

基于通用串预测算法的 AVS2 屏幕混合内容 帧间编码优化

赵利平^{1),3)} 林 涛²⁾ 郭 靖²⁾ 周开伦²⁾

¹⁾(绍兴文理学院计算机科学与工程系 浙江 绍兴 312000)

²⁾(同济大学超大规模集成电路研究所 上海 200092)

³⁾(嘉兴学院数理与信息工程学院 浙江 嘉兴 314000)

摘 要 我国第二代数字音视频编解码技术标准的屏幕与混合内容扩展版(AVS2-SMCC)中,对于全色度屏幕混合内容的编解码其采用融合 4:4:4 格式的通用串预测编码算法和传统 4:2:0 格式的块预测、变换编码算法的编解码框架。在此框架中,根据屏幕混合内容的内在帧间特性和不同帧间子模式的特点,该框架对帧间编码算法进行了优化:首先提出了支持 4:4:4 格式的帧间编解码算法;其次,提出了内容自适应的不同帧间子模式色度失真权重因子算法和帧间子模式预编码与编码树单元划分的提前终止算法;最后,提出了一种新的帧间第一受限串模式,即将通用串预测算法的第一受限串的参考位置扩展至参考图像队列中的第一帧。实验结果表明,对于 AVS2-SMCC 标准测试序列中的移动文本和图形类别序列,提出的算法和 AVS2-SMCC 参考软件 RMD1.0 相比,在有损低延迟配置(一种含帧间编码的配置)下,Y、U、V 分量的平均 BD-rate 下降率分别可达 14.35%、90.73%、88.85%,编码时间降低 50.86%,解码时间增加 15.24%。

关键词 高效视频编码;音视频编解码技术标准;屏幕混合内容编码;帧间编码;快速算法

中图法分类号 TP391

DOI号 10.11897/SP.J.1016.2019.02190

Universal String Prediction-Based Inter Coding Algorithm Optimization in AVS2 Mixed Content Coding

ZHAO Li-Ping^{1),3)} LIN Tao²⁾ GUO Jing²⁾ ZHOU Kai-Lun²⁾

¹⁾(Department of Computer Science and Engineering, Shaoxing University, Shaoxing, Zhejiang 312000)

²⁾(Institute of VLSI, Tongji University, Shanghai 200092)

³⁾(College of Mathematics, Physics and Information Engineering, Jiaxing University, Jiaxing, Zhejiang 314000)

Abstract With the rapid development of technologies in networking and thin-client devices, such as remote desktop, video conferencing with documents or slides sharing, television programs and many curriculum videos playing in the Internet that contain mixtures of camera-captured content (CC) and screen content (SC), efficient screen content coding (SCC) becomes a hot topic in multimedia applications and has attracted increasing researcher attention from both academia and industry. Two international video coding standards include efficient SCC capability. One is High Efficiency Video Coding (HEVC), the other is the second-generation of the Audio Video Coding Standard (AVS2). In recent years, AVS is developing an AVS2 Screen and Mixed Content Coding (SMCC) extension (AVS2-SMCC). In our previous work, a 4:4:4 format universal string

收稿日期:2018-01-10;在线出版日期:2019-03-14. 本课题得到浙江省自然科学基金项目(LY19F020015)、国家自然科学基金项目(61871289,61601200)、绍兴市公益性技术应用研究计划项目(2018C10015)、上海市自然科学基金项目(18ZR1440600,19ZR1461100)资助。

赵利平,博士,副教授,中国计算机学会(CCF)会员,主要研究方向为视频编码、屏幕图像编码。E-mail: zhaoliping_jian@126.com.

林 涛(通信作者),博士,博士生导师,长江学者特聘教授,主要研究领域为视频编码、屏幕图像编码。E-mail: lintao@tongji.edu.cn.

郭 靖,硕士,主要研究方向为屏幕图像编码。周开伦,博士,讲师,主要研究方向为视频编码、超大规模集成电路架构、设计与验证。

prediction (USP) algorithm integrated with the tradition 4:2:0 format block prediction and transform coding framework is used to code 4:4:4 format mixed content screen. To fully exploit both local and non-local, both general and special, and both complex and simple matching patterns with a variety of sizes and/or shapes and/or positions in a wide range of commonly seen screen content, a USP approach and its key technologies with three modes: general string (GS) mode, constrained string 1 (CS1) mode, and constrained string 2 (CS2) mode are proposed. The three constrained string modes are implemented with one of the three types of strings: offset string, coordinate string, and unpredictable pixel, or its combination of them. When using USP algorithm to code a coding unit, one of the three constrained modes which gets the minimum rate distortion value is selected to code the CU. Experimental results show that, for text and graphics with motion of AVS2-SMCC test sequences, USP algorithm achieves significantly improved coding efficiency in the All Intra configuration when compared with the latest HEVC-SCC extension. In order to further improve the coding efficiency in the Low Delay configuration and reduce the complexity, in this framework, according to the inherent inter-frame characteristics of the mixed content and the different characteristics for different inter-sub-coding modes, an optimized inter coding algorithm which consists of a 4:4:4 format inter-coding algorithm, a content adaptive weighted chroma distortion algorithm for different inter-sub-coding modes and an early terminate algorithm for inter-sub-coding modes pre-coding and coding tree unit partition, and a newly inter constrained string 1 mode (the search area of CS1 mode extends to the first frame in the reference picture queue) are proposed. Compared with the reference software RMD1.0 of AVS2-SMCC, for the category of text and graphics with motion in the AVS2-SMCC test sequences, the proposed algorithm achieves the average BD-rate reduction up to 14.35%, 90.73% and 88.85% respectively, in lossy low delay configuration, with a nearly 50.86% decrease in encoding runtime and 15.24% increase in decoding runtime. Especially for the video category of text and graphics with motion, the proposed algorithm achieves significantly improved coding efficiency not only in the All Intra configuration but also in the Low Delay configuration when compared with the HEVC-SCC extension.

Keywords high efficiency video coding; audio video coding standard; screen mixed content coding; inter coding; fast algorithm

1 引言

屏幕混合内容编码 (Screen Mixed Content Coding, SMCC) 也称作屏幕图像编码 (Screen Content Coding, SCC), 是当前云计算、云教育、云游戏、云桌面、屏幕分享、智能手机和平板电脑的第二显示、无线投屏等应用中解决数据传输瓶颈问题的关键技术^[1-4].

屏幕图像是指直接从各类设备的屏幕图像显示单元捕获的视频或图像. 与传统视频相比, 屏幕图像主要由平滑的、具有丰富纹理的、连续色调的、相机拍摄的自然图像区域和具有尖锐的像素边缘的、不连续色调的、计算机产生的非自然图像区域 (如文

本、图形、图标和图表等) 两部分组成. 从视频编码的角度来看, 屏幕图像表现出显著不同的特点. 这些特点不仅包括其帧内特性, 也包括其帧间特性.

(1) 帧内特性. 一个编码块 (Coding Unit, CU) 中不同颜色的数目迥异; 图像中存在具有各种各样的形状、大小和相对位置的匹配图样; 一帧图像中匹配图样的数目很不相等等.

(2) 帧间特性. 文本、图表等非自然图像或视频区域中, 存在着大量的静止区域, 这些区域的原始像素与前一帧相同位置的原始像素是完全一致的. 而自然图像或视频区域中的静止区域由于自然界中光效的原因导致原始像素和前一帧相同位置的原始像素只是接近的; 一般由计算机软件产生的运动是整像素的运动, 而由摄像机记录下来的对信号的时域

采样产生的自然运动则需要利用分像素插值得到更精确的运动估计。

针对屏幕图像的这些内在特性,工业界和学术界开展了高效的屏幕图像编码标准工作:主要包括由国际电信联盟、国际标准化组织和国际电工委员会三大国际组织联合制定的 HEVC(High Efficiency Video Coding)^[5] SCC 扩展版^[2-5] 和由我国数字音视频编码技术标准(Audio Video Coding Standard, AVS)工作组制定的第二代 AVS(简称 AVS2)^[6-7] 标准的屏幕混合内容视频编码(Screen and Mixed Content Coding, SMCC)扩展版^[8] 两大标准。

