

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

赵利平^{1),2),3)} 周开伦²⁾ 林 涛²⁾ 郭 靖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 针对屏幕与混合内容视频具有多样性的特点,在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视频编码(the second-generation Audio Video Coding Standard Screen and Mixed Content Coding, AVS2-SMCC)扩展版中,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(Universal String Prediction, USP)及其关键技术,其主要部分已经被 AVS2-SMCC 扩展版工作草案 3.0 稿采纳. 三种受限串模式统一采用偏移串、坐标串和不可预测像素串三种串类型之一或其组合来实现. 当编码单元采用 USP 算法进行编码时,用三种受限串模式对其分别进行预编码,自适应采用三种受限串模式中取得率失真值最小的受限串模式对其进行编码. 实验结果表明,对于 AVS2-SMCC 通用测试序列中移动的文字和图形类别,提出的 USP 算法与国际上最新的(High Efficiency Video Coding, HEVC)标准 SMCC 扩展版(High Efficiency Video Coding-Screen Content Coding, HEVC-SCC)相比,在编解码复杂度增加较少的情况下,对于有损全帧内配置 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%,基于 CPSNR 的三分量综合的平均 BD-rate 降低率为 22.1%,编码效率已经大幅超过 HEVC-SCC.

关键词 高效视频编码;数字音视频编解码技术标准;屏幕混合内容;串匹配;通用串预测;受限串模式

中图法分类号 TP391

DOI号 10.11897/SP.J.1016.2019.02100

A Universal String Prediction Approach and Its Application in AVS2 Mixed Content Coding

ZHAO Li-Ping^{1),2),3)} ZHOU Kai-Lun²⁾ LIN Tao²⁾ GUO Jing²⁾

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 Screen content coding (SCC) is a key technology for many popular applications in traditional and emerging market sectors, such as cloud computing, remote desktop, screen sharing, etc. Typical computer screen content has very different characteristics from traditional camera-captured content. Therefore, in recent years, SCC has become 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), a joint effort of the ISO/IEC MPEG and ITU-T SG16 WP3 Q6. The other is the second-generation AVS (AVS2), a joint effort of the Audio Video Coding Standard (AVS) Workgroup of China and the IEEE. Repeated identical patterns (i. e., matching patterns) are often

收稿日期:2017-07-26;在线出版日期:2019-04-03. 本课题得到浙江省自然科学基金(LY19F020015)、绍兴市公益性技术应用研究计划项目(2018C10015)、国家自然科学基金(61871289,61601200)及上海市自然科学基金(18ZR1440600,19ZR1461100)资助. 赵利平(通信作者),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. 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. 林 涛(通信作者),博士,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,硕士研究生,主要研究方向为屏幕图像编码.

observed on the same picture of screen content. Two major SCC tools in HEVC SCC developed in recent years to exploit those repeated identical patterns with a variety of sizes and/or shapes are: Intra Block Copy (IBC), palette coding (PLT). IBC is efficient for coding repeated identical patterns with a few fixed sizes and shapes. PLT can code repeated identical patterns of two simple cases called LEFT run and ABOVE run inside a CU using two modes (LEFT and ABOVE) of intra-CU. Beyond the HEVC-SCC extension, another tool String Matching(SM) can code general matching patterns with a variety of sizes, shapes, and positions efficiently. Each of the three tools plays an indispensable role in coding some types of screen content. AVS2 is the second-generation video coding standard developed by the Audio and Video Coding Standard (AVS) Working Group of China. Compared to HEVC, AVS2 can achieve significant coding efficiency improvement for scene video coding by adopting intelligent coding tools. In recent years, AVS is developing an AVS2 Screen and Mixed Content Coding (SMCC) extension (AVS2-SMCC). In the early stage of the development, a pixel string matching technique was proposed for AVS2 SMCC extension. Although the technique provided significant coding gain for screen content, it was still less efficient than HEVC SCC extensions for some screen and mixed contents. 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 universal string prediction(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, which have been adopted by the AVS2-SMCC extension draft version. 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. Compared with the latest HEVC-SCC extension, experimental results show that, for text and graphics with motion of AVS2-SMCC test sequences, the proposed algorithm achieves the average BD-rate reduction of 23.2%, 18.3% and 19.4% for Y, U and V, respectively in All Intra configuration at about the same degree of the encoding and decoding complexity percentage increase.

Keywords high efficiency video coding; audio video coding standard; screen and mixed content; string matching; universal string prediction; constrained string mode

1 引言

新一代主流云计算与信息处理平台中,直接传输云端产生的屏幕图像的像素数据到客户端是数据传输所需带宽最省、安全性较高的一种系统架构^[1]. 任何直接从各类设备的屏幕图像显示单元捕获的视频或图像,或者作为信息载体显示在显示屏上的内容都称为屏幕图像. 屏幕无处不在,各种各样日益剧增的应用产生的屏幕内容错综复杂,种类繁多. 因此,对屏幕图像进行超高压缩比和极高质量的数据压缩技术,即屏幕图像编码(Screen Content Coding, SCC)技术^[2-3],成为当前云计算、云移动计算、桌面云、远程桌面、智能手机和平板电脑的第二显示、屏

幕分享等新一代云计算与信息处理应用中解决客户端与云服务器端或客户端与客户端之间数据传输瓶颈问题中亟需的技术.

屏幕图像编码算法的研究主要围绕由国际电信联盟、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^[4] (High Efficiency Video Coding) SCC 扩展版和由我国数字音视频编码技术标准(Audio Video Coding Standard, AVS)^[5] 工作组制定 AVS 第二代(简称 AVS2)标准^[6] 的屏幕混合内容视频编码(Screen and Mixed Content Coding, SMCC)扩展版两大标准展开. HEVC-SCC 版的标准制定工作于 2014 年 1 月份启动并发布了标准提案征集公告,到 2016 年 3 月份标准制定完成. AVS2-SMCC 扩展版的标准制定工作自 2015 年 8 月份启

动^[7],尚处于制定与完善阶段.屏幕混合内容视频编码也称屏幕图像编码,SMCC 也称 SCC.

从视频编码角度来看,屏幕图像具有多样性的特点.屏幕图像的多样性主要体现在屏幕图像局部或全局范围中具有各种各样的形状和大小不一的重复样图.针对这一特性,目前主流的去重重复样图冗余的屏幕图像编码算法主要包括帧内块复制或匹配(Intra Block Copy, IBC)算法^[8]、调色板(Palette, PLT)算法^[9]和串匹配(String Matching, SM)^[10-20].现有的屏幕图像编码算法对屏幕图像的编码效率已有相当大的提高,但是与满足屏幕图像广阔的应用市场需求还有相当大的距离,面临着更艰巨的困难与挑战.

(1)一方面,HEVC-SCC 标准中的 IBC 算法和 PLT 算法仅仅使用了最简单的帧内块匹配和完全在当前编码单元(Coding Unit, CU)内部进行的特殊形式的索引串匹配.因此,HEVC-SCC 标准对在 HEVC-SCC 标准制定工作选定的相对简单和易于压缩的 HEVC-SCC 标准测试数据集的编码效率很高,但是对于很多其它类型的典型屏幕图像的编码效率并不高,难以适应屏幕图像的多样性.

(2)另一方面,在 AVS2-SMCC 标准中,文献^[20]提出了一种融合像素串匹配的 AVS2 全色度屏幕与混合内容视频编码算法,具有低复杂度和高编码效率的综合优势,与国际最新 HEVC-SCC 标准相比降低了一定复杂度,但整体编码效率仍然还有很大的提高余地.

为了进一步提高屏幕图像的编码效率,在 HEVC-SCC 标准上实现了串匹配算法的典型代表——仿二维串匹配算法(Pseudo 2D String Matching, P2SM)^[16].P2SM 算法的参考缓冲区由当前帧重建像素的一个预定区域构成的主参考缓冲区和保存着当前或附近编码单元中出现频度较高的像素集合的次参考缓冲区组成.P2SM 算法在主参考缓冲区和次参考缓冲区同时进行串匹配,并采用了一系列串匹配算法的改进方案,从而达到更多更好地找到屏幕图像中局部或全局范围的各种形状与大小不一的重复样图的目的.

