

屏幕内容编码方法研究进展

刘丹^{1,2} 陈规胜^{1,3} 宋传鸣^{1,2} 何兴¹ 王相海¹

¹(辽宁师范大学计算机与信息技术学院 辽宁大连 116029)

²(大连理工大学电子信息与电气工程学部 辽宁大连 116024)

³(吉林大学计算机科学与技术学院 长春 130012)

(liudan_dl@lnnu.edu.cn)

Research Advances in Screen Content Coding Methods

Liu Dan^{1,2}, Chen Guisheng^{1,3}, Song Chuanming^{1,2}, He Xing¹, and Wang Xianghai¹

¹(School of Computer and Information Technology, Liaoning Normal University, Dalian, Liaoning 116029)

²(Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, Liaoning 116024)

³(College of Computer Science and Technology, Jilin University, Changchun 130012)

Abstract With the widespread promotion of the applications such as cloud computing, virtual desktop, and so on, screen content image has become an integral part of the new generation of cloud—mobile computing model. It is one of the hot issues of video coding field to investigate the screen content coding methods with high compression efficiency, good real-time performance, and moderate computational complexity. On introducing the statistical characteristics of screen content image presented in the spatial domain, the frequency domain, the temporal domain, as well as the color space respectively, this study focuses on typical coding methods of the discontinuous tone images. The state-of-art methods are classified into seven categories, namely the palette-index map based methods, the template matching based methods, the block matching based methods, the dictionary-based methods, the shape representation based methods, the temporal-domain coding methods, as well as the chroma component coding methods. Then the screen content coding methods using a hybrid framework is further summarized. Meanwhile, the advantages vs. disadvantages of various methods are also compared, analyzed and discussed. Based on the above, the progress of drafting the international HEVC-SCC coding standard is introduced, and the development trend of the screen content coding is forecast in the near future.

Key words video coding; image coding; screen content; screen image; compound image; survey

摘要 随着云计算、虚拟桌面等的普遍推广,屏幕内容图像已成为新一代云——移动计算模型——不可或缺的一部分。研究压缩效率高、实时性好、复杂性适中的屏幕内容编码方法是目前视频编码领域的热点问题之一。从空间域、频率域、时间域和颜色空间4方面分析了屏幕内容图像的数据统计特性,进而

收稿日期:2016-08-22 修回日期:2017-02-07

基金项目:国家自然科学基金项目(61402214,41271422);教育部高等学校博士学科点专项科研基金项目(20132136110002);辽宁省教育厅科学研究一般项目(L201683681);大连市青年科技之星项目支持计划项目(2015R069,2016RQ046)

This work was supported by the National Natural Science Foundation of China (61402214, 41271422), the Specialized Research Fund for the Doctoral Program of Higher Education of China (20132136110002), the Foundation of Science and Research for Higher Education of Liaoning Province (L201683681), and the Dalian Foundation for Youth Science and Technology Star (2015R069, 2016RQ046).

通信作者:宋传鸣(chmsong@lnnu.edu.cn)

重点阐述不连续色调图像的典型编码方法,将现有方法分为基于调色板-索引图的编码算法、基于模板匹配的编码算法、基于块匹配的编码算法、基于字典的编码算法、基于形状表示的编码算法以及时间域编码方法、色度编码方法,并进一步总结基于混合框架的屏幕内容编码方法,对各类算法的优势和存在的不足进行比较、分析和讨论.在此基础上,介绍了 HEVC-SCC 编码国际标准制定工作的进展,并对屏幕内容编码的未来发展趋势进行了展望.

关键词 视频编码;图像编码;屏幕内容;屏幕图像;复合图像;综述

中图分类号 TN911.73; TP37

随着计算机和网络技术的发展,数字化的文本、图像、图形和视频等替代了传统的模拟媒体,这使得媒体的编辑和传播变得愈加便捷,又促使媒体由单一形式向复合形式演变,如网页、PDF (portable document format) 文档、扫描的电子文档、幻灯片和海报等图像中均包含了多种形式的媒体,研究者们将此类图像称为“复合图像”(compound image)^[1]. 而近 5 年,视频会议、在线教学、远程医疗和远程桌面等逐年普及,尤其是云计算取得了迅猛发展,又出现了虚拟桌面、桌面云、WiFi 显示、无线 HDMI (high definition multimedia interface) 等应用^[2],这些应用均要求把本地计算机屏幕显示的内容传输到远程终端上并显示,以实现屏幕共享.由于屏幕内容一般由各种应用软件生成,如办公软件、3D 游戏、电影动画、地理信息系统、网络浏览器等,图像往往由文本、图形和自然图像等若干不同类型的区域混合而成.这类复合图像又被称为“屏幕内容图像”,它已成为新一代云——移动计算模型——中不可或缺的一部分.一方面,屏幕视频的数据量庞大(1 s 的 1920×1080 分辨率@50 Hz 的屏幕视频数据量高达 297 MB),且屏幕共享的交互性对编码、传输的实时性提出了较高要求,许多应用中甚至在 10 Gb/s 的网络带宽下都不能满足高清屏幕视频的传输需求^[3].另一方面,典型的 JPEG(joint picture experts group),JPEG 2000, JPEG-LS 对屏幕内容中的自然图像部分有很高的压缩效率,却对诸如线条、文字、图形边界等非连续色调内容显得无能为力;而一些对非连续色调内容编码效率较高的算法,如 DjVu, MRC(mixed raster content),又对自然图像内容的压缩能力有限,且计算复杂度高.在这样的情况下,研究压缩效率高、实时性好、复杂性适中的面向复合图像,尤其是屏幕图像的编码方法则显得尤其重要,也为实际应用所亟需.

最初,研究人员通过改变参数的方式来利用 JPEG, JPEG 2000 等标准算法压缩复合图像.例如,

文献[4-5]通过自适应调整量化步长提出了 JPEG 的一种扩展框架.为了保持清晰的文本/图形边界,该方法采用较小的步长量化文本/图形区域.文献[6]在 JPEG 的基础上,根据图像内容自适应地调整编码器的率失真,从而为文本/图形区域分配更多的位.文献[7]则采用了 H.264/AVC 的帧内编码模式,并通过改变不同宏块的量化参数来分别满足文本块和图像块的压缩要求.尽管上述算法的参数可根据内容的不同进行自适应地调整或者交互设定,其编码效率往往不能令人满意^[8].于是,研究人员一方面逐步总结出复合图像,特别是屏幕内容图像,相对于自然图像的特殊性;另一方面,他们也发现屏幕图像和其他类型的复合图像相比,同样存在一定不同之处,例如:由于尘埃污染和扫描仪本身的缘故,扫描复合图像存在大量固有的噪声,而屏幕内容图像则不受噪声影响;多幅连续的屏幕内容图像之间存在沿着时间维的强相关性,可以像视频序列那样采用帧内预测编码;而且,屏幕内容图像编码和扫描复合图像编码在应用上也有不同,前者需保证屏幕共享的交互性和实时性,编码算法需兼顾压缩效率和计算复杂度,而后者则主要应用于替代纸质文件实现文档数字化,其压缩效率需要率先考虑,算法复杂度则相对次要.