近年来,屏幕混合内容视频编码算法主要针对屏幕图像的帧内特性提出。主流的去除重复样图冗余的主要算法包括:HEVC 版或 HEVC-SCC 版的标准制定期间提出 SM(String Matching)^[9]、PLT(Palette)^[10-11]、IBC(Intra Block Copy)^[12]、ILC(Intra Line Copy)^[13]、P2SM(Pseudo 2D String Matching)^[14] 和 USM(Universal String Matching)算法^[15] 和 AVS2-SMCC 扩展版的标准制定期间提出的对不同屏幕内容特性的区域自适应采用一般串模式、第一受限串模式和第二受限串模式三种受限串模式之一进行编解码的通用串预测算法(Universal String Prediction, USP)^[16-18]。目前,USP 算法对于 AVS2-SMCC 通用测试序列中的移动的文字和图形类别,在编解码复杂度增加较少、有损全帧内配置的情况下,Y 分量、U 分量、V 分量平均 BD-rate 降低率分别可达 23.2%、18.3%、19.4%,编码效率已超越 HEVC-SCC^[18]。

一般情况下,视频序列可看作物体的平滑运动,除运动的物体和场景变化之外,大部分内容相对于前一帧相同位置是接近静止的。屏幕混合内容也是如此。在低延迟配置(一种含帧间编码的配置)下,如何根据屏幕混合内容的帧间特性消除时域冗余至关重要。文献[19]提出了一种快速的帧间哈希块匹配算法。文献[20]提出了一种带权重的率失真方案,通过降低多次被参考的编码块的编码误差来提高屏幕图像的编码效率。文献[21]提出了低延迟配置下 Diversity-Based 的屏幕图像参考帧管理方案。这些算法都是在 HEVC-SCC 标准中实现的。

本文为了进一步提高 AVS2-SMCC 中含帧间编码低延迟配置的编码效率,在基于 AVS2-SMCC 的 USP 算法基础上,充分考虑了屏幕混合内容的内在帧间特性和不同帧间子模式的特点,做了以下优化工作:(1)实现了支持 4:4:4 格式的帧间编解码

框架;(2)提出了内容自适应的不同帧间子模式色度失真权重因子算法;(3)提出了帧间子模式预编码和编码树单元划分的提前终止算法;(4)将通用串预测算法的第一受限串的参考位置扩展至参考图像队列中的第一帧,提出了一种新的帧间第一受限串模式;(5)对所有各种色度格式的编码模式的编码单元都统一采用 4:4:4 格式的环内滤波,进一步提高了色度分量的编码质量。实验结果表明,与 AVS2-SMCC 参考软件 RMD1.0 相比,提出的帧间优化算法(简称 USP_Inter 算法)在编码时间降低 50.86%,同时解码时间增加 15.24%的前提下,低延迟配置的部分序列类别编码效率显著提高。提出的算法基于 AVS2 提案^[22-23]。

提出基于 AVS2 框架的 USP_Inter 算法与基于 HEVC-SCC 标准框架的文献[20]和[21]在很大程度上是互相独立,没有重叠部分,不具有竞争性、排他性和对比性,而是互补共赢的。因此,对编码效率的提升也是叠加的,各种方法都能够在其它方法基础上进一步提升编码效率。

2 相关工作及算法提出的动机

2.1 AVS2-SMCC 现状

2016 年底颁布的被列为国家标准的 AVS2 标准,其性能已优于同期 H.265/HEVC 国际标准。尤其是对于场景视频编码,AVS2 标准具有显著优势,能够比 HEVC 节省 40%左右的码率^[7]。

AVS2 中采用基于块的预测变换编码技术。具体而言,其采用了基于四叉树递归划分的块预测编码架构。在 AVS2 中,最大编码单元(Largest Coding Unit, LCU)大小为 64×64 ,最小编码单元(Coding Unit, CU)大小为 8×8 。AVS2 帧间编码的特色包括:除了传统的 P 帧和 B 帧预测,还定义了 F 帧,同时支持单前向和双前向帧间预测,使用跳过/直接(skip/direct)模式、时域多假设预测模式和空域方向性多假设模式等编码技术。帧间模式的预编码的顺序是:支持 $2N \times 2N$ 和 $N \times N$ 划分的 skip/direct 子模式、 $2N \times 2N$ 、 $2N \times N$ 、 $N \times 2N$ 、 $2N \times nU$ 、 $2N \times nD$ 、 $nL \times 2N$ 、 $nR \times 2N$ 和 $N \times N$ 的预测单元划分子模式和四叉树递归划分的方式,从而进一步提高编码效率。与 HEVC 不同的是,AVS2 的四叉树递归划分的方式是从 8×8 像素递归至 64×64 像素的编码单元。

2015 年 8 月,AVS 正式启动了 AVS2 标准的

屏幕混合内容视频编码的制定工作^[8]. AVS2 仅支持 4:2:0 采样格式的视频编码,而屏幕图像往往具有很多不连续的区域,采用 4:2:0 和 4:2:2 采样格式会对色度分量造成较大的损失,特别是分散的线条、渲染的文字等区域可能出现较差的主观质量问题,因此屏幕图像的测试序列一般是 4:4:4 采样格式的.为了解决该问题和提高屏幕图像的编码效率,文献[17]提出了融合像素串匹配的 AVS2 全色度屏幕与混合内容视频编码算法.在此基础上,文献[18]提出了由一般串模式、第一受限串模式和第二受限串模式三种受限串模式组成的通用串预测算法(USP)及其关键技术.

目前,基于 USP 算法的 AVS2-SMCC 的全帧内配置的性能已基本超过 HEVC-SCC 扩展版,充分表明了 USP 算法能大幅提高屏幕混合内容的帧内编码性能.但是其低延迟 F 帧(Low Delay F, LDF)配置的帧间编码性能有待进一步提升.

在现有集成了 USP 算法的 AVS2-SMCC 框架中,通用串预测模式是采用 YUV4:4:4 格式编码模

式,而 AVS2 的帧内和帧间模式只支持 YUV4:2:0 格式.不同模式的色度失真组成不同,从而导致 RDcost 中色度分量的失真计算方式不平衡.

表 1 列出了帧内、帧间和 USP 预测模式的色度失真情况.传统的帧内和帧间编码过程中,除量化失真之外,还有上下采样造成的失真.

表 1 AVS2-SMCC 中不同编码模式失真组成		
编码模式	支持采样格式	色度失真组成
通用串预测	YUV4:4:4	量化
帧内预测	YUV4:2:0	上下采样和量化
帧间预测	YUV4:2:0	上下采样和量化

文献[18]通过设置不同的色度失真调节因子使得整体的编码效率进一步提高,对于全帧内编码配置非常有效.但是其对于帧间编码效率的提升有限.

图 1 给出了 AVS2-SMCC 标准测试序列之一的 webb 序列(见表 2)LDF 配置下(配置量化参数为 22)前 30 帧的编码图像质量和编码模式选择百分比的统计分析结果.从图中我们可以发现:

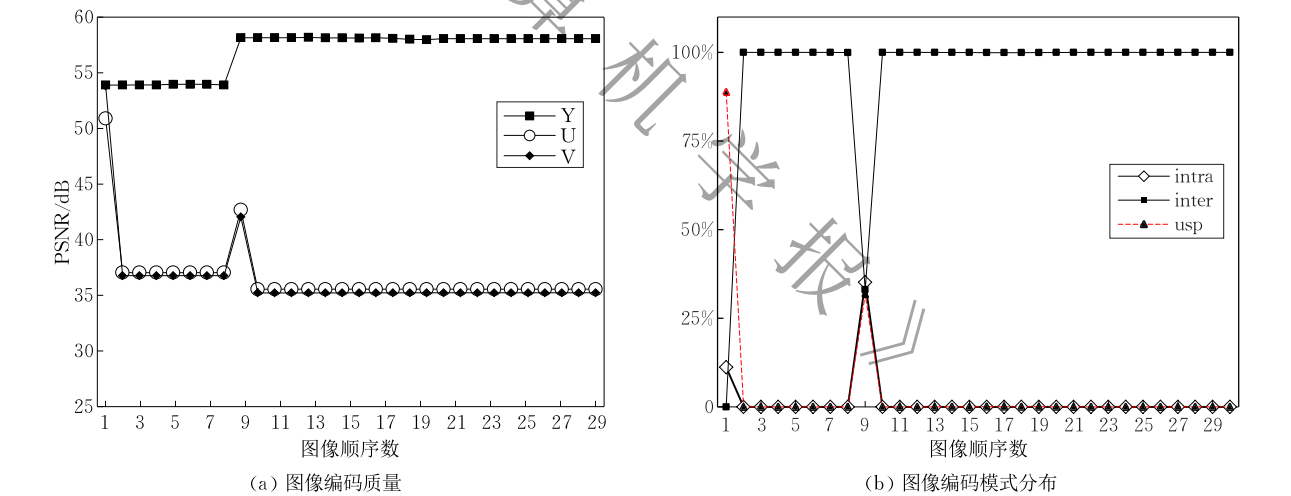


图 1 webb 序列前 30 帧 LDF 配置编码分析

(1)从图 1(a)中图像编码质量来看,亮度分量的 PSNR 能保持与第一帧或者前一个场景变化图像接近.而色度分量的 PSNR 从第二帧开始便急剧下降,后续帧的编码质量保持在该值左右,直到下一个场景变化的图像(由于 USP 模式的 YUV4:4:4 格式的色度分量由于没有色度上下采样,这一帧的 U 和 V 分量的 PSNR 提升,但是由于还是存在 30%左右的帧间和帧内模式,色度分量的 PSNR 提升幅度较小).后续随着选择帧间模式的比例增加到接近 100%,相应的色度 PSNR 保持在一个较低的 PSNR 值.