屏幕图像的多样性决定了不同内容特性的屏幕图像区域需要采用不同的屏幕图像编码算法进行编解码.主流的 IBC 算法、PLT 算法和 P2SM 算法对不同内容特性的区域都发挥着不可或缺的作用.P2SM 算法对于具有复杂的、不同形状与大小的重复样图的屏幕内容区域具有很好的编码效率,但是

IBC 算法或 PLT 算法对于具有简单的块匹配或特殊匹配方式的重复样图的屏幕内容区域具有更好的编码效率.IBC 算法或 PLT 算法可以看成是 P2SM 算法对搜索范围或匹配参数进行取值范围限定后的串匹配算法的特例.

本文在 AVS2-SMCC 标准中,对多种基于匹配方式的屏幕图像编码算法进行各取所长的整合以及进一步改进,提出了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)及其关键技术. USP 算法是 AVS2 提案^[21-25]中的核心算法,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的情况下,有损全帧内配置 Y 分量、U 分量、V 分量平均 BD-rate^[26-27]降低率分别可达 23.2%、18.3%、19.4%,编码效率已大幅超越 HEVC-SCC.通用串预测算法的核心技术已经被 AVS2-SMCC 工作草案(Working Draft, WD)3.0 稿^[28]采纳.

2 通用串预测算法提出的动机

屏幕图像局部或全局范围中具有各种各样的形状和大小不一的重复样图,如何充分挖掘出这些重复样图的匹配特性并对匹配关系参数进行高效编码是提高屏幕图像编码效率的关键.

图 1 给出了基于匹配方式的屏幕图像编码算法的典型代表 IBC 算法、PLT 算法和 P2SM 算法示意图.从图 1 中可知,各种基于匹配方式的 SCC 算法的主要区别在于参考范围的大小、匹配位移矢量(Offset)、匹配长度(Length)与形状、扫描顺序与方式、匹配方式等方面的不同.具体如下:

(1)图 1(a)给出了 IBC 算法^[8]的图示. IBC 算法是帧间算法在当前帧的扩展,在由当前帧重建区域或者重建区域一部分构成的搜索范围内,采用有限种固定大小的块匹配,与屏幕图像中存在各种固定大小与形状的重复样图的特点相适应. IBC 算法的参考范围为当前帧重建像素的一个预定区域,匹配位置是参考范围内的任意合法位置,匹配长度与形状是固定的有限种,采用块匹配方式.

(2)图 1(b)中给出了 P2SM 算法^[16]的图示. P2SM 算法的参考范围为当前帧重建像素的一个预定区域构成的主参考缓冲区(Primary Reference Buffer, PRB)和保存着当前或附近编码单元中出现

频度较高的像素集合的次参考缓冲区 (Secondary Reference Buffer, SRB), 匹配位置是参考范围内的任意合法位置, 匹配长度与形状是任意的长度与形状, 采用行列自适应弓形扫描方式. PRB 串采用仿二维串匹配 (当前样值串和匹配样值串具有相同长度和完全相同的二维形状, 但不一定是二维矩形形状) 方式; SRB 串采用自我复制串匹配方式.

(3) 图 1(c) 中给出了 PLT 算法^[9]的图示. 对于

一个输入 CU, 首先将 CU 内的像素通过调色板生成算法获得调色板 Palette, 然后对每个调色板中的主颜色用索引表示, 最后将 CU 内的像素映射成主颜色对应的索引表. 对当前 CU 映射后的索引表进行特殊形式 (Copy-left 和 Copy-above) 的索引串匹配. PLT 算法的参考范围为当前编码 CU 内的索引, 匹配位置只有左边位置和上边位置. 其它的同 P2SM 算法.

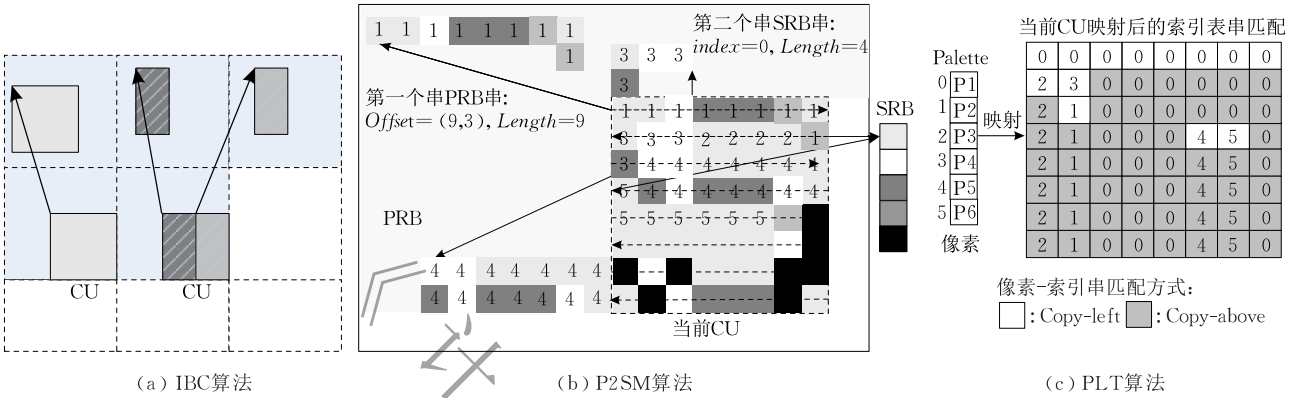


图 1 基于匹配方式的屏幕图像编码算法图示

图 2(a) 给出了 HEVC-SCC 标准对 13 个取名为 FLYG 等序列的编码效率. 这 13 个测试序列分别来自 HEVC-SCC 通用测试条件^[29]和 AVS2-SMCC 通用测试条件^[30]中的移动的文字和图形 (Text and Graphics with Motion, TGM) 类别. 统计

算法采用 HEVC-SCC 参考软件 SCM5.2 版本^①. 对测试序列采用 QP 为 22, 27, 32, 37 进行编码, 获得码流的平均比特数. 原始序列的比特数与码流的平均比特数的比值为压缩率.

模式分布 (序列中选中各种模式进行编码的像素数目占总像素数目的比例) 是一种常用的分析序列的各种模式特点的衡量方式. 图 2(b) 给出了 TGM 序列 P2SM 模式、IBC 模式和 PLT 模式的模式分布情况. 统计算法采用的是 P2SM 算法^[16]. 对测试序列采用 QP 为 22, 27, 32, 37 进行编码, 获得不同屏幕图像序列各种模式所占的平均百分比.

从图 2 中可以得出以下结论:

(1) 根据图 2(a) 可知, 不同屏幕图像序列的压缩率从 32.69 至 178.77 倍不等, 说明 HEVC-SCC 技术对于有些数据的编码效率高, 但对于有些数据的编码效率不高. HEVC-SCC 技术中仅仅支持具有固定形状与大小的块匹配的 IBC 算法和 CU 内进行两种最简单索引串匹配的 PLT 算法, 难以适应屏幕图像的多样性.

(2) 根据图 2(b) 可知, 不同的屏幕图像序列具有不同的模式分布特性. 有些序列选中 P2SM 模式的比例高达 58% 以上; 有些序列选择 IBC 模式的比

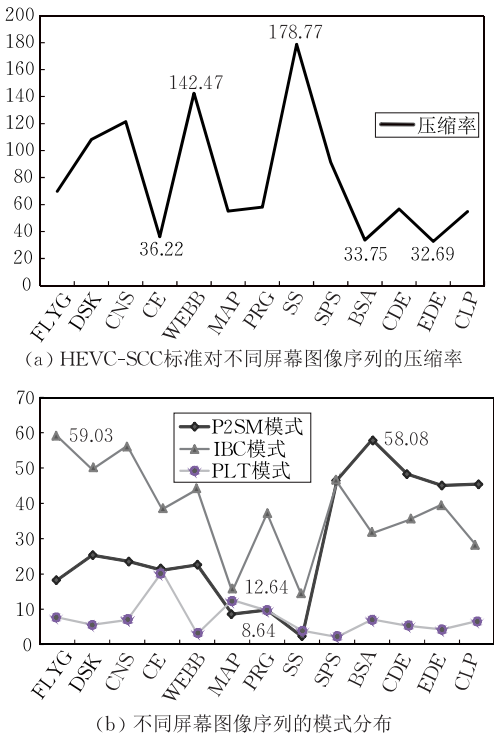


图 2 基于匹配方式的屏幕图像编码算法 (模式) 对不同屏幕图像序列的作用分析

① https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.6+SCM-5.2

例高达 59% 以上;有些序列选中 PLT 模式的比例比 P2SM 模式的比例更高. 因此,每一种模式对于不同内容特性的屏幕图像区域都发挥着不可替代的作用.

