度相关性,周围的像素以多种模式对待编码宏块进行帧内预测,并利用 RDO 技术选择最优模式,从而减少了空间冗余,有效地提高了帧内编码的效率。与 H. 264 相

比,虽然 AVS 的帧内预测技术的复杂度降低了很多,但其运算复杂度依然很高。当最为复杂的运动估计部分运用快速运动搜索算法后^[2],帧内预测已成为影响其编码速度的主要瓶颈之一。因此,有必要优化帧内模式预测,降低模式选择过程的复杂度,以满足实时应用的需求。

AVS 标准提出较晚,目前针对其帧内预测的研究相对比较少,但它与 H. 264 采用了相似的技术,因此可以借鉴已有的针对 H. 264 的快速算法的思路。Pan 等人^[3]最早提出对帧内预测进行优化,采用 Sobel 算子作边缘检测,统计得到图像块的边缘方向直方图,据此判定图像纹理最强的方向,进而确定候选模式。受 Pan 方法的启发,很多学者利用最优的帧

到稿日期:2013-05-21 返修日期:2013-07-02 本文受国家自然科学基金(61203242)资助。

陈云善(1984—),男,博士,助理研究员,主要研究方向为达芬奇技术、视频编码算法等,E-mail: yiyunsn@163.com;苏宛新(1962—),男,研究员,主要研究方向为光电跟踪与测量技术、跟踪图像处理等;王春霞(1962—),女,高级实验师,主要研究方向为光电跟踪瞄准技术;刘玉生(1986—),男,硕士,主要研究方向为嵌入式系统开发。

内预测模式与视频图像块的纹理方向密切相关的特点,对其进行了一定改进^[4,5],但编码效率降低明显。邵娟等人^[6]则将文献[3]的思想应用到 AVS 帧内预测中,以边缘强度最强的方向和 DC 模式为候选模式,本文将其简称为 Edge 算法。这些算法在减少编码时间、保持编码性能方面各有优劣。本文在吸取相关快速算法优点的基础上,针对 AVS-P2 帧内预测的特点,提出了基于 SATD 准则和空间相关性的快速帧内预测算法(Fast Intra Prediction Algorithm Based on SATD criterion and Spatial correlation, FIPABSS)。FIPABSS 算法利用 SATD 准则筛选出可能性最大的两种模式,有效地降低了帧内预测的计算复杂度。实验结果表明,与 AVS 的标准算法相比,FIPABSS 算法在保证编码效率几乎不变的情况下,编码时间平均减少 53.56%。

2 AVS 帧内预测技术

2.1 AVS 帧内预测模式

为了充分利用相邻宏块间的空间相关性,AVS 采用空域内的多方向帧内预测技术,即利用相邻块的像素直接对当前块的像素进行预测,传输预测块和当前块差值的编码,更有效地去除相邻块之间的相关性,从而提高帧内编码效率。AVS 视频标准的帧内预测基于 8×8 块大小,对亮度分量和色度分量分别进行预测。亮度分量有 5 种预测模式,如表 1 所列。

表 1 8×8 亮度块帧内预测模式

IntraLumaPredMode	Name
0	Intra_8x8_Vertical
1	Intra_8x8_Horizontal
2	Intra_8x8_DC
3	Intra_8x8_Down_Left
4	Intra_8x8_Down_Right

图 1 给出了 8×8 亮度块帧内预测示意图。其中的 4 种预测方向分别与表 1 相对应,分别为模式 0(垂直预测)、模式 1(水平预测)、模式 3(左下预测)、模式 4(右下预测),模式 2(DC 预测)没有预测方向。当前帧内预测块由其上边和左边的参考样本 $r[i](i=0..16)$ 和 $c[i](i=0..16)$ 来预测,其中 $r[0]$ 等于 $c[0]$ 。色度块的帧内预测模式和亮度块类似,分别为:模式 0(DC 预测)、模式 1(水平预测)、模式 2(垂直预测)、模式 3(平面预测),相同位置的两个色度块 Cb、Cr 具有相同的最佳模式。