(2)从图 1(b)中图像的编码模式分布来看,第一帧图像选择 USP 编码模式的比例在 90%左右;之后的图像中,选择帧间编码的比例接近 100%,一直到第 9 帧发生场景切换,选择 USP 和帧内编码模式的比例升高,帧间编码模式的比例降低;随后选择帧间编码的比例再次接近 100%.

综合以上发现,AVS2-SMCC 框架中 LDF 配置的帧间编码性能亟需进一步提升的原因主要包括:

(1)一方面,基于 YUV4:2:0 的帧间预测模式对 YUV4:4:4 的屏幕内容进行低延迟配置的编码时,色度分量的上下采样带来的失真对编码效率影

响较大,导致效率不高,因此亟需支持 YUV4:4:4 的帧间预测编码模式.

(2)另一方面,未考虑屏幕内容的内在帧间特性和 AVS2-SMCC 中不同帧间子编码模式的特点,因此亟需根据其特点进行帧间编码算法的优化.

2.2 屏幕内容帧间特性和帧间子模式分析

表 2 是 AVS2 屏幕混合内容标准测试序列集^[24],包含 13 个测试序列,分为 3 类:

- (1)移动的文字和图形(Text and Graphics with Motion,TGM)类别;
- (2)含自然视频、不规则纹理或渐变亮度场景的混合内容(Mixed Content,MC)类别;
- (3)游戏动画(Game,G)类别.

表 2 AVS2-SMCC 标准测试数据			
分辨率	测试序列名称(缩写名)	序列特征	类别
1920×1080	sc_flyingGraphics (flyg)	移动或旋转的图形	TGM
	sc_desktop (dsk)	桌面常见编辑环境	TGM
	sc_console (cns)	命令行窗口等	TGM
	ChineseDocumentEditing (cde)	中文 Word 和 PPT 编辑	TGM
	EnglishDocumentEditing (ede)	英文 Word 和 PPT 编辑	TGM
	ClearTypeSpreadsheet (sps)	Excel 使用环境	TGM
	BitstreamAnalyzer (bsa)	码流分析软件使用	TGM
	CircuitLayoutPresentation (clp)	集成电路图、版图展示	TGM
	PeopleInVehicle_Spreadsheet (pvs)	视频播放与 Excel 使用	MC
	sc_web_browsing (webb)	网页浏览	MC
1280×720	sc_programming (prg)	编程及运行环境	MC
	sc_map (map)	在线地图使用环境	MC
	sc_robot (rob)	精细渲染的动画	G

以上测试序列都是 YUV4:4:4 格式,像素样本为 8 比特.其中序列 flyg、dsk、scns、webb、prg、map 和 robot 也是 HEVC-SCC 的标准测试序列^[25].本文的实验分析中,编码帧数为 30 帧,编码配置为低延迟 F 帧(LDF),即配置是 IFFFF……(第一帧是 I 帧,后续全是 F 帧),量化参数(Quantization Parameter,QP)为 25、32、38 和 45.实验使用 AVS2-SMCC 的参考软件 RMD1.0^①,对每个序列有损编码配置(4 个 QP)分别进行统计,再求所有序列的平均值.

首先,统计每个测试序列当前帧相对于前一帧的内容的变化情况.图 2 给出了当前帧与前一帧相同位置 8×8 像素的绝对值差值之和为零块(简记为 zeroSAD8×8 块)的数目占整个图像 8×8 块数目的百分比情况;图 2 也给出了 zeroSAD8×8 块中选择 64×64 划分 CU 的百分比情况(图 2 中折线表示)和 zeroSAD8×8 块中选择 F_Skip 编码模式的百分比情况(图 2 中灰色条表示).

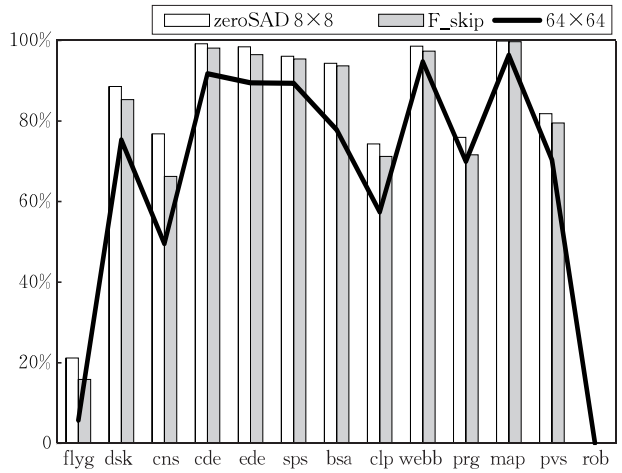


图 2 zeroSAD8×8 和 F_Skip 子模式、zeroSAD8×8 的块中选择 64×64 划分的百分比

从图 2 中可知,除了 flyg 和 rob 序列之外,其余序列的 zeroSAD8×8 的比例都在 75% 以上,表明大部分帧间编码的图像中至少有 3/4 的区域是完全静止的.除此之外,可以看到选中 64×64 编码块的百分比和选中 F_Skip 编码模式的百分比非常接近 zeroSAD8×8 块的百分比,进一步得出 zeroSAD8×8 的块主要是选中 F_Skip 编码模式的 64×64 划分 CU.

图 3 给出了 F 图像各序列模式分析结果. AVS2-SMCC 中,编码模式主要包括帧内模式(intra 4:2:0)、帧间编码模式(inter 4:2:0)和 USP 编码模式.对于 F 图像,帧间编码模式主要包括 F_Skip 子模式、F_Direct 子模式(F_Skip 和 F_Direct 子模式不需要编码运动矢量和参考帧,它们的运动矢量和参考帧都是直接从时域或者空域导出. F_Skip 子模式和 F_Direct 子模式的主要区别在于是否编码残差)和其余的帧间子模式.从图 3 中可知,F 帧的大部分 CU 都是选择帧间 4:2:0 模式中 F_Skip 子模式进行编码. F_Skip 子模式编码只需要编码少量的模式信息,无需编码运动矢量和参考帧,就可提高编码效率.其中,F_Skip 子模式中的 F_Skip_Temporal 子类型编码的比特数最少.根据以上分析,可以得出:

- (1)屏幕内容的帧间编码图像中完全静止的区域,选择 F_Skip 子模式可进一步提高编码效率.
- (2)大部分屏幕内容的帧间编码图像中存在很多完全静止的区域,可通过提前终止算法进一步提高编码速度.