为此,为了应对诸多应用愈加迫切的屏幕内容图像/视频的编码需求,国际视频编码联合协作组(Joint Collaborative Team on Video Coding, JCT-VC)近 2 年来成立了专门工作组负责制定高效率视频编码(high efficiency video coding, HEVC)的扩展标准——屏幕内容编码(HEVC-screen content coding extension, HEVC-SCC),并在世界范围内征集方案^[9-10],成为视频/图像编码领域的研究热点之一,并形成了一系列适合于屏幕内容编码的技术标准.本文也将注意力集中于此.

本文从屏幕内容图像的数据统计特性出发,重点阐述屏幕内容图像/视频的编码进展.将现有方法

分为7类,即基于调色板-索引图的编码算法、基于模板匹配的编码算法、基于块匹配的编码算法、基于字典的编码算法、基于形状表示的编码算法、时间域编码方法,以及色度分量编码方法,并总结了屏幕内容编码的混合框架,对各类算法的优势和不足进行了比较和分析.最后,对屏幕内容编码方法的未来发展进行了展望.

1 屏幕内容图像的统计特性

对于 JPEG, JPEG 2000 和 H.264/AVC, HEVC 等面向自然图像、视频的预测+变换结构的编码器,其有效性建立在自然图像和视频数据的2个统计特性上:1)空间、时间局部相关性;2)能量集中特性,即经过适当的数学变换(如离散余弦变换、小波变换等)之后,图像的绝大部分能量集中在低频分量,保存少数低频系数就可以非常近似地重构原始图像,从而取得很高的压缩率和主观质量.可见,对待编码数据统计特性的充分认识和利用对于编码算法的设计及其性能提升具有重要意义.在本节重点讨论屏幕内容的统计特性.

1.1 空间域的统计特性

屏幕内容往往由不连续色调的文本、图表、图形、图标等图像和连续色调的自然图像、视频片段等不同类型的区域混合而成,而其中的文本、图表、图形等元素一般由显示适配器生成,包含的纹理信息较多,但复杂纹理少,重复图案多,对比度高,线条细腻,边缘锐利,颜色种类少^[11-12];而自然图像则相反,颜色种类多,复杂纹理也多,边缘一般较为平滑,二者的对比情况如图1所示.

为了更好地验证这一点,文献[8]采用空间频率测度(spatial frequency measure, SFM)^[13]统计了3150个文本/图形块和4552个自然图像块的相邻像素值的变化程度:

$$SFM =$$

$$\sqrt{\frac{1}{MN} \left(\sum_{j=0}^{N-1} \sum_{i=0}^{M-1} (x_{i,j} - x_{i-1,j})^2 + \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (x_{i,j} - x_{i,j-1})^2 \right)}$$

其中, $x_{i,j}$ 表示图像块中 i 行 j 列的像素值, M 和 N 分别表示图像块的高度和宽度(实验中均设置为16).如图2所示,文本/图形内容块中相邻像素值的变化量广泛分布于10~160之间,而自然图像块的变化量则几乎全部集中在0~20之间,可见,文本/图形内容块的像素值变化量明显高于自然图像,其像素值的局部相关性不同于自然图像.

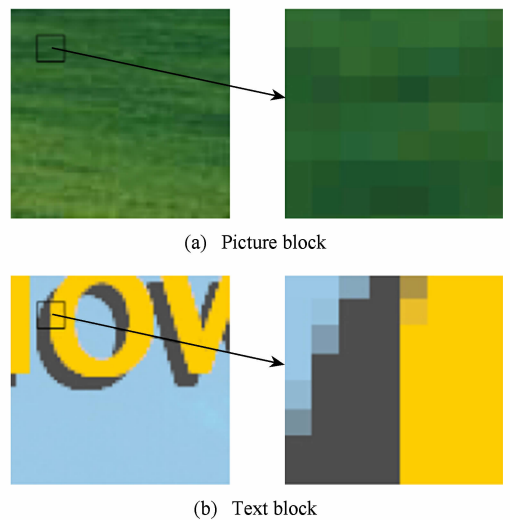


Fig. 1 Comparison between screen content images and natural images^[11]

图1 屏幕内容图像和自然图像的比较^[11]

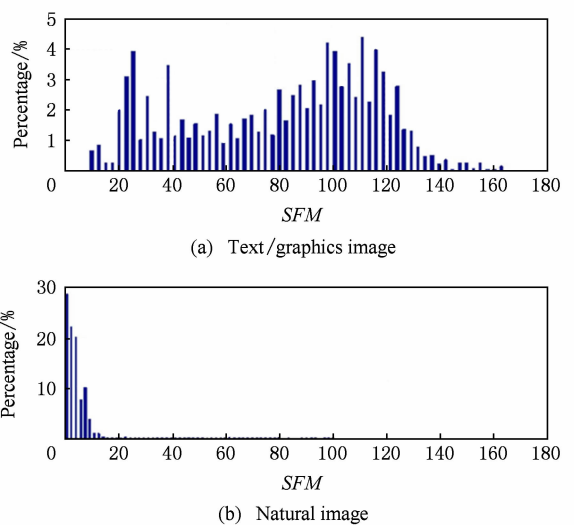


Fig. 2 SFM distribution of screen content images^[8]

图2 屏幕内容图像的 SFM 分布^[8]

1.2 频率域的统计特性

屏幕内容图像与自然图像在像素值分布方面的差异,导致了二者在频率域统计特性上的不同.图3给出了图1的2个图像块经过离散余弦变换(discrete cosine transform, DCT)后的交流系数绝对值分布.可见,自然图像块的能量集中于3个大幅值变换系数,而屏幕内容图像的能量分布则更加均匀,并未呈现明显的能量集中现象.文献[8]进一步在更大的数据集上采用频谱活动测度(spectral activity measure, SAM)度量了3150个文本/图形图像块和4552个自然图像块的图像变换系数幅值分布:

$$SAM = \frac{1}{MN} \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} |F(i,j)|^2}{\left(\prod_{i=0}^{M-1} \prod_{j=0}^{N-1} |F(i,j)|^2 \right)^{\frac{1}{MN}}}$$

其中, $F(i,j)$ 表示 i 行 j 列的 DCT 系数. SAM 的值越小, 变换系数的幅值分布就越均匀; 否则, 该分布就越不均匀, 能量越集中. 由图 4, 在较大数据集上的 SAM 统计结果与图 3 的情况一致, 这说明采用传统 JPEG 或 MPEG (moving picture experts group) 的“变换+量化”策略会损失图像的大部分能量, 产生明显的信息失真和主观观降质.