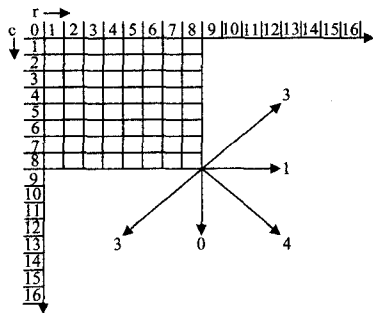


图 1 8×8 亮度块帧内预测方向示意图

2.2 AVS 帧内模式选择过程

AVS 对每个宏块进行帧内编码时,在可选的模式中遍历每种模式,计算出每种模式的图像失真和所需的比特数,利用

式(1)比较各模式率失真代价,选取率失真代价最小的模式作为最佳编码模式。

$$J(s, c, \text{MODE} | QP) = D(s, c, \text{MODE} | QP) + \lambda_{\text{MODE}} R(s, c, \text{MODE} | QP) \quad (1)$$

式中, QP 为量化参数; MODE 是宏块候选编码模式; D 为原始信号 s 与其重建信号 c 的差值平方和,作为失真度量; R 为量化参数 QP 和模式 MODE 下编码此宏块所需的比特数(包括宏块头信息、运动矢量及残差变换量化后的系数等)。其中 λ_{MODE} 为拉格朗日参数,由式(2)计算。

$$\lambda_{\text{MODE}} = 0.85 \times 2^{(QP-11)/4} \quad (2)$$

由于帧内编码对亮度和色度帧内预测模式的每一种组合都要利用式(1)进行计算,从中筛选一个率失真最小的组合模式,因此每一个宏块都要进行 $M8_chroma \times (M8_luma \times 4)$ 次 RDO 计算, $M8_chroma$ 和 $M8_luma$ 分别表示 8×8 色度块、 8×8 亮度块的预测模式个数,这就意味着对每个宏块都要进行 80 次 RDO 计算,这极大地增加了编码器的计算复杂度,很难适应实时性要求较高的场合。

3 快速帧内预测算法

通过以上分析可以看出,虽然 AVS 编码标准采用率失真优化技术可以找到最优预测模式,但其复杂度很高,急需运用快速算法降低其复杂度,提高编码速度。如果能预测出最可能成为最佳模式的几种模式作为候选模式,而排除那些不太可能提高编码效率的模式,就可以有效地降低复杂度。

3.1 FIPABSS 算法的提出

FIPABSS 算法的提出,主要基于以下两点原则:

(1) SATD 与率失真性能的相关性原则

当前编码块原始像素值与预测像素值差经哈达玛变换后所得系数绝对值之和(SATD)可以看作是简单的时频变换,它在一定程度上可以反映生成码流的大小,因此将其作为模式选择的依据,可以同时兼顾预测模式对图像失真和码率的影响。通常 SATD 值越小越有可能是最优模式,显然选取的候选模式越多,编码性能越好,但编码复杂度越高。为了描述方便,将 SATD 值最小和次最小两个模式分别定义为 SATD_MIN 和 SATD_SECOND。表 2 给出了 3 个典型 CIF 格式测试序列的亮度块以 SATD_MIN 和 SATD_SECOND 作为候选模式时,预测正确的概率。可见,准确度均可以达到 84% 以上,对大量测试序列进行的实验发现此时编码效率并没有明显地下降,而色度块直接以 SATD 值最小的模式作为最佳模式即可获得较好的编码效果。由此,RDO 计算就由原来的 80 次减少为 8 次,极大地降低了复杂度。

表 2 使用 SATD 时预测准确度

Sequences	QP=28	QP=32	QP=36
Akiyo	90.75%	87.18%	90.05%
Foreman	88.56%	85.35%	87.90%
Mobile	86.25%	84.80%	85.54%

(2) 相邻宏块预测模式相关性原则

自然图像中相邻块之间的帧内预测模式有着较强的相关性,已编码的左边块和上边块的预测模式很有可能是当前块的最佳模式。因此,可以根据当前块相邻块的最佳模式预测

出当前块最可能的最佳模式(Most Probable Mode,MPM)。MPM确定的具体方法是:如果当前块的左边块和上边块的帧内预测模式都可用,则MPM为两个模式中值较小的模式,否则MPM为DC预测模式。AVS标准利用MPM减少编码帧内预测模式所需要的比特数,如果当前块的最佳帧内预测模式和其MPM相同,则只需1bit确认,无需编码当前块的最佳帧内预测模式值。