① ftp://47.93.196.121/Public/codec/video_codec/AVS2_P2/MCC

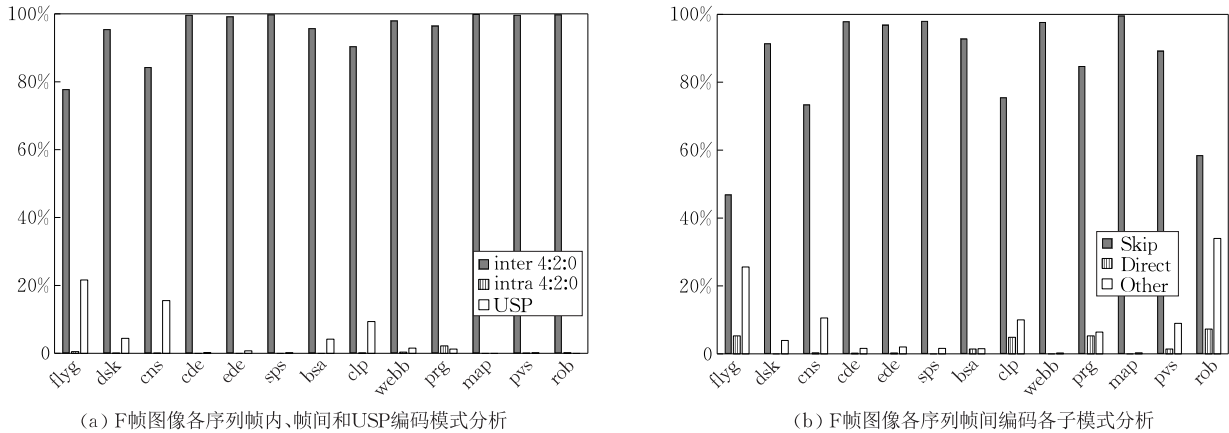


图 3 F 图像各序列模式分析

综上,本文在 AVS2-SMCC 扩展版参考软件 RMD1.0 基础上的帧间优化工作正是在这样的背景下提出的。

3 帧间编码算法的优化

3.1 支持 4:4:4 格式帧间编码的 AVS2-SMCC 算法

图 4 给出了支持 4:4:4 格式帧间编码框架的 AVS2-SMCC 编解码器(灰色部分是新加的模块)。如图 4 所示,支持 4:4:4 格式帧间编码框架主要通过以下方式实现:

(1)增加了 4:4:4 格式的帧间色度参考帧缓冲

区,用于保存 4:4:4 格式的色度重建像素;

(2)进行色度帧间 4:4:4 预测(残差由原始 4:4:4 格式的色度像素和 4:4:4 格式的色度重建像素之差计算得到),预测残差经过以下采样之后进入传统 AVS2 的帧间残差编码模块;

(3)色度帧间 4:4:4 补偿时,把反量化之后的残差上采样之后与色度帧间 4:4:4 预测值相加得到最终帧间编码的 4:4:4 格式的色度重建像素;

(4)将通用串预测算法的第一受限串的参考位置扩展至参考图像队列中的第一帧,提出了一种新的帧间第一受限串模式(见 3.4 节)。

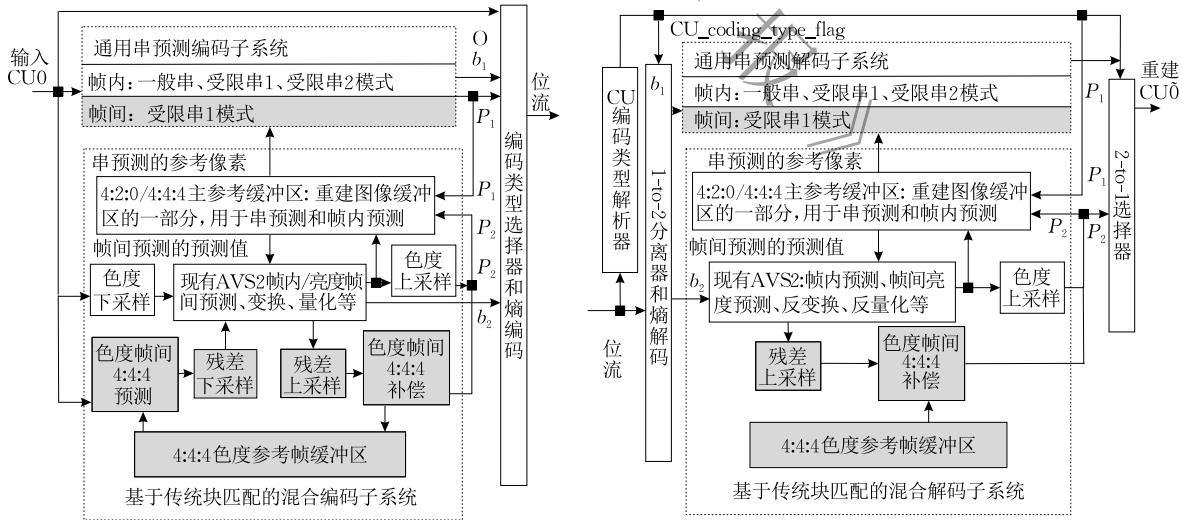


图 4 支持 4:4:4 格式帧间编码的 AVS2-SMCC 框架

本文采用的色度下采样算法是 4:2:0 格式的色度分量取 4:4:4 格式中四个像素的色度分量的平均值(四舍五入),而色度上采样算法则是重复色度分量四次.在此编解码框架下,除了编码失真,依然存在帧内、帧间模式的色度上下采样失真.与文献

[18]类似地,不同模式采用不同的色度失真调节因子:LDF 配置下帧内预测模式的色度调节因子设置为 0.5,通用串预测模式的色度失真权重因子设置为 0.7,帧间模式的色度因子设置方案见 3.2 节.与文献[18]不同的是,为了进一步提高色度分量的图

像质量,对所有各种色度格式的编码模式的编码单元,都统一采用 4:4:4 格式的环内滤波.

支持 4:4:4 格式帧间编码框架的 AVS2-SMCC 算法,大幅提高了帧间图像色度分量的编码质量,从而提高了参考图像队列中各个参考帧的图像质量,使得整体的 LDF 编码配置的编码效率大幅提升.

3.2 基于内容自适应的不同帧间子模式色度失真权重因子设置算法

设 m 表示不同编码模式, $S(m)$ 表示编码模式的集合. m^* 表示当前 CU 选中的模式. $D(m)$ 与 $R(m)$ 分别表示对当前 CU 用模式 m 编码获得的失真与比特数. AVS2 中采用基于率失真值 (Rate Distortion Cost, RDCost) 模式选择机制自适应选择最佳编码模式. 具体而言,即遍历 $S(m)$ 中所有 m ,对当前 CU 用模式 m 编码获得 $RDcost$ 值,从中选择最小者作

为 m^* ,见式(1).

$$m^* = \arg \min_{\forall m \in S(m)} D(m) + \lambda \times R(m) \tag{1}$$

其中, λ 表示失真与比特数的权重因子.

在 2.2 节的分析中,大部分屏幕内容的帧间编码图像中存在很多完全静止的区域,对这些完全静止的区域从理论上应该选择 F_Skip 子模式,从而进一步提高编码效率.但是在实际有损编码中,由于重建像素会存在一定的误差,采用基于 RDCost 模式选择机制会导致这些区域不一定选中 F_Skip 子模式.因此,本文充分利用屏幕图像的帧间全局和局部特性,提出了内容自适应的不同帧间子模式色度失真权重因子算法.图 5 给出了 webb 序列前 9 帧 F 图像分类的情况.第 1 个 F 图像(第 2 帧)至第 7 个 F 图像(第 8 帧)都是完全匹配图像,事实,前面 8 帧都是完全一样的.第 9 帧为场景切换图像.



图 5 webb 序列前 9 帧 F 图像分类结果

定义 1. 完全匹配、基本匹配和场景切换图像.
根据 zeroSAD8×8 块的比例,将帧间编码的 F 图像分为完全匹配图像、基本匹配图像、场景切换图像三类,定义如下:

- 完全匹配图像, zeroSAD8×8 块的比例 ≥ 0.99
- 基本匹配图像, 0.99 > zeroSAD8×8 块的比例 > 0.5.
- 场景切换图像, zeroSAD8×8 块的比例 ≤ 0.5

定义 2. 局部静止 LCU 和局部运动 LCU.
根据当前编码 LCU 和相邻可用 LCU(左边、右边和上边 LCU)的像素值与前一帧相同位置 LCU 和相邻可用 LCU 的像素值是否完全相同,将当前编码 LCU 分为局部静止 LCU 和局部运动 LCU.需用一个标志位来记载当前 LCU 的类别.如果像素值完全相同,当前 LCU 标识为 1(表示局部静止 LCU),否则标识为 0(表示局部运动 LCU).该标识仅用于编码器端辅助选择出更合适的编码模式,无需写入码流.

定义 3. 不同帧间子模式色度失真权重因子 $\omega\lambda_{c1}, \omega\lambda_{c2}, \omega\lambda_{c3}$.

根据 AVS2 中不同帧间编码子模式消耗编码比特数从小到多的顺序,本文将帧间编码子模式分为 F_Skip_Temporal 模式,其它 F_Skip 子模式和其它帧间子模式三类. $\omega\lambda_{c1}, \omega\lambda_{c2}, \omega\lambda_{c3}$ 分别定义为 F_Skip_Temporal 模式,其它 F_Skip 子模式和其它帧间子模式对应的色度失真权重因子.对不同 LCU 的屏幕内容进行分类,我们对不同类别的 LCU(或者 CU)通过设置不同的帧间子模式色度权重因子,使编码比特数更少的 F_Skip 子模式更容易被选中,从而进一步提高编码效率.本文的权重因子通过实验尝试法设定.实验结果表明,内容自适应的不同模式色度失真权重因子算法,对 TGM 类别能带来约 1.4% 的增益.