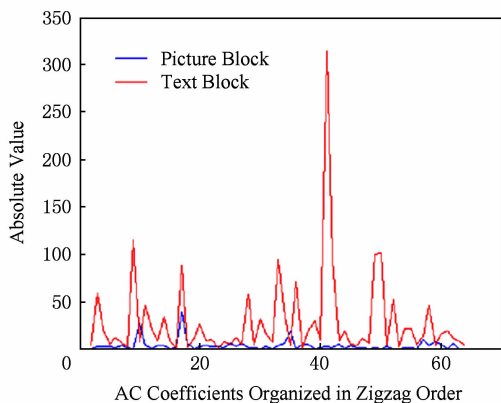


Fig. 3 DCT AC coefficients distribution of the two kinds of image blocks in Fig. 1^[11]

图 3 图 1 中 2 类图像块的 DCT 交流系数分布^[11]

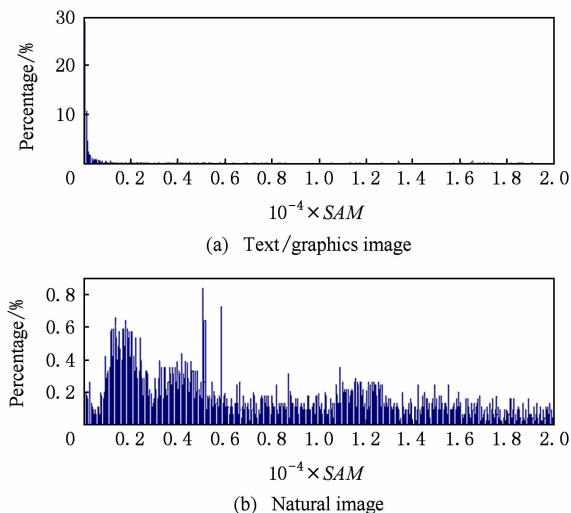


Fig. 4 SAM distribution of screen content images^[8]

图 4 屏幕内容图像的 SAM 分布^[8]

1.3 时间域统计特性

屏幕内容由应用软件和显示适配器捕获生成, 且主要记录软件界面和用户操作序列, 所以物体的运动特点与摄像机拍摄的自然视频也存在不同. 首先, 由于自然视频反映了物体在客观世界的运动, 它

的运动向量在理论上应是连续的; 而屏幕内容的运动在本质上是离散的^[14], 其最小运动幅度取决于显示设备的定位精度, 即运动向量为整数像素精度. 其次, 屏幕内容包含了窗口移动/放大/缩小、页面滚动、翻页、淡入淡出等运动, 其运动幅度较之自然视频更大, 模式也更复杂^[15]. 文献^[15]通过比较自然视频序列“Johnny”和屏幕视频“SlideShow”的相邻帧的均方差 (MSE) 发现, 前者的帧差变化非常平缓, 而后者的帧差则表现出剧烈变化, 既存在长时间的零值, 又有短时间的突变, 如图 5 所示. 屏幕视频的这个特点加大了帧间预测的难度, 导致编码器会产生波动较大的输出码率.

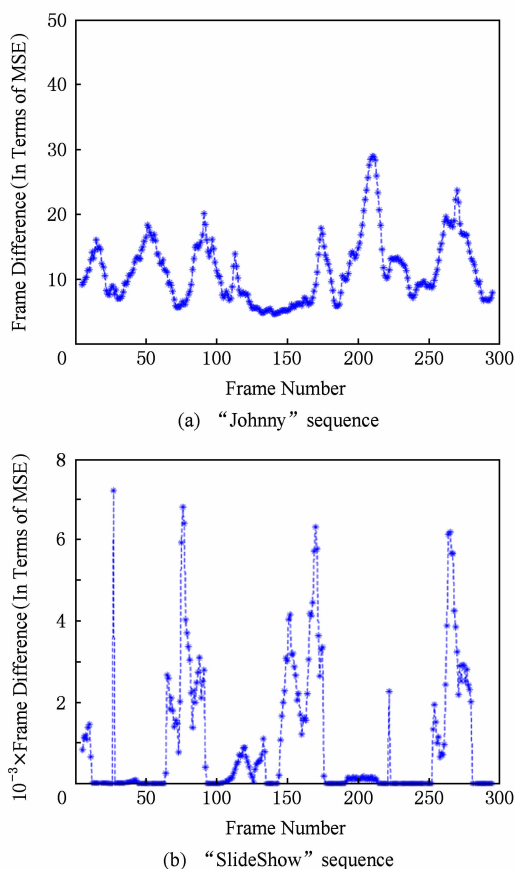


Fig. 5 Frame difference comparison between screen content sequence and natural video sequence^[15]

图 5 屏幕内容序列与自然视频序列的帧差对比^[15]

1.4 颜色空间采样的差异

典型自然图像和视频的标准格式普遍采用 $4:2:0$ 颜色空间采样, 如 YUV $4:2:0$. 然而, 由于屏幕内容包含大量细线条, 有些甚至是 1 个像素宽的单色线条, 如果仍旧采用 $4:2:0$ 的颜色采样方式, 就会产生人眼可见的明显失真^[16]. 为了实现画面质量的高保真, 目前屏幕内容图像和视频格式一般采用

4:4:4的颜色采样方式,如 RGB 4:4:4,这使得屏幕内容图像的色度冗余高于自然图像,前者的色度分量在编码过程中会消耗比后者更多的位。

1.5 小结

由 1.1~1.4 节比较可知,屏幕内容有着复杂的空间和频谱特征,变换编码已不适用于其不连续色调部分的高效率编码,这同时也解释了文献[4-7]等力图通过传统“帧内/帧间预测+变换结构”的单一编码器压缩屏幕内容等复合图像的效率不够理想的根本原因。于是,研究人员开始针对屏幕内容图像,尤其是不连续色调的文本、图表、图形等图像的数据分布特性,展开了以下 4 方面研究:

1) 根据不连续图像的纹理信息多、边缘锐利、颜色种类少等特点,研究适用于文本、图表、图形等屏幕图像/视频的空间域编码方法。

2) 根据不连续图像序列的运动幅度大、运动模式复杂、离散运动等特点,研究有效的、向量精度自适应的时间域编码(即帧间预测)方法。

3) 根据不连续图像的色度采样特点,研究减少屏幕内容图像/视频的色度信息冗余的方法。

4) 根据屏幕内容是连续色调图像和不连续色调图像混合体的特点,研究适用于屏幕内容编码的总体框架。

本文将分别详细阐述这 4 个方面的典型研究工作。

2 不连续色调图像的空间域编码方法

对于第 1 个方面,为了保证文本、图表、图形等不连续色调图像的主观保真度,目前通常采用无损或近无损编码算法对其进行压缩。早期,文献[17]采用 JPEG-LS 算法编码文本/图形图像块,但 JPEG-LS 是针对自然图像统计特性设计的,不完全适用;文献[18]采用 PNG(portable network graphic)算法编码不连续色调部分,但计算复杂度高,压缩比有限。除了应用标准的编码算法以外,研究人员针对不连续色调图像的像素分布特性提出了 5 类更加有效的方法,即基于调色板-索引图的编码算法、基于模板匹配的编码算法、基于块匹配的编码算法、基于字典的编码算法和基于形状表示的编码算法。

2.1 基于调色板-索引图的编码算法

考虑到不连续色调图像包含颜色种类少的统计特性,文献[11]提出将图像中出现次数较多的几种灰度值作为基本颜色(base color),分别为每种基本

颜色指定一个索引值构成调色板,再将原图像的每个灰度值用对应的索引值替代便形成一张索引图(index map);只要将经过熵编码后的调色板和索引图传输至解码端,即可利用调色板和索引图重构出复合图像,如图 6 所示。由于该算法发掘了文本、图表、图形等图像的数据特点,成为不连续色调图像的重要空间域编码方法之一。