事实上,IBC 算法、PLT 算法都是 P2SM 算法的特例. 为了适应屏幕图像的多样性,对不同的屏幕内容采用不同受限串模式的通用串预测算法正是在这样的背景下产生的.

相对于兼具低复杂度和高编码效率综合优势的融合像素串匹配的 AVS2 全色度屏幕与混合内容视频编码算法^[20],本文以在一定的复杂度限制下追求编码效率最高化为目标,提出与实现了基于 SM 的通用串预测算法. 主要创新点如下:

(1) 提出了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. 每种受限串模式统一采用偏移串、坐标串和不可预测像素串三种类型的串之一或其组合来实现.

(2) 提出了三种受限串模式中的一系列关键技术. 主要包括融合多种子模式的第一受限串模式方案、混合搜索技术方案、第一受限串模式和一般串模式中的 Offset 编码直接或预测编码方案等.

3 通用串预测算法

3.1 通用串预测算法的相关定义

本节给出通用串预测算法的相关定义.

定义 1. SM 属性.

SM 属性是 SM 算法中可以用来区分不同受限 SM 模式类别的特性. SM 属性至少包括参考集、匹配关系、匹配尺寸等属性. 参考集是指重建图像的

一个预定的区域或按照预定规则选择的用于参考的若干重建像素. 匹配关系是描述参考集中的像素与当前编解码像素之间的关系,包括匹配形状、匹配方式和扫描方式等. 匹配尺寸是具有同一匹配关系的当前编解码像素的个数,也称匹配长度. 用 $SMP_1, SMP_2, \dots, SMP_n$ 分别表示 SM 属性 1, SM 属性 2, $\dots$, SM 属性 n . 例如 SMP_1 表示匹配长度.

定义 2. SM 属性的取值范围.

用 $RSMP_1, RSMP_2, \dots, RSMP_i$ 分别表示 SM 属性 1 的取值范围, SM 属性 2 的取值范围, $\dots$, SM 属性 n 的取值范围. 例如 $RSMP_1$ 表示匹配长度的取值范围为 $[1, CU$ 的大小].

定义 3. 约束条件 (Constraint Condition, CC).

约束条件指 SM 属性的取值需要满足某种约束条件或约束条件的组合. 一般而言,约束条件是对 SM 属性的取值进行限制,使得 SM 属性的取值被限定在一个预定的 SM 属性取值范围内,用式(1)表示.

$$CC_i: SMP_i \text{ 的取值} \in RSMP_i \quad (1)$$

定义 4. 受限 SM 模式.

受限 SM 模式 CSMMode 是由若干个 SM 属性构成的 SM 算法,每个 SM 属性都需要满足相对应的约束条件,见式(2).

$$CSMMode = \{SMP_1, SMP_2, \dots, SMP_i | SMP_i \text{ 满足 } CC_i\} \quad (2)$$

定义 5. USP 算法框架.

USP 算法框架是由两种或者两种以上的受限 SM 模式组成的通用串预测算法框架,见式(3).

$$USP = \{CSMMode_1, CSMMode_2, \dots, CSMMode_n | n \geq 2\} \quad (3)$$

图 3 给出了 USP 算法框架示意图. 通过对屏幕图像特征的挖掘,发现不同内容特征的屏幕图像区

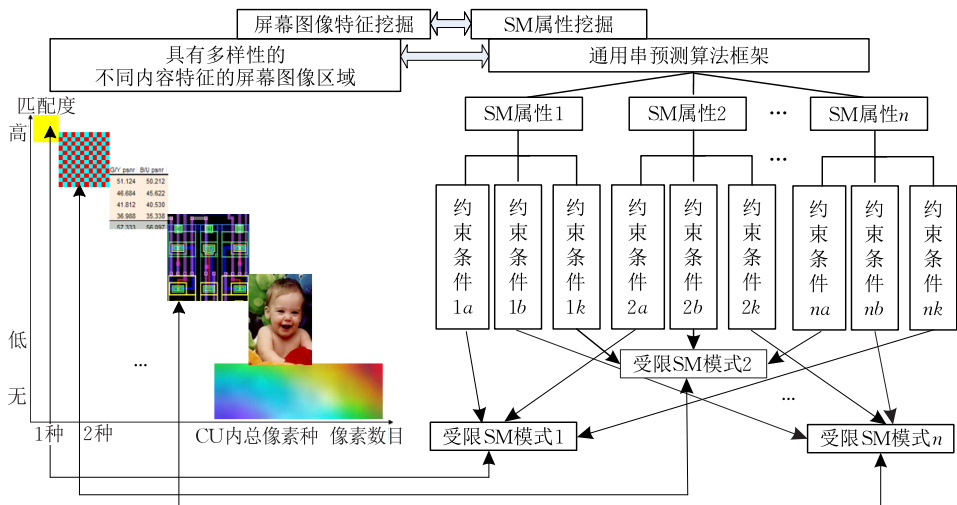


图 3 通用串预测算法框架示意图

域在采用 SM 算法编码时, SM 的属性体现出不同的特点. 将这些特点通过对 SM 属性挖掘, 并对 SM 属性的取值通过约束条件进行取值范围的限制, 从而衍生出各种受限 SM 模式. 各种受限 SM 模式因其 SM 属性的取值进行取值限制, 可以进行串匹配搜索与串匹配参数熵编码等方面的优化, 使得具有不同内容特征的屏幕图像区域自适应地采用不同的受限串模式, 从而达到编码效率与复杂度的最佳平衡.

3.2 受限串模式的串预测编解码操作

图 4 给出了受限串模式的串预测示意图. 采用受限串模式进行编解码单元统一使用 4:4:4 采样格式的 3 种像素串: 偏移串(Offset String, OFS)、坐标串(Coordinate String, CDS)和不可预测像素串(Unpredictable Pixel String, UPS).

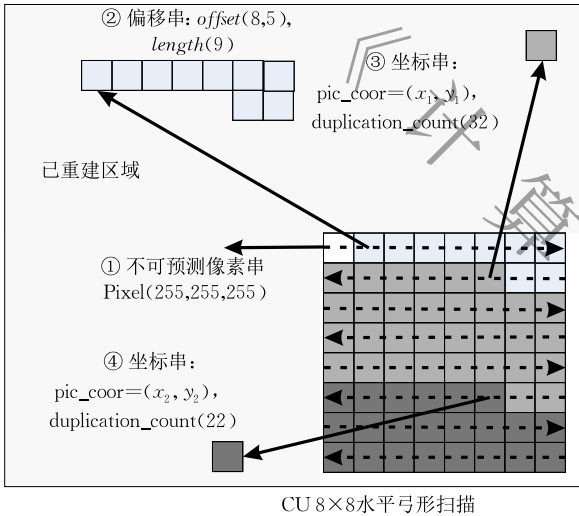


图 4 受限串模式的串预测示意图

如图 4 所示, 当前像素在已重建区域范围内找不到任何偏移串或者坐标串时, 当前像素为不可预测像素串类型. 当前 CU 的第一个像素为不可预测像素串. 不可预测像素串用像素值本身来表示; 当前 CU 的第二个串为偏移串, 偏移串用串位移矢量 **Offset** (当前串的第一个像素所在图像坐标与参考串的第一个像素所在图像坐标的水平偏移量和垂直偏移量称为位移矢量 **Offset** = (**OffsetX**, **OffsetY**)) 和串长度 **Length** 两个参数来表示当前串和参考串的匹配关系; 当前 CU 的第三个和第四个串为坐标串, 坐标串用图像坐标 **pic_coor** = (**x**, **y**) 和重复次数 **duplication_count** 两个参数来表示当前串和参考串的匹配关系.

下面给出受限串模式的串预测编解码操作过程.

设 m 表示受限串模式, $S(m)$ 表示受限串模式的集合. 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.