算法 1 给出了内容自适应的不同编码模式色度失真权重因子算法的伪代码.

算法 1. 内容自适应的不同编码模式色度失真权重因子算法.

```
IF (当前 LCU 为局部静止 LCU)
BEGIN
    IF (当前帧是完全匹配图像)
```

```

 $w\lambda_{c1}=0.2; w\lambda_{c2}=0.4; w\lambda_{c3}=0.5$ 
ELSE IF (当前帧是基本匹配图像)
 $w\lambda_{c1}=0.3; w\lambda_{c2}=0.4; w\lambda_{c3}=0.5$ 
ELSE
 $w\lambda_{c1}=0.4; w\lambda_{c2}=0.45; w\lambda_{c3}=0.5$ 
END IF
ELSE //当前 LCU 为局部运动 LCU
BEGIN
IF (当前帧是完全匹配图像)
 $w\lambda_{c1}=0.45; w\lambda_{c2}=0.45; w\lambda_{c3}=0.5$ 
ELSE
 $w\lambda_{c1}=0.5; w\lambda_{c2}=0.5; w\lambda_{c3}=0.5$ 
END ELSE

```

3.3 帧间子模式预编码和 LCU 划分的提前终止算法

在帧间子模式预编码和 LCU 划分之前, 本文算法根据当前编码块内容特性, 跳过或不跳过帧间子模式和/或 LCU 划分, 进一步提高编码速度。

考虑到 AVS2 的四叉树递归划分的方式是从 8×8 像素递归至 64×64 像素的编码单元, 若当前 LCU 为局部静止 LCU, 对于完全匹配图像, 大部分 CU 的编码模式选择 64×64 划分的 F_Skip 子模式编码, 因此可以跳过帧间子模式和其它编码模式的预编码以及 8×8 、 16×16 和 32×32 大小 CU 的递归划分; 对于非完全匹配图像, 跳过帧间子模式和其它编码模式的预编码, 能大幅减少编码时间。

算法 2 给出了帧间子模式预编码和 LCU 划分的提前终止算法。该提前终止算法能在 TGM 和 MC 类别的测试序列的平均编码效率提升 0.2% 的情况下, 减少 50% 左右的编码时间。

算法 2. 帧间子模式预编码和 LCU 划分的提前终止算法。

```

IF (当前 LCU 为局部静止 LCU)
BEGIN
IF (当前帧是完全匹配图像)
跳过 LCU 的递归划分;
跳过除 F_Skip 子模式之外的模式预编码;
ELSE
跳过除 F_Skip 子模式之外的模式预编码;
END IF
ELSE
当前 LCU 进行正常的递归划分和对每个划分的 CU
进行全部模式的预编码;
END ELSE

```

3.4 帧间第一受限串模式

AVS2 的帧间算法采用分像素精度的运动矢

量。非自然内容区域(如文本和图表等)的运动一般是整像素精度的, 采用分像素精度的运动矢量编码需要消耗更多的比特数。

在 RMD1.0 中, USP 中的第一受限串(Constrained String 1, CS1) 的搜索范围仅局限于当前帧内已重建的区域, 根据串的大小, 采用混合搜索技术, 达到尽快找到最佳参考串的目的; 并将当前串与最佳串的运动矢量(整像素精度)写入码流。

本文在帧内 CS1 串基础上, 提出了一种新的 CS1 模式。该方案对运动矢量的编码方案采用与帧内 CS1 模式相同的方案, 与帧内 CS1 模式不同的是:

(1) 搜索范围。帧间第一受限串的搜索范围扩展至参考图像队列中的第一帧; 对于非 I 帧图像, 引入语法元素 `usp_offset_base_flag` 来标识 CS1 串的参考串是来自于参考图像队列中的第一帧还是当前帧^[26]。

(2) 搜索算法。对于非 I 图像来说, 首先对参考图像队列中的第一帧的原始像素值根据不同分块大小计算出来的哈希值建立哈希链表(将相同的哈希值链接在一起)。哈希值的计算方式采用文献[27]中的方式。其次, 本文用同样的哈希值计算方式计算当前编码块的哈希值, 从哈希链表中搜索得到匹配的串(块)。

帧间 CS1 模式能有 4% 左右的编码效率的提升, 相比传统的帧间编码有两个优势:

(1) 由计算机产生的文本和图表等非自然内容区域产生的运动一般是整像素精度的, 因此帧间 CS1 模式采用整像素精度运动矢量比帧间编码模式采用分像素精度的运动矢量编码消耗的比特数更少。

(2) 帧间 CS1 模式的搜索范围为参考图像队列中的第一帧的整个范围, 搜索的范围比帧间编码的搜索范围更大, 能找到更好的匹配串(块)。

4 实验结果与分析

4.1 实验配置和相关性能评价指标

本文的实验数据采用表 2 中 13 个 AVS2-SMCC 标准测试序列。采用 AVS2-SMCC 通用测试条件中的标准 LDF 配置^[24] 和 HEVC-SCC 的标准 LB 配置^[26] (LDF、LB 配置统一称为 LD 配置)。HEVC-SCC 的量化参数为 22、27、32 和 37; AVS2-SMCC 的量化参数为 25、32、38 和 45。编码帧数为 30 帧。本文中的算法及其比较的基准算法包括以下 5 种:

- (1) RD17.0Pseudo444,简称 RDP444:在 AVS2-P2 基准档次的参考软件 RD17.0 基础上,增加了对输入原始 4:4:4 数据的色度下采样和对输出 4:2:0 重构图像的色度上采样,实现对 YUV4:4:4 序列的编码和解码。
- (2) USP: AVS2-SMCC 的参考软件 RMD1.0^[26]。
- (3) USP_Inter: 在 USP 的基础上,实现本文的帧间编码优化算法。
- (4) SCM: HEVC-SCC 的参考软件 HM16.18-SCM8.7^①。
- (5) SCM_SCCOff: SCM 关闭 Intra Block Copy、Palette、Adaptive Color Transform 等 SCC 扩展工具。

实验结果从编码效率和复杂度两方面来衡量算法性能。所有序列的总体编码效率采用 Y、U、V 分量的平均 BD-rate^[28-29] (分别记为 Y-BDrate、U-BDrate、V-BDrate)来衡量。BD-rate 是 HEVC 核心实验中比较两种算法有损编码模式编码效率优劣的客观评价方法:根据比较的两种算法 4 个 QP 的比特率(bit-rate,单位为 kb/s)与峰值信噪比(Peak

Signal to Noise Ratio, PSNR)数据形成的率失真(Rate Distortion, RD)曲线,对其进行曲线拟合,计算两种算法在相同 PSNR 下比特率的变化情况,最后取平均值得到两种算法编码效率的比较结果。BD-rate 为负值表示相同 PSNR 下比特率降低,编码效率提高。所有序列的总体编码效率同时也采用基于 CPSNR(一种综合的 PSNR 的衡量方法^[30]),该衡量方法已经被 AVS2 标准采纳)的三分量综合的平均 BD-rate 降低率(用 C(CPSNR)表示)来衡量;单个序列的编码效率采用 RD 曲线图来衡量。算法复杂度的比较分别用国际标准制定工作中通用的编、解码运行时间的比值来衡量。

实验平台为双路 CPU(Intel Xeon E5-2680 v2)@2.3 GHz 20 核、内存为 80 GB 的服务器。所有的算法都在相同的环境和同样的命令行窗口数目情况下执行。

图 6 给出了典型屏幕图像 cde 和 pvs 的第一帧图像。其中,包含中文 Word 和 PowerPoint 的文档编辑场景的 cde 序列是 TGM 类别代表序列;包含视频播放与 Excel 使用的场景的 pvs 序列是 MC 代表序列。