表3 MPM为最佳模式的概率

Sequences	probability	QP=28	QP=32	QP=36
Akiyo	A	61.42%	64.45%	67.53%
	B	59.50%	68.81%	68.25%
	C	89.66%	83.30%	84.98%
Foreman	A	48.42%	50.15%	52.98%
	B	41.47%	47.23%	46.97%
	C	86.54%	82.22%	87.96%
Mobile	A	45.55%	46.71%	47.74%
	B	41.40%	43.47%	43.21%
	C	81.54%	80.64%	82.49%

表3列出了不同QP下MPM为当前亮度块最佳模式的概率。其中,A栏是MPM为最佳模式的概率,可以看出45%以上的宏块最佳模式即为MPM,对于纹理较为简单的Akiyo序列,MPM为当前块最佳模式所占比例达到60%以上。B栏代表MPM和SATD_MIN相等的概率,C栏表示当MPM=SATD_MIN时,MPM为最佳模式的概率,可见此时把MPM当作亮度块最佳模式的正确率在80%以上,由此,有40%以上的块可直接确定最佳模式,仅需一次RDO计算。

3.2 FIPABSS算法具体步骤

(1)基于以上分析,本文提出了基于SATD准则和空间相关性的快速帧内预测算法(FIPABSS),其流程如图3所示,具体步骤如下:计算各模式的SATD值,比较各模式的SATD值大小,得出SATD_MIN和SATD_SECOND,同时计算当前块的MPM值,如果MPM与SATD_MIN相同,则以当前模式为最佳模式,转步骤(3);

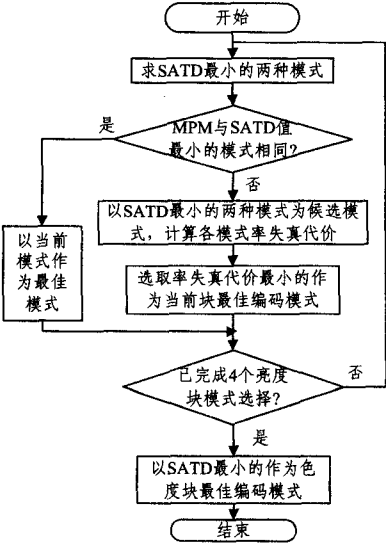


图2 FIPABSS算法流程图

(2)以SATD_MIN和SATD_SECOND两种模式为候选模式,利用式(1)计算它们的率失真代价,以率失真代价最小的模式作为最佳模式;

(3)重复步骤(1)–(2),直到求得当前宏块所有亮度块的

最佳帧内预测模式;

(4)对于两个色度块,计算4种模式下的SATD值,SATD值最小的模式即为最佳色度模式。

可见,运用本文算法后,每个亮度块的预测模式数为2种,而色度块仅1种候选模式,因此,完成一个宏块的帧内预测最多需要8次亮度块RDO运算,最好的情况下,只需4次,而Edge算法^[6]中基于边缘检测的方法需要进行24次或16次RDO运算。

表4 预测准确度和平均预测模式个数的比较

Sequence	预测准确度		平均预测模式个数	
	FIPABSS	Edge	FIPABSS	Edge
Akiyo	89.4%	77.8%	1.284	2.658
Foreman	86.8%	77.1%	1.496	2.655
Mobile	85.2%	81.5%	1.505	2.563

表4给出了QP=30时,使用本文算法后,亮度块预测准确度和平均预测模式个数与Edge算法的比较。可见本文算法在复杂度和准确度两方面均优于经典的基于边缘检测的方法。

4 实验结果与分析

为了检验FIPABSS算法的有效性,将其在AVS编解码器rm52j上进行实现。本实验选取具有不同空间细节的10个标准测试序列的前100帧进行测试,各序列均为CIF格式,帧率30f/s。QP值分别选取为28、32、36,全部采用I帧编码,开率失真优化。