根据式(4), 对不同内容特性的当前编码单元, 采用基于率失真值(Rate Distortion Cost, RDCost)模式选择机制自适应选择一种受限串模式, 即遍历 $S(m)$ 中的所有 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(4).

$$m^* = \arg \min_{\forall m \in S(m)} D(m) + \lambda \times R(m) \quad (4)$$

其中, λ 表示失真与比特数的权重因子.

对当前 CU 用模式 m 进行编码, 是指按照图 4 所示在已重建区域内进行串匹配, 将当前 CU 中的像素集合采用式(5)的方式分解成若干个偏移串和/或坐标串和/或不可预测像素串的集合.

$$CU = \{str_1, str_2, \dots, str_n \mid str_i \in OFS, CDS, UPS\} \quad (5)$$

受限串的编码操作主要是对当前编码 CU, 首先编码受限串模式的标识; 然后依次编码受限串模式中的串类型标识及该串类型的匹配关系参数, 重复该步骤直至所有的像素编码完成. 过程 1 $Encode(CU)$ 给出了受限串预测编码过程的伪代码.

过程 1. $Encode(CU)$.

输入: CU 中的像素集合

输出: 受限串模式码流

编码受限模式标识;

$numCodedPixel = 0$;

WHILE($numCodedPixel < CU$ 总像素数目)

 编码串类型标识;

 IF(串类型是偏移串)

 编码 **Offset**;

 编码 **Length**;

 ELSE IF(串类型是坐标串)

 编码 **pic_coor** 所在数组的索引;

 编码 **duplication_count**;

Length = **duplication_count**;

 ENDELSE

 ELSE(串类型是不可预测像素串)

 编码不可预测像素;

Length = 1;

 ENDELSE

ENDIF

$numCodedPixel += Length$;

ENDWHILE

受限串的解码操作主要是对当前解码 CU, 首先解码受限串模式标识; 然后依次解码受限串模式中的串类型标识及该串类型的匹配关系参数: 如果是偏移串类型, 则将重建串的 **Length** 个像素值从由 **Offset** 确定的参考串位置复制到当前串位置; 如果是坐标串类型, 则将重建串的像素值从由 **pic_coor** 确定的图像位置所在的像素值重复 **duplication_count**

次至当前串位置;如果是不可预测像素串则从码流中解码一个像素的值,该像素的值即重建像素,重复该步骤直至所有的像素解码完成.

3.3 三种受限串模式介绍

根据已挖掘的屏幕图像的特征,下面分别介绍一般串模式、第一受限串模式和第二受限串模式三种 CU 层次的受限串模式.

(1) 一般串(General String,GS)模式

一般串模式采用图 4 中的受限串预测方式,包括偏移串、坐标串和不可预测像素串三种串类型.考虑到硬件和软件实现时的带宽问题,至少满足约束条件 1~3.

约束条件 1. 对于至少包含一个参考像素在 2 个 CTU(当前 CTU 和左边 CTU)范围之外的偏移串,其 SM 属性匹配长度必须大于等于 32 个像素.

约束条件 2. 每个 CU 所允许的偏移串的个数必须不大于当前 CU 的像素数的四分之一.

约束条件 3. 串匹配算法仅应用于 32×32 或以下的 CU.

(2) 第一受限串(Constrained String 1,CS1)模式

CS1 模式是满足约束条件 4 的 GS 模式的特例.在 CS1 模式中,一个 CU 只有一个偏移串或 N 个相等长度的偏移串,没有坐标串和不可预测像素串.

约束条件 4. 偏移串的长度取值为 CU 像素数目或者 CU 像素数目的 N 分之一.

(3) 第二受限串(Constrained String 2,CS2)模式

CS2 模式的偏移串模式是至少满足约束条件 5 的受限串模式,也是 GS 模式的另一个特例.

约束条件 5. 偏移串的 *Offset* (*OffsetX*, *OffsetY*) 的取值仅限于:当 CU 是垂直扫描时, *Offset* 的取值只能等于(1,0);当 CU 是水平扫描时, *Offset* 的取值只能等于(0,1).

基于三种受限串模式的 USP 算法是 USP 算法的一个实现特例.可以看出,USP 算法一方面对不同内容特征的屏幕图像区域自适应地采用不同的受限串模式,能够对各种各样的屏幕图像序列进行有效编码,从而适应屏幕图像的多样性;另一方面统一采用串预测方式进行编解码,具有统一的串预测编解码操作、语法元素和编解码框架.

4 USP 算法在 AVS2-SMCC 中的应用

4.1 编解码框架

融合通用串预测技术的 AVS2-SMCC 编解码系统架构见图 5 所示.编码器端分为通用串预测子系统与基于块匹配的传统混合编码子系统.对于输入 CU O ,通用串预测编码器分别进行 GS 模式、CS1 模式、CS2 模式的编码,然后进行基于 RDCost 选择器的选择,选择最小 RDCost 的受限串模式进行编码形成位流 b_1 并获得重建 CU P_1 ;与此同时,AVS2 基准档次中传统帧内与帧间模式选择出最优的模式,对其进行编码形成位流 b_2 并获得重建 CU P_2 .两者经过 RDCost 比较,最终选择其中的一种模式为最优的模式,输出至位流中.在 AVS2-SMCC 的实现中,采用 USP 编码的 CU 中用到的图像坐标的个数小于 32.

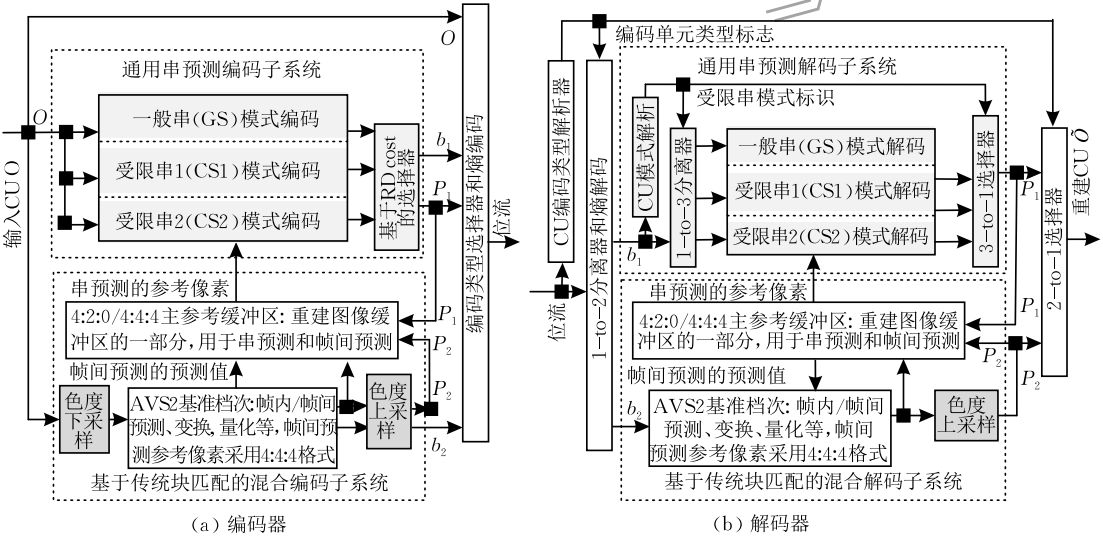


图 5 融合通用串预测算法的 AVS2-SMCC 编解码架构示意图

屏幕图像中的有些内容,使用 4:2:0 格式的既有 AVS2 基准档次可以达到更高的编码效率.因

此,在将像素数据送入 AVS2 基准档次编码器前,首先将 4:4:4 采样格式的数据经过色度下采样,转

换成 4:2:0 采样格式的数据,再进入编码器,编码完成后,再把重建的 4:2:0 采样格式的数据经过色度上采样转换为 4:4:4 采样格式的重构图像.本文采用的色度上下采样算法同文献[20].由于对于 AVS2 帧内与帧间编码器,重构图像的失真包括编码失真和色度下采样失真.而对通用串预测编码器,仅存在编码失真.为了解决这一问题,不同模式的 RDcost 计算时采用不同的色度失真调节因子(帧内模式调节因子为 0.2,USP 模式和帧间模式调节因子为 0.5).

解码器原理类似.文献[28]给出了 AVS2-SMCC 中 USP 算法详细的语法元素、相关语义、解析过程和解码过程.