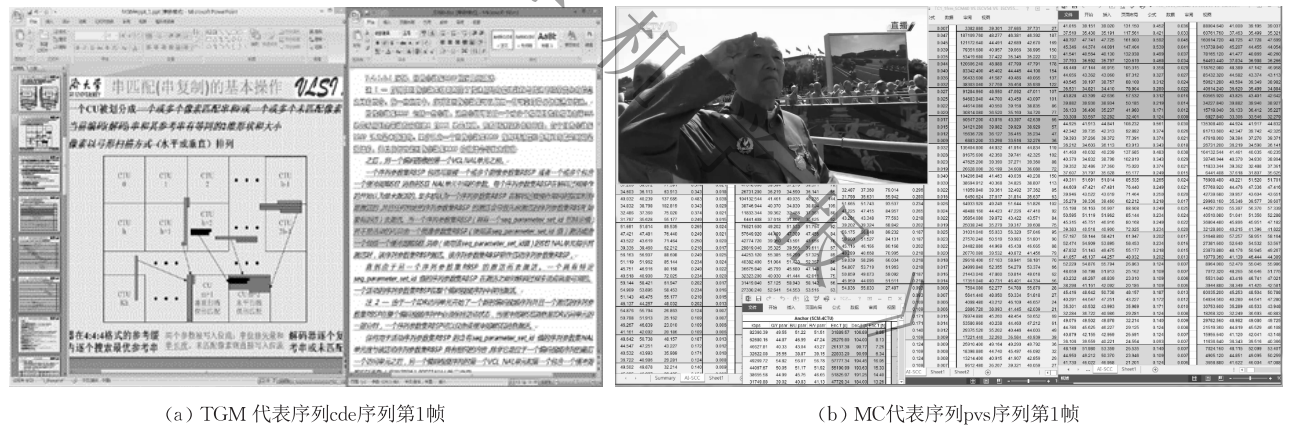


图 6 典型的屏幕混合内容序列

4.2 实验结果与分析

表 3 给出了 USP 和 USP_Inter、SCM 和 USP_Inter 两组算法总体编码效率的比较结果。

表 3 USP 与 USP_Inter 算法、SCM 与 USP_Inter 算法 LD 配置编码效率两组比较结果 (单位: %)					
对比算法	测试序列类别	LD 配置			
		Y-BDrate	U-BDrate	V-BDrate	C(CPSNR)
USP VS. USP_Inter	TGM	-14.35	-90.73	-88.85	-81.79
	MC	-6.28	-99.99	-98.92	-18.26
	G	5.20	-77.34	-70.24	-14.36
SCM VS. USP_Inter	TGM	-14.79	-7.03	-7.69	-13.56
	MC	-3.53	11.23	8.57	-0.81
	G	-10.02	162.19	153.31	3.69

表 4 给出了两组实验 LD 配置的复杂度比较。第一组是 USP_Inter 与 USP 两者复杂度的比较;第二组是 USP_Inter 算法以 RDP444 为基础的相对复杂度与 SCM 算法以 SCM_SCCOff 为基础的相对复杂度进行比较。对于 USP_Inter 算法(在 AVS2 的非 SCC 参考平台 RDP444 上添加 USP 和 USP_inter 等 SCC 扩展工具)与 SCM 算法(在 HEVC 的非 SCC 参考平台 SCM_SCCOff 上添加 Intra Block Copy、Palette、Adaptive Color Transform 等 SCC 扩展工具)的复杂度比较,由于两者的基础平台 RDP444 和

① https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.18+SCM-8.7/

表 4 LD 配置两组实验的复杂度比较 (单位: %)

	对比算法	编码时间比	解码时间比
第一组实验	USP_Inter/USP	49.14	115.24
	SCM/SCM_SCCoff	88.98	100.82
第二组实验	USP_Inter/RDP444	73.16	251.84

SCM_SCCoff 本来就有复杂度的较大差异(尤其是 LD 配置),直接将 USP_Inter 算法与 SCM 算法进行

比较不能体现两者中 SCC 扩展工具的复杂度的差异.因此,本文实验中将 USP_Inter 算法以 RDP444 为基础的相对复杂度与 SCM 算法以 SCM_SCCoff 为基础的相对复杂度进行比较.

图 7 和图 8 分别给出了 cde 和 pvs 序列的 USP、USP_Inter 和 SCM 算法有损 LD 配置编码下的 Y、U、V 分量的 RD 曲线图.

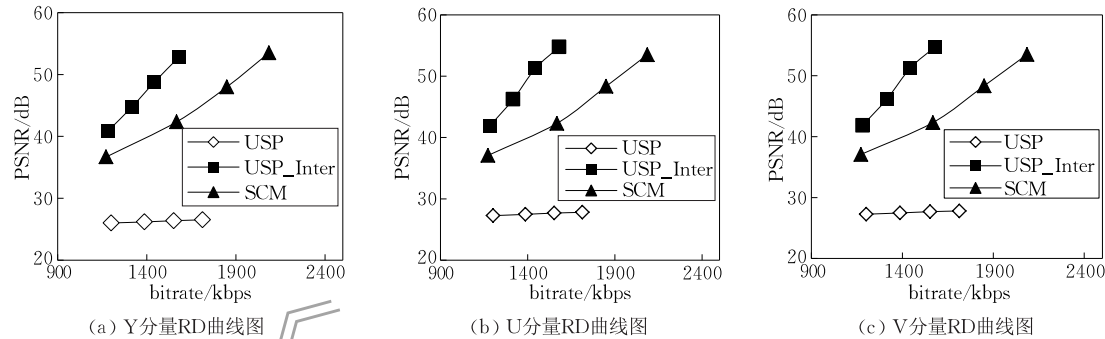


图 7 TGM 代表序列 cde 序列有损 LD 配置下的 RD 曲线图

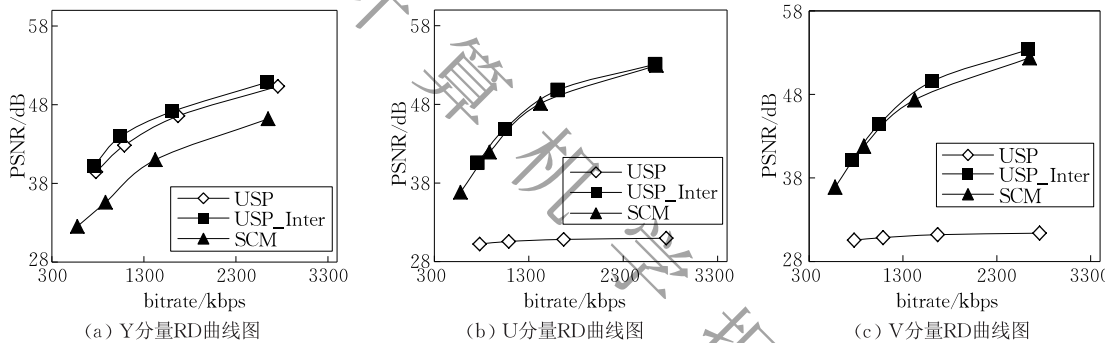


图 8 MC 代表序列 pvs 序列有损 LD 配置下的 RD 曲线图

图 9 给出了 webb 序列在量化参数为 22 时, USP_Inter 算法 LDF 配置的前 30 帧的图像质量.

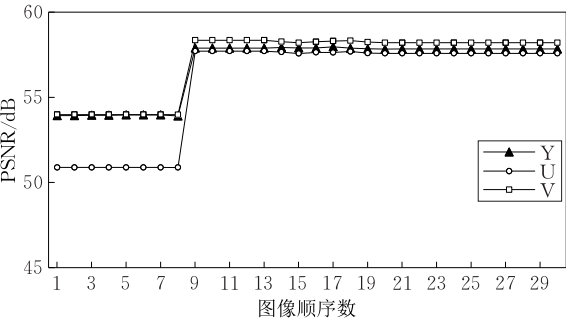


图 9 webb 序列前 30 帧的图像质量分布图

根据以上实验结果,我们可以得出以下结论:

(1) 从编码效率来看,USP_Inter 算法平均编码效率大幅提升,尤其是色度分量.从表 3 中可知,USP_Inter 算法较 USP 算法平均编码效率提升明显;对于 TGM 序列类别,LD 配置下 Y、U、V 分量平均 BD-rate 降低率分别为 14.35%、90.73%、

88.85%,UV 色度分量编码效率提升更显著的主要原因是支持 4:4:4 采样格式帧间编码框架消除了 USP 算法中 4:2:0 采样格式帧间预测模式色度分量的上下采样误差,进一步提高了色度分量的编码和参考质量;USP_Inter 算法较 SCM 算法相比:对于 TGM 序列类别,LD 配置下 Y、U、V 分量平均 BD-rate 降低率分别为 14.79%、7.03%、7.69%;对于 TGM 和 MC 序列类别,基于 CPSNR 的三分量综合的平均 BD-rate 降低率分别为 13.56%和 0.81%.