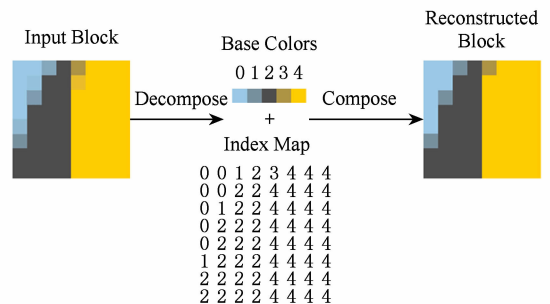


Fig. 6 Diagram of palette-index map coding algorithm^[11]

图 6 调色板-索引图编码算法示意图^[11]

该类算法包括 3 个主要步骤:调色板和索引图的生成环节、调色板编码环节和索引图编码环节,后续的研究工作便从这 3 方面分别展开。

2.1.1 调色板和索引图的生成方法

为了获得不连续色调图像的调色板和索引图,文献[11]提出首先采用 K -均值算法将像素值聚合为若干类,并用每个聚类中心的值替代相应类中的所有像素值;然后利用树结构向量量化(tree structure vector quantization, TSVQ)方法将像素值量化成 1~8 种基本颜色。文献[19-22]进一步将这种调色板-索引图方法作为一种帧内编码模式增加到 HEVC-SCC 标准中,并且文献[12, 22]提出一种新的调色板生成方法,其主要步骤是先将图像直方图进行排序,再从出现频率最高的像素值开始扫描,若当前像素值不能被调色板中的颜色以低于量化因子的误差来表示,则将当前像素值作为新的基本颜色增加到调色板。同时,文献[23]利用 HEVC 帧差编码将量化误差也传输至解码端,实现了屏幕内容的无损压缩。不过,由于基本颜色的数量 num 因图像内容而异,编码端就需要为每个图像块向解码端传送一个边信息来标识 num ,在一定程度上影响了编码效率。

于是,文献[24]提出为每个图像选取固定数量 b 的基本颜色,从而减少传输 num 的边信息量,其主要思路是选取图像中出现次数最多的 b 种像素值作为基本颜色,将其余像素值映射成离其最近且距离不超过阈值 Δ_1 的基本颜色,而将那些与最近的基本

颜色距离超过 Δ_1 的像素值映射成逃逸色 (escape color). 经过大量的实验统计, 文献[24]认为 4 种基本颜色为最佳数目. 该方法选取基本颜色的计算过程简单快速, 但是无法避免由于基本颜色过于集中、逃逸色导致图像失真偏大的情况. 因此, 文献[25]提出了一种构造最优调色板和索引图的率失真模型, 有效控制了屏幕内容图像的整体失真, 其 BD-Rate 指标较文献[20]降低了 4.7%, 不足之处是计算复杂度有所提高. 文献[8, 26]也讨论了类似的率失真模型, 并且给出一种率失真模型的动态规划解法, 其计算量低于文献[20].

此外, 文献[27]提出了一种索引图的优化方法, 其中心思想是: 若 2 个索引值相邻的次数较多, 则尽可能重新为它们分配 2 个连续的索引值. 该思路较为新颖, 将索引图的压缩效率提高了 43%, 但涉及较多的统计运算.

2.1.2 调色板编码方法

在建立了图像块(或编码单元)的调色板后, 文献[11]采用上下文自适应的算术编码生成调色板信息的压缩码流; 文献[8]通过量化将基本颜色的值域从 0~255 调整成 $\{0, 8, 16, 24, 32, \dots, 248, 255\}$, 再利用多符号算术编码进行压缩; 文献[24]则采用差分脉冲编码调制(differential pulse code modulator, DPCM)和霍夫曼编码对调色板进行压缩. 这 3 种方法均利用了调色板内部的数据冗余, 而考虑到不连续色调图像颜色种类少的特点, 相邻的或者位于相似纹理区域的不同图像块(或编码单元)的调色板之间必然也存在一定相关性, 显然上述 3 种方法并未发掘这种相关性. 鉴于此种情况, 就有研究者提出了局部调色板和全局调色板的解决思路.

所谓的“局部调色板”是指, 利用调色板数据的局部相关性而设计的编码方法. 文献[21, 28]采用左侧相邻编码单元的调色板来预测当前编码单元的调色板. 文献[29]提出一种调色板合并算法, 其主要思路是将当前编码单元上方和左侧相邻编码单元的调色板合并, 若当前编码单元的某个基本颜色在该合并调色板 C 中, 则把它在 C 中的索引值传输至解码端; 否则, 将该基本颜色直接传输至解码端. 除了调色板合并方法以外, 文献[12]建立了一个参考调色板, 若当前调色板的某个基本颜色位于参考调色板中, 则将该颜色在 2 个调色板中的索引之差传输至解码端; 否则, 就利用当前调色板中的前一种基本颜色预测待编码的基本颜色, 只将预测误差传输至解码端. 该方法利用了图像中不同块的基本颜色间的

相关性和同一图像块调色板内部的颜色相关性, 取得了不错的编码效率. 与此类似, 文献[26, 30]也给出了调色板的 2 种编码模式: 隐式模式和显示模式. 前者是用已编码图像块(编码单元)的调色板作为当前图像块(编码单元)的调色板; 后者则是从已编码图像块(编码单元)的调色板中提取出常用的基本颜色加入到一定容量的参考调色板, 再为每种参考颜色设置一个重用标识用于表示其在当前调色板中是否重用. 对于那些不在参考调色板、而出现在当前调色板中的颜色, 就直接将其编码进压缩码流. 同时, 为了保持参考调色板有较高的预测准确率, 文献[26, 30]又定义了参考调色板的更新操作, 实时地将当前调色板中出现的新基本颜色放入参考调色板, 而将其中不常用的基本颜色删除. 需要指出, 该编码方案已被 HEVC-SCC 测试模型接纳.

有研究表明^[31], 图像中广泛存在着非局部相关性, 即一个图像块往往与同一图像中位置不相邻的某块具有相似的边缘、纹理或轮廓等, 这样调色板数据也应具有非局部相关性. 而所谓的“全局调色板”就是利用调色板的这种非局部相关性所设计的编码方法. 文献[25]首先利用率失真方法建立整幅图像的调色板, 再采用凸优化方法从中选择出预测误差最小的若干种基本颜色建立全局调色板, 进而实现对每个图像块调色板的最优预测. 理论上, 该方法可达到优于上述所有调色板编码方法的效率, 但是它需要进行 2 轮调色板扫描和求解优化问题的多次迭代, 故此其计算量明显高于其他方法, 实用性受到了一定限制.

2.1.3 索引图编码方法

索引图是基于调色板-索引图的编码算法中体量最大的一部分数据, 其编码效率直接影响调色板编码算法的整体压缩性能, 所以索引图编码是调色板编码最主要的组成部分.