4.2 关键技术

通用串预测算法的研究主要围绕如何确定通用串预测算法中的各种受限 SM 模式的组成和如何进

行参考串搜索、最佳参考串选择和串预测匹配关系参数熵编码等关键技术而展开,探索各种受限 SM 模式如何达到编码效率与复杂度的综合平衡.

4.2.1 GS 模式和 CS2 模式的关键技术

USP 的 GS 模式与 P2SM 算法类似,但是 GS 模式中对非 Above 的 *Offset* 采用最近串预测编码,而 P2SM 算法没有采用. GS 模式中采用坐标串而 P2SM 中采用 SRB 串;CS2 模式与 PLT 算法都是对匹配的位置进行限定的串匹配算法,在一定程度上有类似的地方,但是 CS2 模式中使用特殊位置的偏移串和坐标串,而 PLT 中使用索引串.

4.2.2 CS1 模式的关键技术

4.2.2.1 基于多种子模式的 CS1 模式

AVS2 中 CS1 模式包括大串子模式, N 等长串子模式与帧间 CS1 模式.图 6 给出了 3 种子模式图示.

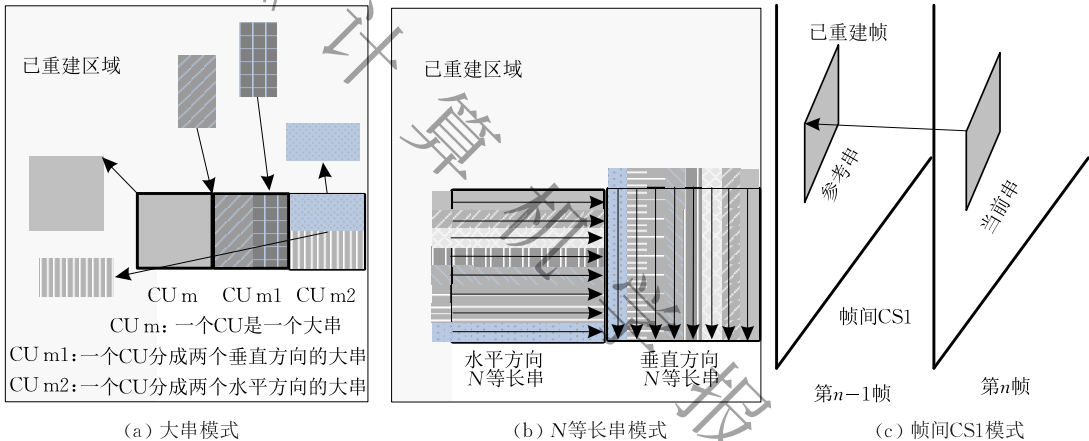


图 6 第一受限串模式中的 3 种子模式图示

可以看出,大串模式是指一个 CU 中只有一个偏移串或两个相等长度的偏移串. N 等长串模式是指一个 CU 中所有的串都与 CU 的高度或者宽度等长,且都是复制上方一行的像素或者左方一列的像素. N 等长串模式有 0.5% 左右编码效率的提升^[25]. 帧间 CS1 模式是指参考串在当前编码图像的前一帧图像中找到. 与传统的帧间预测算法不同的是,帧间 CS1 模式仅仅只是将参考串的位置扩展至按照编解码顺序的前一帧,并且是通过哈希搜索的方式在全球范围寻找最佳参考串. 帧间 CS1 模式对于 LDF 配置的编码效率有约 3% 提升^[25]. 对 USP 算法中 CS1 模式的多种子模式进行预编码,选出 RDCost 值最小的模式作为 CS1 模式的最佳模式.

4.2.2.2 基于混合搜索技术的大串模式

CS1 模式的大串模式根据串的大小,采用不同的搜索策略,即混合搜索技术. 混合搜索技术对不同

内容特性的 CU,自适应采用特殊位置搜索、一维搜索、二维搜索和哈希搜索相结合的混合搜索技术,达到尽快找地最佳参考串的目的.

表 1 给出了大串模式不同大小的串采用的混合搜索方式. 图 7 给出了一维搜索、二维搜索和哈希搜索的图示. 如图 7 所示,一维搜索是仅对 *OffsetX* 为 0 或者 *OffsetY* 为 0 的部分像素点(图示中黑色的像素点)进行的搜索. 二维搜索是对局部范围(如 2 个

表 1 大串模式不同大小的串采用的混合搜索技术		
串的大小	搜索范围	搜索策略
32×32,16×16,32×16,16×8、 16×32,8×16,4×8,8×4	全帧内 合法范围	特殊位置搜索 和一维搜索
	2CTU 内	一维搜索或一 维搜索和二维 搜索
	2CTU 外	哈希搜索
	全帧内 合法范围	特殊位置搜索

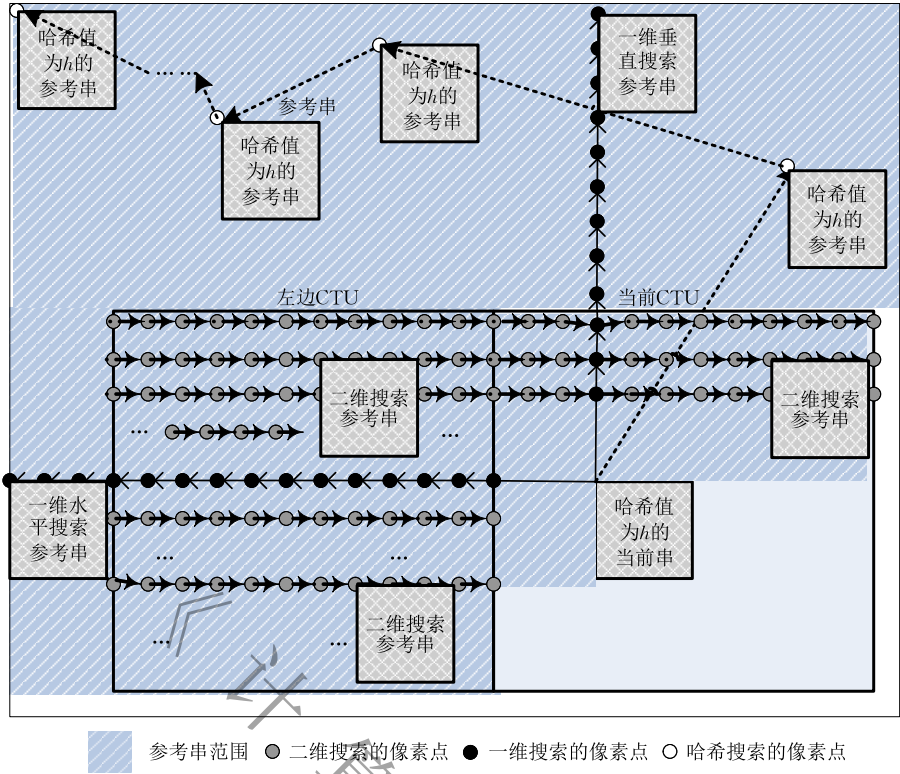


图 7 一维搜索、二维搜索和哈希搜索方式图示

编码树单元 (Coding Tree Unit, CTU) 内的像素 (图示中灰色的像素点) 进行的逐点搜索. 哈希搜索是对于局部范围以外的全局范围的与当前串具有相同哈希值的像素点 (图示中白色的像素点) 进行的逐点搜索. 哈希值的计算方式同 IBC 算法^[8].

特殊位置搜索是在进入其他搜索方式之前进行的最多可达 64 个特殊候选最佳参考串 *Offset* 集合中的快速搜索. 特殊候选最佳参考串 *Offset* 集合由下述 3 部分组成:

- (1) 若干个最近串 (编解码顺序上最近完成编解码的偏移串, 也称为 *N* 元组) 的 *Offset* (USP 算法中取值为 12 个).
- (2) 若干个相邻串 (帧内空间位置或帧间时间顺序上相邻的偏移串) *Offset* (USP 算法中取值为 2 个).
- (3) 其他搜索方式搜索过程中保留下来的 8 个 *Y* 分量的失真最小的候选最佳参考串 *Offset*.

随着搜索的进行, 上述三部分不重复的 *Offset* 以“先入先出”的规则累加至候选最佳参考串 *Offset* 集合. 因此特殊位置可以看成是最有潜力的“少而精”的候选最佳参考串的集合, 用于又快又好地找到最佳参考串.