(2) 在复杂度方面,一方面(见表 4 第一组实验数据),USP_Inter 能大幅降低 USP 算法的编码时间复杂度(从表 4 可知,USP_Inter 相比 USP,能降低编码时间至 49.14%);另一方面(见表 4 第二组实验数据),USP_Inter 和 SCM 分别与 AVS2 非 SCC 版本和 HEVC 非 SCC 版本相比,编码时间分别减少了 26.84%和 11.02%,可见,USP_inter 比 SCM 有更大的相对编码时间减少百分比的优势.由

于 USP_Inter 的环内滤波(采用 4:4:4 的采样格式)做得比 USP(采用 4:2:0 的采样格式)更全面,解码时间增加了 15.24%. USP_Inter 较 RDP444 算法增加的解码复杂度主要原因是 LDF 配置除了第一帧外都是 F 图像,RDP444 中大多数 CU 选中的模式是基于 4:2:0 采样格式的帧间 skip 模式,该模式的解码复杂度很低;而 USP 中选 4:4:4 的 USP 模式需要处理的 UV 分量数目增加四倍,总像素分量数目增加一倍,导致解码时间增加比较多.

(3)从图像的客观质量来看,USP_Inter 算法的编码质量大幅提高,尤其是色度分量.从图 7 和图 8 中的 RD 曲线可以看出,对于 cde 和 pvs 数据,USP_Inter 较 USP 算法的各分量的 PSNR 都有大幅提升,与 SCM 算法相比也有些提升.从图 9 中 webb 序列前 30 帧的图像质量分布图与图 1(a)中的图像质量分布图可以看出,USP_Inter 算法取得的色度分量 PSNR 有了大幅的提高,并且已经接近亮度分量的 PSNR.

(4)提出的 USP_Inter 算法性能的提升程度与屏幕图像的类别相关.从表 3 中可以看出,与 SCM 相比,对于 TGM 类别的三分量综合的编码效率提升非常显著;对于 M 类别的三分量综合的编码效率稍有提升;而对 G 类别的三分量综合的编码效率却有降低,主要原因包括:① USP_Inter 算法是基于 AVS2 标准实现的,采用 4:2:0 采样格式的帧内预测算法,对于色度分量同时存在预测误差和上下采样误差,而 SCM 是基于 HEVC 标准 RExt 扩展版实现的,色度分量只有预测误差(这是 MC 类别和 G 类别 UV 分量平均 BD-rate 较高的原因之一);② SCM 中还包括一些其它技术,比如自适应颜色变换(Adaptive Color Transform,ACT)、分量交叉预测(Cross Component Prediction,CCP)等.这些技术对 G 类数据的编码效率的提升(包括 UV 分量)很有帮助.而 USP_Inter 算法中目前还没有这些技术.

综上,基于 AVS2 标准的 USP_Inter 算法与国际上最新的 HEVC-SCC 标准的最新参考软件 SCM8.7 相比,对于 TGM 类别和 MC 类别,全帧内配置和 LD 配置都能获得更好的综合编码效率(USP_Inter 算法对 AI 配置的编码性能几乎没有影响,因此本文仅仅给出 LD 配置的实验结果),同时与其它基于 HEVC-SCC 标准的帧间优化算法^[20-21]不具有竞争性、排他性和对比性,而是具有互补和叠加的编码效率提升作用.

5 结论与下一步工作

本文在目前 AVS2-SMCC 扩展版基础上,对 LDF 配置的帧间编码算法做了 5 项优化工作:(1)针对基于 YUV4:2:0 的帧间预测模式对 LDF 配置的编码效率影响比较大的问题,实现了支持 4:4:4 格式的帧间编解码框架;(2)考虑到屏幕内容的帧间编码图像中完全静止的区域选择 F_Skip 子模式可进一步提高编码效率,提出了内容自适应的不同帧间子模式色度失真权重因子算法,对完全静止的区域块 F_Skip 子模式采用更小的色度失真权重因子;(3)根据大部分屏幕内容的帧间编码图像中存在很多完全静止的区域特征,提出了帧间子模式预编码和编码树单元划分的提前终止算法;(4)根据屏幕内容中很多运动是整像素精度等特征,提出了帧间 CSI 模式;(5)对所有各种色度格式的编码模式的编码单元,都统一采用 4:4:4 格式的环内滤波.其中优化工作(1)、(2)、(4)、(5)进一步提高了编码效率,而优化工作(3)进一步减少了编码时间复杂度.因此,优化后的 AVS2-SMCC 算法能大幅提高屏幕混合内容的帧间编码效率,特别是对于色度分量有大幅的提升,并且减少了编码时间复杂度.优化后的 AVS2-SMCC 算法与 HEVC-SCC 相比,TGM 序列类别和 MC 序列的编码效率在低延迟配置下体现出较大的优势.

本文下一步的工作包括在下一代 AVS 视频编码标准^[31]中继续探索屏幕混合内容视频中帧间编码的运动矢量、残差、参考帧等特点,根据其特点进行编码性能的优化(包括编解码复杂度的优化(比如,在北京大学的开源的 AVS2 高清实时编码器 xavs2^①中实现低复杂度的 USP_Inter 算法))以及对现有通用串预测中的各种受限串进行扩展.同时,我们应探索超低复杂度兼高效的串匹配算法及其应用^[32-34].

参 考 文 献

- [1] Lin Tao, Zhou Kai-Lun, Wang Shu-Hui. Cloudlet-screen computing: A client-server architecture with top graphics performance. International Journal of Ad Hoc and Ubiquitous Computing, 2013, 13(2): 96-108
- [2] Xu Ji-Zheng, Joshi R, Cohen R A. Overview of the emerging HEVC screen content coding extension. IEEE Transactions on Circuits and Systems for Video Technology, 2016, 26(1):