仿真环境:AMD Athlon(tm) 64 X2 Dual Core Processor 4800+ 2.51GHz,2GB,VS2005。

表5 FIPABSS算法和AVS标准算法编码性能对比

Sequence	QP=28			QP=32			QP=36		
	Δs (%)	Δb (%)	Δt (%)	Δs (%)	Δb (%)	Δt (%)	Δs (%)	Δb (%)	Δt (%)
Akiyo	-0.01	0.97	-51.74	-0.02	2.00	-52.84	-0.07	2.20	-52.56
News	-0.01	0.81	-52.54	-0.04	1.26	-53.20	-0.03	2.34	-52.98
Foreman	-0.01	0.75	-52.42	-0.03	1.09	-52.59	-0.03	2.65	-51.25
Football	-0.01	0.59	-51.35	-0.01	1.20	-53.51	-0.01	1.70	-51.36
City	-0.01	0.26	-52.21	-0.03	0.52	-52.45	-0.05	0.77	-52.27
Bus	-0.01	0.35	-53.40	-0.02	0.42	-53.55	-0.03	0.91	-52.64
Paris	-0.01	0.44	-53.51	-0.01	0.63	-54.29	-0.02	0.84	-53.21
Flower	-0.02	0.38	-55.28	-0.05	0.47	-55.21	-0.05	0.52	-55.05
Mobile	-0.01	0.35	-55.02	-0.02	0.40	-55.49	-0.04	0.48	-54.47
Husky	-0.01	0.22	-56.76	-0.02	0.18	-56.83	-0.03	0.21	-56.84
Average	-0.01	0.51	-53.42	-0.03	0.82	-54.00	-0.04	1.26	-53.26

表6 FIPABSS算法和Edge算法编码性能对比

Sequence	QP=28			QP=32			QP=36		
	Δs (%)	Δb (%)	Δt (%)	Δs (%)	Δb (%)	Δt (%)	Δs (%)	Δb (%)	Δt (%)
Akiyo	0.01	-2.35	-18.81	-0.01	-1.50	-20.98	-0.01	-1.70	-21.69
News	0.02	-1.77	-17.68	0.00	-1.53	-21.27	0.04	-0.60	-21.54
Foreman	0.01	-0.61	-14.47	0.01	-0.53	-16.26	0.02	0.62	-16.76
Football	0.01	-0.87	-13.25	0.01	-0.70	-17.80	0.02	-0.16	-15.00
City	0.01	-0.51	-11.24	0.00	-0.37	-16.20	0.00	-0.27	-14.44
Bus	0.01	-0.59	-14.57	0.01	-0.63	-17.75	0.02	-0.25	-16.55
Paris	0.02	-0.66	-12.02	0.02	-0.60	-16.24	0.03	-0.56	-14.84
Flower	0.00	-0.22	-13.70	-0.01	-0.09	-16.33	-0.01	-0.03	-17.33
Mobile	0.01	-0.43	-13.09	0.01	-0.46	-15.13	0.02	-0.45	-15.63
Husky	0.00	-0.30	-16.39	0.00	-0.30	-16.94	0.01	-0.23	-17.75
Average	0.01	-0.83	-14.52	0.00	-0.67	-17.49	0.01	-0.36	-17.15

将FIPABSS算法与AVS标准算法、经典的Edge算法在

3个方面进行了对比:(1)结构相似度(SSIM, structural similarity):本文用比峰值信噪比(PSNR)更符合人眼视觉特性的结构相似度来衡量编码重建图像和原图像的差别。结构相似度的值为介于0和1之间的浮点数,它反映的是两个图像块的相似程度,SSIM值越大说明这两个图像块越相似^[7-11];(2)码率:用来反映编码压缩效率,单位为 kbit/s;(3)编码时间:编码总时间,代表编码复杂度,单位为 s。为了便于比较,用 Δs 、 Δb 和 Δt 分别表示结构相似度、码率和编码时间变化的百分比,其中正值表示增加,负值表示减少。

FIPABSS算法和 AVS 标准算法编码结果对比如表 5 所列,其中各序列纹理复杂度从上到下依次增加。从表中数据可以看出,FIPABSS算法与 AVS 标准算法相比,在获得相近编码效率的前提下,编码时间大幅减少,10 个测试序列的编码时间平均减少 53.56%,而 SSIM 平均仅下降 0.027%,码率平均增加 0.86%。另外,由表 5 可见,随着序列纹理复杂度的增加,编码时间也相应减少,这主要是由于纹理较复杂的序列本身编码时间较长的缘故。此外,由于 FIPABSS 算法综合运用了前文所述两点原则,对于具有不同纹理复杂度的序列,在不同 QP 下,虽然节约的编码时间略有不同,但基本相当。可见,本文算法具有较好的适应性。

表 6 给出了 FIPABSS 算法和 Edge 算法的编码性能对比。可见,在编码质量略有提高的情况下,编码时间平均减少 16.39%,同时码率还有 0.62%的减少。而细节较少的序列编码速度提高 20%左右,码率节省也比较多,可见 FIPABSS 算法在编码质量、码率和编码时间方面均优于 Edge 算法。