对于 8×8 的 CU, 首先进行特殊位置的搜索和

一维搜索, 然后根据 CU 的数据特点 (如水平梯度和垂直梯度的最小值小于固定阈值) 自适应进行或不进行二维搜索, 最后进行哈希搜索. 每种搜索都按一定的起始像素顺序 (如图 7 所示的箭头表示每种搜索的起始像素顺序) 进行最佳参考串搜索, 一旦满足提前终止的条件, 搜索则结束.

4. 2. 2. 3 *Offset* 的自适应预测或直接编码方案

以编解码顺序最近完成编解码的 *Offset* 常常会重复出现. 根据这一特性, GS 模式的 *Offset* 编码采用文献^[15]中提出的一种自适应选择最近串预测方案或直接编码方案的算法.

以上特性同样适用于 CS1 模式中大串模式的 *Offset* 编码. 因此, 在借鉴 AVS2 帧间算法中位移矢量 (Motion Vector, MV) 采用相邻块预测编码方案的基础上, 提出了大串模式中基于自适应相邻串或最近串预测方案或直接编码的 *Offset* 编码方案. 如图 8 所示, 对于待编码的 *Offset*, 首先用空间或时间上相邻块的 *Offset* 进行预测编码, 若取得的 RDcost 是最优的, 则采用相邻串模式编码; 否则, 进行最近串 *Offset* 预测编码, 如果当前 *Offset* 在最近 *N* 元组中, 则采用 *N* 元组模式 (*N* 取值为 12) 编码; 否则采用直接编码方案.

大串模式中相邻串的选取规则同 AVS2 帧间

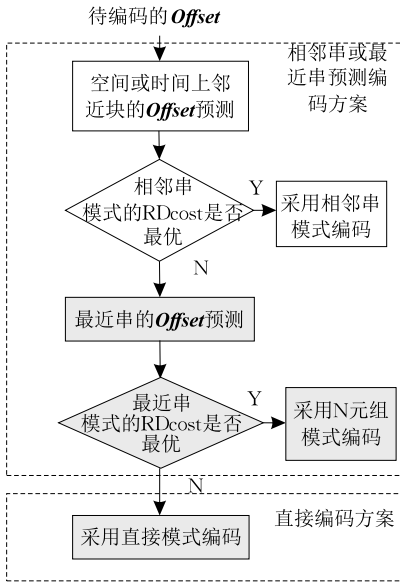


图 8 大串模式中的 Offset 编码算法流程图

预测中相邻块的选取规则^[25]. 大串模式中 N 元组的建立与更新算法与 GS 模式基本相同^[20], 不同的是大串模式中 N 元组不包括采用相邻串模式编码的 Offset. 直接编码方案是直接编码 Offset 是否为零位、符号位(是否为零位和符号位称为前缀)以及绝对值(后缀). 根据 Offset 具有的相关性和固有的帧内匹配特性^[23], 采用基于区域划分与调整的 Offset 直接编码算法. 其中 $adjustedOffsetX$ 与 $adjustedOffsetY$ 是对 OffsetX 与 OffsetY 的绝对值进行调整后的数值^[23].

算法 1 给出了基于区域划分与调整的 Offset 直接编码算法. 其中 EG5 与 EG1 分别表示 5 阶和 1 阶指数哥伦布编码.

算法 1. 基于区域划分与调整的 Offset 直接编码算法.

输入: OffsetX, OffsetY, adjustedOffsetX, adjustedOffsetY

输出: Offset 码流数据

编码 OffsetY 是否为零位

IF(OffsetY 为零)

编码 adjustedOffsetX 是否为零位;

IF(adjustedOffsetX 不为零)

EG5(abs(adjustedOffsetX)-1);

ENDIF

ELSE

编码 OffsetX 是否为零位

IF(OffsetX 为零)

编码 adjustedOffsetY 是否为零位;

IF (adjustedOffsetY 不为零)

EG5(abs(adjustedOffsetY)-1);

ENDIF

ENDIF

ELSE

EG1(abs(OffsetX)-1);

编码 OffsetX 的符号位;

EG1(abs(OffsetY)-1);

IF(OffsetX<0)

编码 OffsetY 的符号位;

ENDIF

ENDElse

ENDElse

ENDIF

5 实验结果与分析

本文采用的测试数据集来自 AVS2-SMCC 通用测试条件^[30]中的通用测试序列. 使用 13 个 YUV 4:4:4 采样格式, 每个分量为 8 比特的测试序列.

为了便于说明本文算法的特性, 将本文提出的算法及其比较的基准主要分为以下 4 种算法:

(1) RD17.0Pseudo444, 简称 RDP444. 在 AVS2-P2 基准档次的参考软件 RD17.0 上增加对输入原始图像的色度下采样和对输出重构图像的色度上采样, 以实现 YUV 4:4:4 序列的编码和解码, 相当于上图 5 中下方的虚线方框.

(2) USP. 在 RDP444 基础上实现的通用串预测算法^[25,28].

(3) HM. HEVC 4:4:4 扩展版的参考软件 HM16.7^①.

(4) SCM. 基于 HM16.7 的 HEVC-SCC 的参考软件^②, 是目前 HEVC SCC 编码效率最高的版本.

编码参数配置采用 AVS2-SMCC 通用测试条件中的标准配置^[30]. 采用有损编码模式, 有损编码模式的 4 个 QP 值 HEVC 为 {22, 27, 32, 37}, AVS2 为 {25, 32, 38, 45}, 使用 All Intra(AI, 即全部图像都是 I 图像)和 Low Delay F(LDF, 即第一幅图像是 I 图像, 其余都是 F 图像)两种编码图像类型配置.

实验结果从编码效率和复杂度来衡量算法的性能. 由于 USP 算法中引入了 YUV 4:4:4 与

① https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/

② https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.7+SCM-6.0/

YUV 4:2:0 之间的色度下采样和色度上采样,因此除了采用 HEVC 核心实验中用于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——Y、U、V 分量的 BD-rate^[26-27]降低率来衡量,还采用了 CPSNR——一种综合的 PSNR 的衡量方法^[31](该衡量方法已经被 AVS2 标准采纳).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.

表 2 给出了基于 AVS2 的 RDP444、USP 和基

于 HEVC 的 HM、SCM 四种算法 AI 配置、LDF 配置的编码效率和复杂度的两组(RDP444 相对 USP、HM 相对 SCM)比较结果.编码效率比较中,每行数据表示一个序列类别(每个类别由 1~8 个序列组成),分别给出了 AI、LDF 配置分量 Y、U、V 的平均 BD-rate 降低率. BD-rate 降低率表示在相同的峰值信噪比(Peak Signal to Noise Ratio, PSNR)取值下比特率节省的百分比.负值表示比特率降低,编码效率提高.

表 2 RDP444、USP、HM16.7 和 SCM 算法编码效率与复杂度的两组比较结果

对比算法	测试序列类别	编码效率比较/%								复杂度比较/%		
		AI 配置				LDF 配置				AI 配置	LDF 配置	
		Y	U	V	CPNSR	Y	U	V	CPNSR			
USP vs. RDP444	TGM	-81.82	-100.00	-100.00	-98.50	-63.41	-99.99	-100.00	-94.61	编码时间比值	268.27	167.23
	MC	-61.50	-99.96	-99.76	-78.97	-24.63	-99.91	-99.33	-67.65			
	G	2.96	-29.37	-21.99	-4.72	16.23	-77.39	-68.47	-5.75	解码时间比值	108.54	289.32
SCM6.0 vs. HM16.7	TGM	-77.45	-78.74	-78.49	-77.71	-62.70	-66.23	-66.10	-63.56	编码时间比值	345.78	53.66
	MC	-61.70	-65.60	-68.11	-62.33	-41.28	-46.70	-50.03	-42.45			
	G	-2.17	-15.34	-11.75	-3.56	0.25	-7.66	-5.30	-0.47	解码时间比值	60.95	92.42

表 3 给出了基于 AVS2 的 USP 算法和基于 HEVC 的 SCM 算法两种算法 AI 配置总体编码效率比较结果.