早期, 文献[11]采用上下文自适应的算术编码压缩索引图, 文献[8]采用上下文重映射和熵编码压缩索引图, 二者均利用了索引图的统计冗余. 但从 1.1 节可知, 索引图中还会出现重复的图案, 即所谓的局部和非局部数据相关性. 于是, 文献[12]采用 1D 串匹配的方法编码索引图, 如图 7(a)的索引图片段可编码成一系列二元、三元组序列 $(0, 14)(1, 1, 3)(0, 17)(1, 1, 3)\dots(0, 1)(0, 2)(0, 3)(0, 4)(1, 4, 4)$ ^[12]. 然而, 索引图中的重复图案几乎全部是 2D 图案, 如采用前面的 1D 串匹配还存在较大冗余. 如图 7(b), 1D 串匹配产生了多次重复的三元组 $(1, 8, 7)$, 故此

文献[12,32]提出了索引图的2D串匹配编码方法, 将1个重复的2D图案表示成四元组(匹配成功标志, 距离, 宽度, 高度), 例如图7(b)的阴影部分可表示成(0,0)(1,1,7,8)。可见, 充分发掘索引图的数据相关性可带来明显的编码增益。

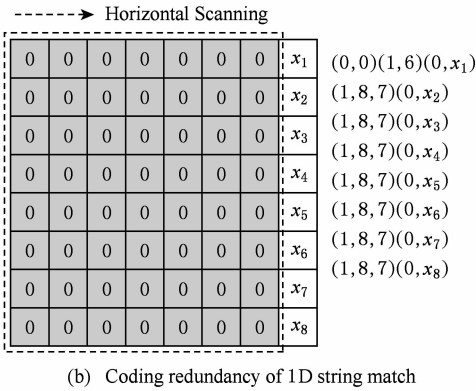
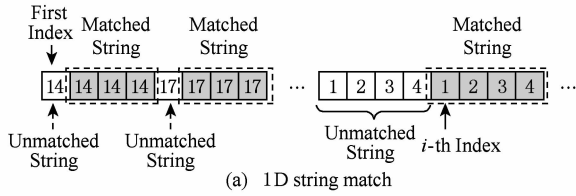


Fig. 7 Diagram of 1D string match based prediction coding^[12]

图7 基于1D串匹配的预测编码示意图^[12]

文献[19]提出了一种2重预测编码方法, 即方向预测和模板预测。其中, 方向预测如图8所示, 首先计算与待编码索引相邻的前一位置的索引值与其水平方向、对角线方向、垂直方向和反对角线方向上各索引值的欧氏距离; 再选取欧氏距离最小的索引值所在的方向作为当前位置的预测方向, 把待编码索引沿着预测方向上的索引值作为其预测。对于预测失败的元素, 则进一步采用模板预测, 其基本思想是利用统计相关性查找待编码索引的最佳预测, 如图9所示, 在模板“0020”作为上文, 下文“1”出现3次, 而“2”出现1次, 所以“1”被作为待编码索引的预测。最后, 采用CABAC和2叉树编码对2个阶段预测生成的匹配表进行压缩。由于方向预测和模板预测分别利用了索引图的局部相关和非局部相关, 索引值被准确预测的概率达到了92%。为此, 文献[24]提出了2级层次预测编码模式来发掘索引图的全局相关性: 第1级将每个与左侧相邻索引值相等的索引标识为符号“L”, 将每个与上方相邻索引值相等的索引标识为符号“U”, 然后将剩余索引标识为“O”, 得到如图10(b)所示的预测符号图; 第2级对每一行预测符号进行分组, 每组包含 $m(m=4)$ 个

相邻的预测符号。对于每个分组, 若其中的预测符号均为“L”, 则将该分组标识为“X”; 若其中的预测符号均为“U”, 则将该分组标识为“Y”; 否则将该分组标识为“Z”, 结果如图10(c)所示。最后, 对各标识符号进行熵编码。虽然文献[19,24]是发掘索引图局部与非局部相关的代表性方法, 可是这2种方法涉及多轮扫描, 计算量偏高。为此, 文献[29]对文献[19]进行了简化, 在水平和垂直方向中选取索引值变化量较小者作为预测方向, 再采用预测方向上与待编码索引直接相邻的索引作为预测值。该方法的计算量仅相当于文献[19]的20%, 但预测方向较少、未兼顾索引图的非局部相关性, 预测效率有所降低, 适合于实时要求较高的应用。

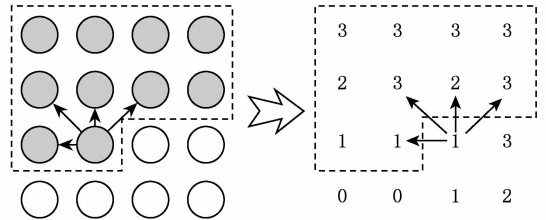


Fig. 8 Diagram of directional prediction^[19]

图8 方向预测示意图^[19]

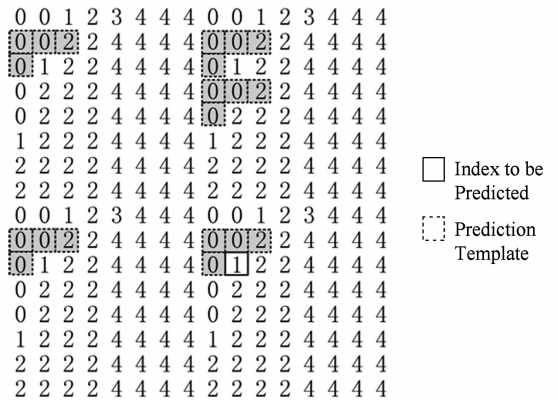
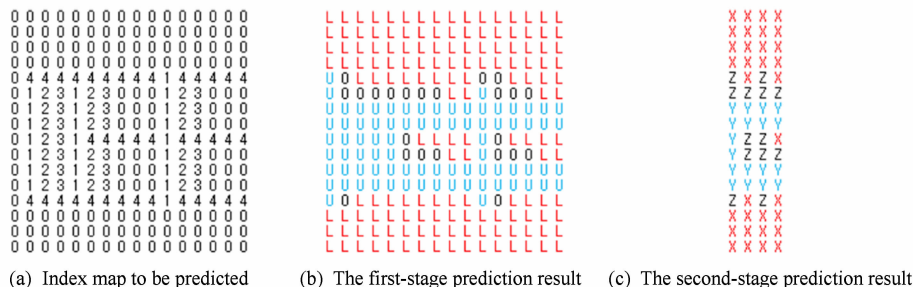


Fig. 9 Diagram of template prediction^[19]

图9 模板预测示意图^[19]

与上述方法不同, 文献[33]发掘了索引图的行、列相关性, 提出水平预测模式、垂直预测模式和逐像素预测模式3种编码模式。具体地讲, 若一个图像块的某一行(列)与其相邻的前一行(列)有相同索引值或者仅有1个索引值不同, 则用垂直(水平)预测模式编码该行(列)像素的索引值。如果不满足以上2种模式, 则采用逐像素预测模式, 即利用左侧相邻像素的索引值预测当前像素的索引值, 再将预测误差写入码流。与文献[8,11,19,29]的每次只能预测1个索引值的方法相比, 该方法每次可预测多个索引值,

Fig. 10 Diagram of two-stage prediction of index map^[24]图 10 索引图的 2 级预测示意图^[24]

其预测效率更高,所需同步信息更少,因此这一思路经过改进后被 HEVC-SCC 接纳. HEVC-SCC 为每个索引值定义 2 种预测模式:“COPY_ABOVE_MODE”和“COPY_INDEX_MODE”^[31,34]. 前者类似文献[33]的水平预测模式,后者相当于垂直预测模式,不同的是,这 2 种模式采用拷贝行程替代了水平(垂直)预测模式的整行(整列)拷贝方式,使得预测不局限于整行(整列)进行,连续预测的索引数目可少于或者多于一行(列),其编码方式更加灵活,效率也更高.