为了更直观地比较编码质量,在图 3 中给出了 QP=30 时,Akiyo 序列第 100 帧各算法的重建图像。可见,FIPABSS 算法的重建图像与 AVS 标准算法和 Pan 算法的重建图像质量基本一致。



图 3 QP=30 时,Akiyo 序列第 100 帧编码质量对比

图 4 和图 5 分别给出了 QP=30 时,Akiyo 序列每帧编码时间和每帧 SSIM 值的对比结果。可以看出,FIPABSS 算法每帧编码时间均明显少于 AVS 标准算法和 Edge 算法。而在编码质量方面,FIPABSS 算法也优于 Edge。

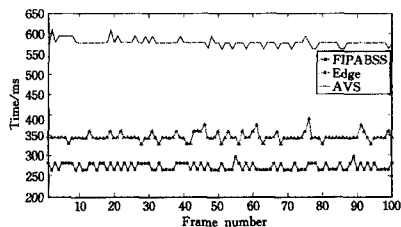


图 4 Akiyo 序列每帧的编码时间

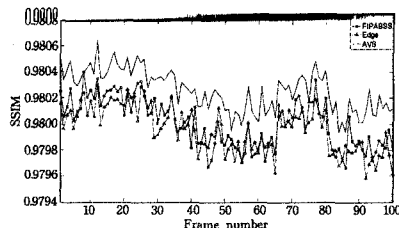


图 5 Akiyo 序列每帧的编码质量

由以上结果可知,FIPABSS 算法对各种测试序列都有较好的适应性,帧内模式选择复杂度明显降低,从而有效地减少了编码时间,而重建图像质量基本不变,对主观视觉的影响可以忽略不计。

结束语 本文在分析 AVS 帧内预测原理以及预测模式选择的基础上,提出了基于 SATD 准则和空间相关性的快速帧内预测算法,该算法利用 SATD 与率失真性能的相关性和相邻宏块预测模式相关性两个原则对候选模式进行筛选,减少了预测模式数量,从而有效地降低了 AVS 帧内预测复杂度。实验测试结果表明,本文算法有较好的适应性,对于不同测试序列均有很好效果,在获得相近编码效率的情况下,较 AVS 标准算法和 Edge 算法分别平均节省了 53.56%和16.39%的编码时间,极大地提高了帧内编码速度,为 AVS 在工程中的实时应用打下了良好的基础。

参考文献

- [1] 国家标准化委员会. GB/T 200090. 2-2006. 信息技术先进音视频编码(第 2 部分):视频[S]. 北京:中国标准出版社
- [2] 祝世平,申晓东. 十字交叉六边形块运动的估计搜索[J]. 光学精密工程,2009,17(12):3069-3076
- [3] Pan F, Lin X. 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
- [4] 何宜宝,陈再秀,毕笃彦,等. 利用边缘方向检测实现 H. 264 帧内预测编码的新算法[J]. 西安电子科技大学学报:自然科学版, 2009, 36(6):1126-1131
- [5] 曹雁,刘宁钟,崔子冠. 基于边缘方向直方图的 H. 264 帧内预测快速算法[J]. 计算机技术与发展,2008,18(5):60-63
- [6] 邵娟,张卫宁,魏磊,等. AVS 中帧内预测模式的快速选择策略[J]. 计算机工程与应用,2009,45(25):163-165
- [7] Wang Z, Bovik A C, Lu L G. Why is image quality assessment so difficult [C]//Proc of IEEE Int Conf on Acoustics, Speech, and Signal Processing. Orlando:[s. n.],2002:3313-3316
- [8] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment:from error visibility to structural similarity [J]. IEEE Trans on Image Processing,2004,13(4):600-612
- [9] 杨春玲,王华兴. 基于结构相似度的 H. 264 快速运动估计算法[J]. 华南理工大学学报:自然科学版,2008,36(8):28-32
- [10] 赵连凤,郭宝龙,韩合民. 基于结构相似度的快速运动估计算法[J]. 计算机应用研究,2007,24(11):306-308
- [11] 陈云善,高慧斌,苏宛新,等. 结构相似度在 AVS 帧间模式选择中的应用[J]. 光电子·激光,2011,22(3):435-439