表 3 USP 算法与 SCM 算法 AI 配置编码效率比较结果
(单位:%)

对比算法	测试序列类别	AI			
		Y	U	V	C(CPSNR)
USP vs. SCM	TGM	-23.22	-18.33	-19.37	-22.09
	MC	-9.86	21.12	19.83	-3.56

USP_CS2Off 算法表示在 USP 算法中仅仅关闭 CS2 模式;同理,USP_GSOff 和 USP_CS1Off 分别表示 USP 中仅仅关闭 GS 模式和 CS1 模式.表 4 给出了 USP、USP_GSOff、USP_CS1Off 和 USP_CS2Off 四种算法的三组(其中 USP 相对 USP_GSOff 的实验用于衡量 GS 模式的作用;USP 相对 USP_CS1Off 的实验用于衡量 CS1 模式的作用;

USP 相对 USP_CS2Off 的实验用于衡量 CS2 模式的作用)AI 配置的编码效率比较结果.

表 4 USP 算法中三种受限串模式编码效率三组比较结果
(单位:%)

对比算法	测试序列类别	AI			
		Y	U	V	C(CPSNR)
USP vs. USP_GSOff	TGM	-31.57	-32.21	-31.29	-31.77
	MC	-15.95	-14.34	-14.75	-15.45
	G	-0.01	-0.52	-0.35	-0.08
USP vs. USP_CS1Off	TGM	-11.12	-10.71	-10.66	-10.95
	MC	6.18	-5.33	-4.26	-5.57
	G	0.21	-0.91	-0.62	-0.33
USP vs. USP_CS2Off	TGM	-1.50	-2.02	-2.34	-1.72
	MC	-4.16	-9.13	-8.80	-5.44
	G	0.54	-16.94	-9.20	-2.33

图 9 给出了 AVS2-SMCC 通用测试条件中的 3 个典型的屏幕图像序列的第一帧;图 10 给出了图 9 中的序列使用这四种算法进行 AI 配置编码的 RD 曲线图.

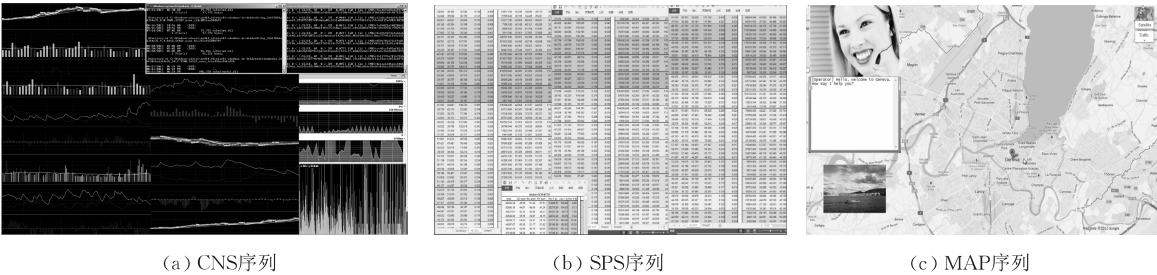


图 9 AVS2-SMCC 通用测试条件^[30]中典型测试序列 CNS、SPS 和 MAP 序列的第一帧图示

从以上实验结果,可以得出以下结论:

(1)从编码效率来看,USP 算法较 RDP444 的平均编码效率大幅提升.从表 2 中可知,USP 算法与 RDP444 算法相比,对于 TGM 序列类别的 Y、U、V 分量和基于 CPNSR 三分量综合平均 BD-rate 降低率,AI 配置下分别为 81.82%、100%、100% 和 98.5%;LDF 配置下分别为 63.41%、99.9%、100% 和 94.61%.同时,从表 3 中可知,USP 算法较 SCM 算法平均编码效率提升明显:对于 TGM 序列类别, AI 配置下 Y、U、V 分量平均 BD-rate 降低率分别为 23.22%、18.33%、19.37%;对于 TGM 和 MC 序列类别,基于 CPNSR 的三分量综合的平均 BD-rate 降低率分别为 22.09% 和 3.56%,已经大幅超过 SCM 的编码效率.

(2)从复杂度来看,对于 AI 配置和 LDF 配置, USP 算法较 RDP444 编码复杂度分别增加了约 168.27%、67.23%.编码时间增加一方面在于 USP 编码器中进行最优匹配串 的搜索;另一方面,AVS2 的帧内与帧间算法是基于 4:2:0 格式的,处理的 UV 分量的像素值数目是 USP 算法的 1/4.解码复杂度 AI 配置增加了 8.54%,LDF 配置增加了 189.32%(LDF 配置除了第一帧外都是 F 图像, RDP444 中大多数 CU 选中 4:2:0 帧间 skip 模式,其解码复杂度很低,相对来说,USP 中 4:4:4 串匹配需要处理的 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).事实上, USP 的 AI 配置增加的编码复杂度远远低于 SCM 增加的编码复杂度(约为 245.78%).

(3)不同受限串模式对屏幕图像的平均编码效率都有一定的提升.从表 4 中可知,USP 算法与 USP_GSOFF 算法相比, AI 配置下,对于 TGM 序列类别, Y、U、V 分量平均 BD-rate 降低率和基于 CPNSR 三分量综合平均 BD-rate 降低率分别为 31.57%、32.21%、31.29% 和 31.77%; USP 算法与 USP_CS1Off 算法相比,对于 TGM 序列类别, Y、U、V 分量和基于 CPNSR 三分量综合平均 BD-rate 降低率分别为 11.12%、10.71%、10.66% 和 10.95%; USP 算法与 USP_CS2Off 算法相比,对于 MC 序列类别, Y、U、V 分量和基于 CPNSR 三分量综合平均 BD-rate 降低率分别为 4.16%、9.13%、8.80% 和 5.44%.

(4)不同受限串模式对不同内容特性的屏幕图像序列作用不同,体现出不可替代的作用.从图 10 中可知,对于 CNS 序列,CS1 模式的作用最大,GS 模式次之,CS2 模式的作用最小;对于 SPS 序列,GS

模式的作用非常大,CS1 模式次之,CS2 模式的作用最小;对于 MAP 序列,CS2 模式的作用最大,GS 模式次之,CS1 模式的作用最小.

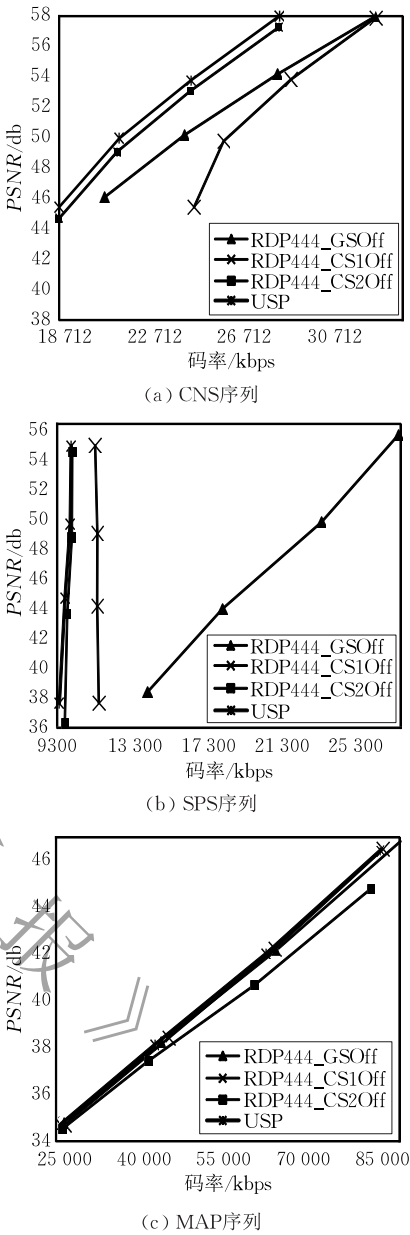


图 10 典型序列有损编码 AI 配置 Y 分量 RD 曲线比较

(5)提出的 USP 算法性能的提升程度与屏幕图像类别相关.从表 2、表 3 和表 4 中可以看出,对于 TGM 类别,提升相当大, M 类别有一定程度提升,而对于 G 类别几乎没有提升.