① <http://www.avs.org.cn/news/readnews.asp?id=168>

- 50-62
- [3] Peng W, Walls F, Cohen R, et al. Overview of screen content video coding technologies, standards, and beyond. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 2016, 6(4): 339-408
- [4] Liu Dan, Chen Gui-Sheng, Song Chuan-Ming. Research advances in screen content coding methods. *Journal of Computer Research and Development*, 2017, 54(9): 2059-2076(in Chinese)
(刘丹, 陈规胜, 宋传鸣. 屏幕内容编码方法研究进展. *计算机研究与发展*, 2017, 54(9): 2059-2076)
- [5] Shen Yan-Fei, Li Jin-Tao, Zhu Zhen-Min, et al. High efficiency video coding. *Chinese Journal of Computers*, 2013, 36(11): 2340-2355(in Chinese)
(沈燕飞, 李锦涛, 朱珍民等. 高效视频编码. *计算机学报*, 2013, 36(11): 2340-2355)
- [6] Gao Wen, Ma Si-Wei. *Advanced Video Coding Systems*. Berlin, Germany: Springer, 2015
- [7] Ma Si-Wei, Luo Fa-Lei, Huang Tie-Jun. Kernel technologies and applications of AVS2 video coding standard. *Telecommunications Science*, 2017, 8: 2-15(in Chinese)
(马思伟, 罗法蕾, 黄铁军. AVS2 视频编码标准技术特色及应用. *电信科学*, 2017, 8: 2-15)
- [8] AVS N14175. AVS Document. Dalian, China, 2015(in Chinese)
(AVS N14175. 数字音视频编解码技术标准工作组文件. 大连, 中国, 2015)
- [9] Lin Tao, Zhang Pei-Jun, Wang Shu-Hui, et al. Mixed chroma sampling-rate high efficiency video coding for full-chroma screen content. *IEEE Transactions on Circuits and Systems for Video Technology*, 2013, 23(1): 173-185
- [10] Guo Li-Wei, Pu Wei, Zou Feng, et al. Color palette for screen content coding // *Proceedings of the IEEE International Conference on Image Process*. Pairs, France, 2013: 5556-5560
- [11] Ma Zhan, Wang Wei, Xu Meng, et al. Advanced screen content coding using color table and index map. *IEEE Transactions on Image Processing*, 2014, 23(10): 4399-4412
- [12] Xu X, Liu S, Chuang S, et al. Intra block copy in HEVC screen content coding extensions. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 2016, 6(4): 409-418
- [13] Chen Chun-Chi, Peng Wen-Hsiao. Intra line copy for HEVC screen content coding. *IEEE Transactions on Circuits and Systems for Video Technology*, 2017, 27(7): 1568-1579
- [14] Zhao Li-Ping, Lin Tao, Zhou Kai-Lun, et al. Pseudo 2D string matching technique for high efficiency screen content coding. *IEEE Transactions on Multimedia*, 2016, 18(3): 339-350
- [15] Zhao Li-Ping, Zhou Kai-Lun, Guo Jing, et al. A universal string matching approach to screen content coding. 2018, 20(4): 796-809
- [16] Zhao Li-Ping, Lin Tao, Zhou Kai-Lun. An efficient ISC offset parameter coding algorithm in screen content coding. *Chinese Journal of Computers*, 2017, 40(5): 1218-1228(in Chinese)
(赵利平, 林涛, 周开伦. 屏幕图像压缩中串复制位移参数的高效编码算法. *计算机学报*, 2017, 40(5): 1218-1228)
- [17] Zhao Li-Ping, Zhou Kai-Lun, Guo Jing, et al. Pixel string matching for full-chroma screen and mixed content coding in AVS2. *Chinese Journal of Computers*, 2018, 41(11): 2482-2495(in Chinese)
(赵利平, 周开伦, 郭靖等. 融合像素串匹配的 AVS2 全色度屏幕与混合内容视频编码算法. *计算机学报*, 2018, 41(11): 2482-2495)
- [18] Zhao Li-Ping, Zhou Kai-Lun, Lin Tao, et al. A universal string prediction approach and its application in AVS2 mixed content coding. To be published in *Chinese Journal of Computers*(in Chinese)
(赵利平, 周开伦, 林涛等. 通用串预测算法及在 AVS2 屏幕与混合内容视频编码中的应用. *计算机学报*, 已录用)
- [19] Xiao Wei, Shi Guang-Ming, Li Bin, et al. Fast hash-based inter block matching for screen content coding. *IEEE Transactions on Circuits and Systems for Video Technology*, 2018, 28(5): 1169-1182
- [20] Xiao Wei, Li Bin, Xu Ji-Zheng, et al. Weighted rate-distortion optimization for screen content coding. *IEEE Transactions on Circuits and Systems for Video Technology*, 2018, 28(2): 499-512
- [21] Li Jia-Hao, Li Bin, Xu Ji-Zheng, et al. Diversity-based reference picture management for low delay screen content coding. *IEEE Transactions on Circuits and Systems for Video Technology*, 2018, 28(6): 1369-1378
- [22] Lin Tao, Guo Jing, Zhao Li-Ping, et al. The improvements on universal string prediction for screen and mixed content coding in AVS2-P2. AVS Document M4081, Haikou, China, 2016(in Chinese)
(林涛, 郭靖, 赵利平等. 屏幕与混合内容编码中通用串预测(USP)方法的几点改进. 数字音视频编解码技术标准工作组文件 M4081, 海口, 中国, 2016)
- [23] Zhou Kai-Lun, Zhao Li-Ping, Lin Tao, et al. The improvements of CS1 mode in universal string prediction for screen and mixed content coding. AVS Document M4153, Beijing, China, 2017(in Chinese)
(周开伦, 赵利平, 林涛等. 屏幕与混合内容编码中通用串预测(USP)方法中 CS1 模式的几点改进. 数字音视频编解码技术标准工作组文件 M4153, 北京, 中国, 2017)
- [24] AVS. Common test condition for screen and mixed content coding in AVS2-P2. AVS Document N2387, Dalian, China, 2016(in Chinese)
(AVS N2387. AVS2-P2 屏幕与混合内容视频编码(S&MCVC)通用测试条件. 大连, 中国, 2016)
- [25] Yu H P, Cohen R, Rapaka K, et al. Common test conditions for screen content coding. Joint Collaborative Team on Video Coding Document X1015, Geneva, 2016
- [26] AVS. AVS2-P2-SMC-WD3.0. AVS Document N2388. Haikou, China, 2016(in Chinese)
(AVS N2388. 通用串预测方式的语法、语义、解析过程、解码过程. 海口, 中国, 2016)

- [27] Li B, Xu J. "Hash-based motion search," document JCTVC-Q0245, JCT-VC, Valencia, Spain, Mar., 2014
- [28] Bjøntegaard G. Calculation of average PSNR differences between RD-curves. ITU-T SG16 Q.6 Document, VCEG-M33, 2001
- [29] Bjøntegaard G. Improvements of the BD-PSNR model. ITU-T SG16 Q.6 Document, VCEG-AI11, 2008
- [30] Wu Cheng-Jia, Chen Da-Peng, Chen Fang, et al. A method for calculating PSNR. AVS Document M3830, Hangzhou, China, 2016(in Chinese)
(吴成家, 陈大鹏, 陈芳等. 一种计算综合 PSNR 的方法. AVS 工作组文件 M3830, 杭州, 中国, 2016)
- [31] AVS N2401. The technical requirements V9.0 for the next generation of AVS video coding standards. AVS Document, Beijing, China, 2017(in Chinese)
(AVS N2401. 下一代 AVS 视频编码标准的技术需求 V9.0, 北京, 中国, 2017)
- [32] Lin Tao, Cai Wen-Ting, Chen Xian-Yi, et al. Lossless compression algorithm based on string matching with high

performance and low complexity for screen content coding. Journal of Electronics & Information Technology, 2017, 39(2): 351-359(in Chinese)

(林涛, 蔡文婷, 陈先义等. 一种高性能低复杂度的基于串匹配的屏幕图像无损压缩算法. 电子与信息学报, 2017, 39(2): 351-359)

- [33] Zhao Li-Ping, Lin Tao, Zhou Kai-Lun, et al. A byte-size multi-variable-length-code based string matching algorithm for Alpha image coding. Telecommunications Science, 2018, 34(11): 96-104(in Chinese)

(赵利平, 林涛, 周开伦等. 一种基于字节型多变长码的串匹配的 Alpha 图像编码算法. 电信科学, 2018, 34(11): 96-104)

- [34] Lin Tao, Zhang Dong-Yu, Zhao Li-Ping. An improved entropy coding algorithm in string matching based alpha image coding. To be published in Journal of Jishou University (Natural Science Edition)(in Chinese)

(林涛, 章栋宇, 赵利平. 一种 Alpha 图像串匹配编码中熵编码的改进算法. 吉首大学学报(自然科学版), 已录用)



ZHAO Li-Ping, Ph. D., associate professor. Her research interests include video coding, screen content coding.

LIN Tao, Ph. D., professor, Ph. D. supervisor. His research interests include video coding, screen content coding.

GUO Jing, M. S. His research interest is screen content coding.

ZHOU Kai-Lun, Ph. D., lecturer. His research interests include video coding and VLSI architecture, design and verification.

Background

Screen content, such as the computer and mobile display content, has recently emerged as a popular video type due to the fast rising demands for transporting or storing screen display in the form of video. This is partly driven by rapid advances in mobile, cloud and display technologies, which enable a variety of screen applications over various networks/links, such as wireless displays, second screen, desktop screen sharing and collaboration, cloud computing and gaming, PC-over-IP, etc.

In recent years, SCC has attracted increasing attention of researchers from both academia and industry. In response to the demand on SCC, three major HEVC-SCC tools developed are: Intra Block Copy (IBC), palette coding (PLT) and String Matching (SM). In China, an Screen and Mixed Content extension is also developing under the second generation of the Audio Video Coding Standard from August 2015.

From video coding point of view, the most significant and relevant characteristic of screen content is its different characteristics with the nature videos, not only the inner frame characteristics but also the inter frame characteristics. In our previous work, a universal string prediction (USP)

approach with three modes: general string (GS) mode, constrained string 1 (CS1) mode, and constrained string 2 (CS2) mode are proposed. Experimental results were shown that USP is very efficient for all intra configuration coding, but is not efficient enough for low delay configuration coding. In this paper, an optimized inter coding algorithm which consists of a 4:4:4 format inter-coding algorithm, a content adaptive weighted chroma distortion algorithm for different inter-sub-coding modes and a early terminate algorithm for inter-sub-coding modes pre-coding and coding tree unit partition, and a newly inter-constrained string 1 mode (the search area of CS1 mode extends to the previous frame in the decoding sequences) is proposed to improve the performance of the low delay configuration coding for Screen and Mixed Content.

This work is partially supported by the Natural Science Foundation of Zhejiang Province (No. LY19F020015), the National Natural Science Foundation of China (Nos. 61871289 and 61601200), the Public Service Technology Application Research Project of Shaoxing City (No. 2018C10015), and the Natural Science Foundation of Shanghai (Nos. 18ZR1440600 and 19ZR1461100).