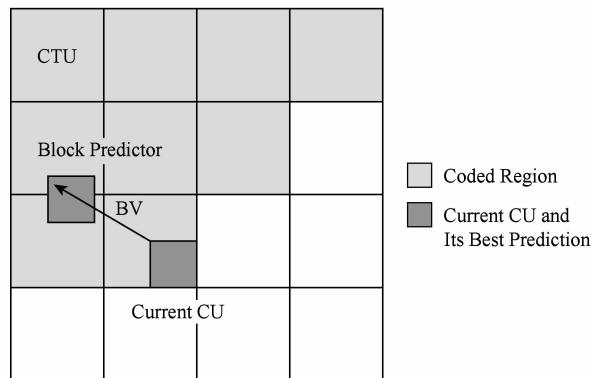
2.2 基于模板匹配的编码算法

针对屏幕图像中往往包含较多的相同或相似文字和图形的特点,研究人员提出采用模板匹配的方法来降低屏幕内容中这种非局部相关的信息冗余. 模板匹配方法是在 H. 264/AVC 广泛应用的一种帧内预测技术,其基本思想是利用待编码像素块的某个邻域的已编码像素集构成一个模板,然后利用该模板在已编码的区域中搜索与待编码块最相似的 1 个或多个块,最后用最佳匹配块^[14]或多个相似块的均值^[35]作为待编码块的预测. 由于这类方法不需向解码端传输同步信息,可有效改善编码效率,但是基于块的模板匹配的计算量偏高,且预测准确度不够稳定. 在这种情况下,文献[36]提出了一种像素粒度的模板匹配预测编码方法,思路与图 9 类似. 实验表明,该方法的预测准确率达到 80%,被准确预测的像素无需编码,只需编码预测残差非零的像素索引、位置和像素值. 文献[37-38]通过统计发现,多数非零残差对应的像素值也分布在基本色中,为了进一步提高非零残差的编码效率,该文采用全局调色板与当前编码单元调色板的差作为非零残差的调色板,对小于一定阈值的残差进行率失真优化下的修正,继而减少了残差的码率,提高了整体的编码效率. 但是,模板匹配的过程仍然非常耗时. 为此,文献[39]将 Hash 表结构引入到了模板匹配中,提出了对一个由 21 个像素组成的模板在整帧范围内进行快速

搜索的方法,并采用 LZMA (Lempel-Ziv-Markov chain algorithm) 熵编码方法压缩预测残差. 与 HEVC 扩展参考软件相比,该方法的运行时间降低了 50%左右,并且编码效率提高了 1 倍.

2.3 基于块匹配的编码算法

除了模板匹配,基于块匹配的编码算法也是一种发掘非局部数据相关性的有效手段,由文献[40]首次引入 H. 264/AVC 中,目前已被采纳为 HEVC-SCC 扩展标准中的预测模式之一,称为“帧内块拷贝”(intra block copy, IBC)^[14]. 其基本思想类似于传统的帧间运动估计,在当前帧的已编码区域内搜索与待编码单元相似的块,再将 2 个块的距离(称为“向量”)及预测误差进行编码. 如图 11 所示,阴影部分为搜索区域,其中,CTU 表示编码树单元(coding tree unit),BV 表示块向量(block vector),CU 表示编码单元(coding unit).

Fig. 11 Diagram of intra block copy method^[14]图 11 帧内块拷贝方法示意图^[14]

一方面,为了在编码效率和计算量之间进行折中,典型的帧内块拷贝方法往往采用较小的搜索范围,无法有效提取出屏幕图像存在的大范围重复图案和冗余. 于是,文献[41]提出了一种基于 2 级 Hash 的块匹配方法,首先为每个待编码块计算出 1 级 Hash 值,将与该块具有相同 1 级 Hash 值的块作为候选块;然后,选取与待编码块具有相同 2 级

Hash 值的候选块作为最佳匹配块. 根据计算量的对比分析, 该方法的搜索速度在一定条件下甚至可超过快速运动估计 TZ Search 方法, 是一种高效率的帧内块拷贝技术.

另一方面, 文献[42]认为方形块结构的预测方式对于屏幕内容中的细腻线条等任意形状、大小的图案缺乏灵活性, 提出了一种类似于 HEVC 的变块尺寸运动估计的非方块匹配模式, 改善了块匹配方法的编码效率, 但是其时间复杂度增加了约 14%. 文献[43]则进一步提出将 1 个编码单元(CU)划分为 16 个矩形的“微块”, 所允许的块最小长度(或宽度)可达到 1 个像素. 由于预测结构更加精细, 该方法的编码效率较之传统方形块结构平均提高了 4.78%, 编解码的时间复杂度则提高约 6.43%. 可见, 形状丰富灵活的块结构有利于改善屏幕内容的预测效率. 实际上, 这一结论也是基于字典的编码算法的研究出发点之一.

2.4 基于字典的编码算法

基于模板匹配和块匹配的编码算法均利用矩形区域的像素集合发掘屏幕内容中蕴含的非局部冗余, 但是文本、图表和图标等很难用固定形状的像素集合实现最佳匹配, 这样就出现了一类基于字典的编码算法. 该算法的主要思想是利用待编码像素所在的 1 个 1D 或 2D 的连续像素串作为模板, 该像素串在空间域上可组织成任意的形状, 再在已编码区域中搜索与其最匹配的像素串, 进而对待编码像素与匹配像素串的距离和匹配串长度进行编码.

2009 年, 文献[44-45]采用基于 Lempel-Ziv 字典的 gzip 算法对复合图像进行编码, 将字典编码技术引进到图像编码中. 其后, 文献[14, 46]将该方法推广到屏幕内容的编码中, 应用 gzip 或 LZMA 等提出了字典熵编码, 作为全色度无损编码器的关键技术之一, 后又经其作者将字典编码的基本执行单元从最大编码单元(largest coding unit, LCU)修改为 CU^[47-48]. 为了加快像素串的匹配速度, 文献[49]提出了 Hash 表结构的 1D 字典编码以及 2 种字典模式, 对应不同的搜索范围; 文献[50]通过实验发现像素串的最佳匹配长度大多是 3 的倍数, 进而提出一种 3B 计算 Hash 值的方法, 使得像素串在匹配过程中不再逐字节地搜索, 并将 Hash 表的存储空间减少了 2/3; 文献[51]则认为可适当降低匹配标准, 将像素串的无损匹配调整为有损匹配来提高匹配效率, 并提出一种基于拉格朗日乘子法的率失真约束的有损字典编码方法.