6 结 论

在 AVS2-P2 框架融合串匹配算法的全色度屏幕混合内容的编码算法基础上,提出了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及其关键技术。每种受限串模式采用偏移串、坐标串和不可预测像素串三种类型的串之一或其组合来实现。提出的 USP 算法能大幅提高不同内容特性的屏幕图像序列的编码效率,尤其对 TGM 类别和 MC 序列特别有效。与目前国际最新 HEVC 中 SCC 扩展版相比,提出的算法在编解码复杂度增加较少的前提下,TGM 和 MC 序列类别的平均 CPSNR 分别降低 22.09% 和 3.56%。

下一步的工作包括(1)进一步提升 LDF 配置的编码效率和降低 USP 算法的编解码复杂度。(2)进一步探索 USP 算法中的各种受限串模式及其与 AVS2-P2 基准档次中的现有帧内、帧间编码模式的兼容性问题。(3)将 USP 算法应用到无再损^[32]、高质量的下一代混合内容视频和互联网内容编码^[33]中。

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): 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] Shen Yan-Fei, Li Jin-Tao, Zhu Zhen-Min, Zhang Yong-Dong. High efficiency video coding. *Chinese Journal of Computers*, 2013, 36(11): 2340-2355 (in Chinese)
(沈燕飞, 李锦涛, 朱珍民, 张勇东. 高效视频编码. *计算机学报*, 2013, 36(11): 2340-2355)
- [5] Gao Wen, Ma Si-Wei. *Advanced Video Coding Systems*. Berlin, Germany: Springer, 2015
- [6] Wang Shan-She, Luo Fa-Lei, Ma Si-Wei. Overview of the second generation AVS video coding standard (AVS2). *ZTE Communications*, 2016, (1): 1-13
- [7] AVS Group. Resolutions of the 54th AVS workgroup meeting. Beijing: AVS workgroup, AVS Document: N2203, 2015 (in Chinese)
(数字音视频编解码技术标准(AVS)工作组. AVS 工作组第 54 次会议大会决议. 北京: AVS 工作组, AVS 文件: N2203, 2015)
- [8] Xu X, Liu S, Chuang T, 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
- [9] 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
- [10] Wang S, Lin T. A unified LZ and hybrid coding for compound image partial-lossless compression//*Proceedings of the IEEE International Conference Image and Signal Processing*. Tianjin, China, 2009: 1-5
- [11] Wang S, Lin T. United coding for compound image compression //*Proceedings of the IEEE International Conference Image and Signal Processing*. Yantai, China, 2010: 566-570
- [12] Wang S, Lin T. Compound image compression based on unified LZ and hybrid coding. *IET Image Processing*, 2013, 7(5): 484-499
- [13] Wang S, Lin T. United coding method for compound image compression. *Multimedia Tools and Applications*, 2014, 71(3): 1263-1282
- [14] 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
- [15] 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
- [16] Zhao Li-Ping, Tin 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
- [17] Zhao Li-Ping, Zhou Kai-Lun, Guo Jing, et al. Universal string matching for ultra high quality and ultra high efficiency SCC. Shanghai: Tongji University, JVET Document: B0048, 2016
- [18] Zhao Li-Ping, Zhou Kai-Lun, Guo Jing, et al. A universal string matching approach to screen content coding. *IEEE Transactions on Multimedia*, 2018, 20(4): 796-809
- [19] 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)
- [20] 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*, to appear (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*, 录用待发表)
- [21] Lin Tao, Zhao Li-Ping, Zhou Kai-Lun, et al. Pixel string matching for screen content coding in AVS2-P2. Shanghai: Tongji University, AVS Document: M3844, 2015 (in Chinese)
(林涛, 赵利平, 周开伦等. 在现有 AVS2-P2 编码框架中增加串匹配编码单元类型的屏幕混合内容编码技术. 上海: 同济大学, AVS 文件: M3844, 2015)
- [22] Lin Tao, Zhao Li-Ping, Zhou Kai-Lun. Universal string prediction for screen and mixed content coding in AVS2-P2. Shanghai: Tongji University, AVS Document: M3915, 2016 (in Chinese)
(林涛, 赵利平, 周开伦等. 在现有 AVS2-P2 框架中增加通用串预测编码单元类型的屏幕与混合内容编码技术. 上海: 同济大学, AVS 文件: M3915, 2016)

- [23] Zhao Li-Ping, Lin Tao, Zhou Kai-Lun, et al. The improvements on universal string prediction for screen and mixed content coding in AVS2-P2. Shanghai: Tongji University, AVS Document: M3958, 2016(in Chinese)
(赵利平, 林涛, 周凯伦等. 屏幕与混合内容编码中通用串预测(USP)方法的改进. 上海: 同济大学, AVS 文件: M3958, 2016)
- [24] Lin Tao, Zhao Li-Ping, Zhou Kai-Lun, et al. The improvements on universal string prediction for screen and mixed content coding in AVS2-P2. Shanghai: Tongji University, AVS Document: M3958, 2016(in Chinese)
(林涛, 赵利平, 周开伦等. 屏幕与混合内容编码中通用串预测(USP)方法的改进. 上海: 同济大学, AVS 文件: M4021, 2016)
- [25] Lin Tao, Guo Jing, Zhao Li-Ping, et al. The improvements on universal string prediction for screen and mixed content coding in AVS2-P2. Shanghai: Tongji University, AVS Document: M4081, 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. 上海: 同济大学, AVS 文件: M4081, 2016)
- [26] Bjøntegaard G. Calculation of average PSNR differences between RD-curves. Norway: Telenor Satellite Services, ITU-T SG16 Q. 6 Document: VCEG-M33, 2001
- [27] Bjøntegaard G. Improvements of the BD-PSNR model. Norway: Telenor Satellite Services, ITU-T SG16 Q. 6 Document: VCEG-A111, 2008
- [28] AVS Workgroup. AVS2-P2-SMC-WD3.0. Beijing: AVS Workgroup, AVS Document: N2388, 2016(in Chinese)
(数字音视频编解码技术标准(AVS)工作组. 通用串预测方式的语法、语义、解析过程、解码过程. 北京: AVS 工作组, AVS 文件: N2388, 2016)
- [29] Yu H P, Cohen R, Rapaka K, et al. Common test conditions for screen content coding. Geneva: Joint Collaborative Team on Video Coding, JCTVC Document: X1015, 2016
- [30] AVS Workgroup. Common test condition far screen and mixed content coding in AVS2-P2. Beijing: AVS Workgroup, AVS Document: N2387, 2016(in Chinese)
(数字音视频编解码技术标准(AVS)工作组. AVS2-P2 屏幕与混合内容视频编码(S8-MCVC)通用测试条件. 北京: AVS 工作组, AVS 文件: N2387, 2016)
- [31] Wu Cheng-Jia, Chen Da-Peng, Chen Fang, et al. A method for calculating PSNR. Shanghai: Shanghai University, AVS Document: M3830, 2016(in Chinese)
(吴成家, 陈大鹏, 陈芳等. 一种计算综合 PSNR 的方法. 上海: 上海大学, 数字音视频编解码技术标准工作组, 文: M3830, 2016)
- [32] Tao Lin. Achieving re-loss-free video coding. IEEE Signal Processing Letters, 2009, 16(4): 323-326
- [33] AVS Group. The technical requirements V9.0 for the next generation of AVS video coding standards. Beijing: AVS Workgroup, AVS Document: N2401, 2017(in Chinese)
(数字音视频编解码技术标准(AVS)工作组. 下一代 AVS 视频编码标准的技术需求 V9.0. 北京: AVS 工作组, AVS 文件: N2401, 2017)



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

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

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

GUO Jing, M. S. candidate. His research interests focus on screen content coding.

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.

Three major HEVC SCC tools developed in recent years are: Intra Block Copy (IBC), palette coding (PLT) and String Matching (SM).

From video coding point of view, the most significant and relevant characteristic of screen content is its diversity.

Due to the diversity of the screen and mixed content, a pixel string matching technique proposed in AVS2 (the second generation of the Audio Video Coding Standard) SMCC (AVS2-SMCC) extension is not efficiency for some screen and mixed contents. In order to solve this problem, a universal string prediction (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, which have been adopted by the AVS2-SMCC extension draft version.

This work is supported by the National Natural Science Foundation of China (Nos. 61871289 and 61601200) and the Natural Science Foundation of Zhejiang Province (No. LY19F020015).