然而, 文献[52-54]认为采用 1D 像素串作为模板无法充分利用图像的 2D 相关性, 并且要求待匹配像素串和匹配像素串做到精确匹配也具有一定局限性. 于是, 文献[52-54]进一步提出一种屏幕内容的 2D 字典编码方法, 将待编码单元的 Hash 值作为字典索引查找到候选的匹配块, 再依据率失真函数确定最佳匹配块, 不过他们所采用的像素串仍然是规则的块结构. 文献[55]则提出一种广义的基于 2D 串拷贝的字典编码. 如图 12 所示, 连续的像素串可组成 2D 空间的任意形状, 对于具有复杂形状的文本文、图标等元素能够实现更加准确的预测, 且匹配过程既能够在编码单元间进行, 也可以在当前待编码的单元中进行, 比上述的帧内块拷贝和基于块的字典编码等方法更加灵活. 同时, 该方法还支持待匹配像素串与匹配像素串发生重叠, 通过增加匹配长度来提高编码效率.

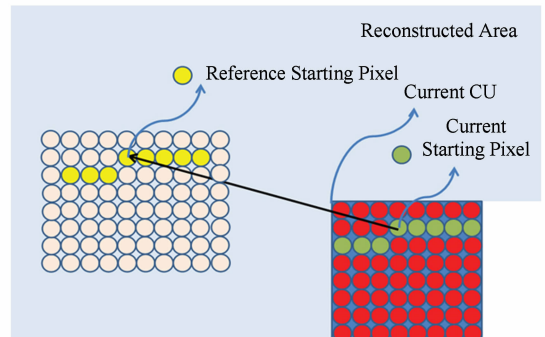


Fig. 12 Diagram of 2D string copy^[55]

图 12 2D 串拷贝示意图^[55]

由于基于模板匹配的编码算法、基于块匹配的编码算法和基于字典的编码算法均充分利用了不连续色调图像包含大量相同字符或者相同纹理结构这一非局部相关性特点, 有效提高了屏幕内容的编码效率, 3 类算法的不同之处表现在基本预测结构和预测参数的表示方式上. 帧内块拷贝模式与帧间运动估计在一定程度上可以统一起来, 在 HEVC 等视频框架下实现较其他两者更加方便, 预测参数表示也较为简单; 而基于字典的编码算法的预测参数表示却比较复杂, 参数数量也较多. 为此, 文献[56]统计分析了字典编码中位移参数的联合概率分布和参数之间的相关性, 进而优化设计了位移参数的码字分配方案, 并提出一种位移参数的联合编码方法, 是对基于字典的编码算法的一种有效改进. 另外, 上述 3 类算法的计算量都偏高, 尽管研究人员通过引进 Hash 函数的方式加速匹配计算, 可是如何在 Hash 表的存储空间、计算量和预测精度之间达到理想的折中仍需探索.

2.5 基于形状表示的编码算法

与上述的 3 种非局部搜索方法思路不同,文献[9]认为屏幕内容图像是由点、线、面、三角形和矩形等一些基本形状组成,并称之为形状原语,进而提出了形状原语提取编码方法.该方法选择了 4 种形状原语:孤立点、水平线、垂直线和矩形,任何一个复杂的形状都可以分解为这 4 种形状原语的组合,如图 13 所示.为了提取屏幕图像中的这些形状原语,从最左下角开始按从左到右、从下到上的顺序扫描图像块,如果当前像素点已包含于前一形状原语,则下一像素点成为当前像素点,继续向右、向上扫描.该算法提出的形状原语比较新颖,而且对形状原语的编码也较容易实现,特别是能比较有效地处理不连续色调图像.

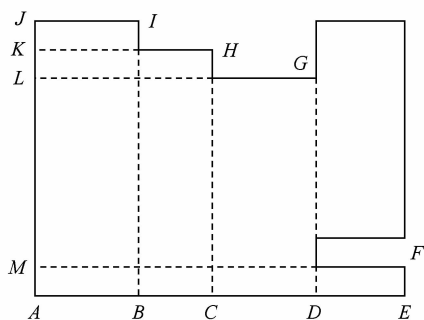


Fig. 13 Diagram of shape primitive extraction^[9]

图 13 形状原语的提取示意图^[9]

除了上述的系列性工作以外,还有一些研究者不断尝试更多的编码方法,如文献[57]在 HEVC 的帧内方向预测基础上,针对屏幕内容图像中包含大量强边缘的特点,提出一种基于梯度的边缘预测的帧内预测模式.其基本思路是利用待编码像素的 3×4 邻域内的像素计算沿着 $\angle 0^\circ$, $\angle 45^\circ$, $\angle 90^\circ$ 和 $\angle 135^\circ$ 方向的梯度值,从中选取梯度值最大方向的相邻像素预测待编码像素.文献[58]则在文献[57]基础上,进一步提出在中值预测、边缘预测和帧内方向预测 3 种模式中选取具有最优率失真性能的模式进行预测的无损编码算法.该方法不仅与 HEVC 有很好的兼容性,还比 HEVC 的帧内编码效率提高了 16.13%.

3 屏幕内容的时间域编码方法

针对第 2 个方面,典型自然视频的物体运动往往是连续的,屏幕内容的物体运动却由于处理设备和显示设备的工作特性呈现出离散的、整数像素精度的特点,文献[59-60]认为若继续采用分数像素精

度的运动估计/补偿可能导致码字的浪费,进而提出了一种自适应确定运动向量精度的方法.其基本思路是将待编码图像分成不重叠的块,利用 Hash 方法在参考图像中为每个分块搜索与其匹配的块,再根据能够准确匹配的分块比例和多个阈值选择合适的运动向量精度.该方法可比单纯采用 $1/4$ 像素精度运动向量的 HEVC 提高约 3.3% 的编码效率,目前已被 HEVC-SCC 采纳.

屏幕视频的另一个特点是存在快速全局运动^[15].在这种情况下,为了获得较高的帧间预测效率,需要在较大范围甚至整帧内展开运动估计和补偿.这既会给编码器带来较大的计算负担,又可能使那些源于运动向量中心偏置假设的快速运动估计算法(如 TZ search^[61]等)陷入局部最优.为此,出现了 2 类研究工作.

第 1 类工作侧重在运动估计中引进低位深度的像素,例如 1b 全搜索仅需简单的位操作即可实现,并且能将多个像素的匹配误差并行处理.文献[62]选取相邻像素差异最大的位平面将视频帧量化成 2 值图像,然后进行带有中止判别的 1b 全搜索,取得了与 8b 全搜索相近的时间域预测效率.文献[63]进一步通过对比实验考察了典型的低位深度运动估计对屏幕内容的适用性,发现基于加权异或匹配准则和最高 3~4 个位平面的运动估计^[64]能够获得较高的预测效率.

第 2 类工作的主要思想是借助 Hash 表提高较大搜索范围内的块(串)匹配的效率.文献[41]提出了基于 2 级 Hash 的块匹配方法,其基本思路详见 2.3 节.该方法不仅能够用于帧内块拷贝,也可有效应用在帧间预测中.与 HEVC 的校验模型相比,它能够将包含快速运动的屏幕视频的编码码率降低 59%.类似地,文献[65]也提出一种基于双缓存和 Hash 表的字典编码方法,其中主缓存相当于传统编码器的帧缓存,用于存储编码端重建的参考像素串,次级缓存用来保存最近或频繁使用的参考像素串.这样,前者可发掘屏幕内容的局部和短时非局部相关性,而后者则能够充分利用屏幕内容的长时非局部相关去除时间域冗余.故此,文献[41,65]均是较为有效的时空域预测方法.

值得注意的是,文献[15]认为屏幕内容中某些对象的快速运动很可能是由用户拖动/最大化/最小化窗口、切换界面等操作所引起的,在这种情况下,编码算法只需保证视觉上的平滑过渡,而没必要高保真地压缩运动对象的内容.于是,文献[15]采用最大后验概率将屏幕内容划分为高实用性内容块

(high utility content) 和低实用性内容块 (low utility content), 并进一步利用圆对称高斯滤波器对低实用性内容块进行模糊处理. 由于考虑了人眼视觉的时域掩蔽效应, 该方法可在保持主观解码质量的前提下, 将所需的码率降低了 24%~40%.

现有基于块平移模型的快速运动估计/补偿技术均是根据自然视频的运动向量统计特点(如中心偏置特性)所设计的. 然而从作者所掌握的文献来看, 目前鲜见有关屏幕内容的运动特性以及基于该特性的帧间预测研究. 从这个意义上讲, 屏幕内容的时间域编码方法尚有一定的改进空间.

4 屏幕内容的色度编码方法

针对第 3 个方面, 典型连续色调的图像和视频的标准格式普遍采用 4:2:0 颜色空间采样, 对不连续色调图像的采样则一般采用 4:4:4 的格式, 其 3 个色彩分量之间存在大量数据冗余, 例如在 YUV 空间, 若相邻的若干像素的亮度值相等, 那么其色度值也极可能相等.

一方面, 文献[66]研究发现, 通过一定的颜色变换可有效提高色度信息的编码效率. 为了尽量减少颜色空间采样产生的冗余, HEVC-SCC 标准采用一种自适应的色彩空间变换将 RGB 色彩空间的像素线性转换到 YC_oC_g 色彩空间^[14]:

$$\begin{pmatrix} Y \\ C_o \\ C_g \end{pmatrix} = \begin{pmatrix} 1/4 & 1/2 & 1/4 \\ 1/2 & 0 & -1/2 \\ -1/4 & 1/2 & -1/4 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix},$$

每个编码单元的预测残差自适应地在 RGB 和 YC_oC_g 中选取合适的颜色空间进行编码.

另一方面, 文献[16,67]提出一种基于混合色度采样率的双编码器联合编码算法, 由 1 个全色度信息编码器和 1 个下采样色度信息的编码器组成, 前者用于编码不连续色调的部分, 后者用于编码连续色调的部分. 同时, 不连续色调部分在全色度空间进行预测, 其残差经色度下采样后在 YUV 4:2:0 空间进行编码. 在编码一个图像块时, 采用一个率失真函数在 2 个编码器中自适应选择, 进而充分发掘了不同颜色通道间的相关性. 进一步地, 文献[68]认为 3 个色彩分量之间或者相邻像素的色彩分量之间存在一定关系, 进而提出一种颜色分量间的预测方法, 实现了用已解码像素的亮度和色度分量之间的线性关系和当前像素的亮度分量预测其色度分量.

总体来看, 有关屏幕内容图像的色度编码方法

的研究还不多, 而且缺少对屏幕内容色度分量特点的特殊考量.

5 屏幕内容的混合编码方法

屏幕内容由不连续色调部分和连续色调部分组成, 一方面, 前者包含丰富的高对比度、简单背景的字符边缘、细线条等, 人眼视觉系统对其边缘失真非常敏感; 后者则包含大量的复杂前景和背景、多个颜色种类、平滑边缘, 而视觉掩蔽效应使得人眼对于其信息损失的敏感度降低, 这表明 2 类组成部分的保真度要求不同. 另一方面, 不连续色调部分和连续色调部分的数据统计特性存在明显差异, 尚不存在某一种编码方法能够同时实现 2 类组成部分的高效率压缩. 故此, 研究人员提出采用混合编码框架, 如图 14 所示. 对屏幕内容的图像/视频进行编码, 其主要思路是首先将图像划分为不连续色调和连续色调 2 种类型的区域, 然后为不同区域选择恰当的编码算法, 最后将压缩码流进行复合.

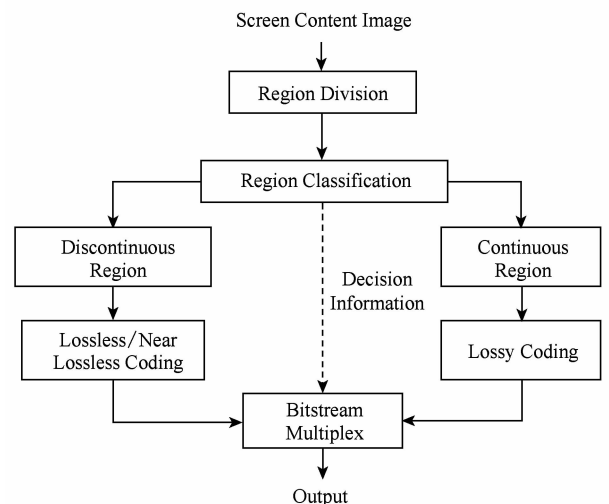


Fig. 14 General procedure of screen content image coding

图 14 屏幕内容图像编码的一般流程

5.1 屏幕内容的区域划分模式

1) 基于对象的区域划分

基于对象的划分模式将图像划分成具有一定语义的区域, 如 1 个图形或字符等. 其优点是区域划分准确; 而缺点在于对分割算法要求高, 且区域形状不规则, 需借助边信息将区域边界传输至解码端, 额外开销较大, 不便于使用现有算法进行编码, 故几乎很少采用.

2) 基于层的区域划分

该模式将每个像素划分到不同的层, 而不同的层可采用不同的压缩算法. 例如, MRC 方法^[69]把图像

划分为前景层、背景层和遮罩层. 前景层包括文本、图形或线条, 背景层包括自然图像和空白区域, 而遮罩层用来指示某一像素输出前景层的值还是背景层的值. 背景层和前景层使用 JPEG, 而遮罩层则使用 JBIG; DjVu 方法^[70]则采用基于小波变换的编码方法 (IW44) 压缩前景层和背景层, 采用 JBIG 2 算法压缩遮罩层. 基于层的区域划分简化了基于对象的划分方法, 但是尚不存在一种适用于所有图像的分层方法^[29], 并且一部分像素可能同时属于不同层, 也可能在同一层内出现不同类型的图像区域, 以致影响编码效率.

3) 基于块的区域划分

文献[17]提出将图像划分为一系列不重叠的、大小一致(如 8×8 像素、 16×16 像素)的块, 再将其分成文本块、图形块和图像块等不同的类型, 进而采用恰当的方法编码每种类型的块. 这种区域划分方法计算简单, 无块间冗余, 不需要边信息实现编解码端的同步, 且与标准的编码方法兼容. 然而, 若当某个像素块处在不同类型区域的交界时, 则可能由于区域划分的不准确、像素统计特性的不同而降低编码效率.

基于对象的区域划分和基于层的区域划分过程较为复杂且技术尚不成熟, 基于块的区域划分模式就逐渐发展为屏幕内容混合编码框架的首选方法.

5.2 基于块的屏幕内容典型混合编码框架

基于块的屏幕内容混合编码的主要思路是先将图像划分为不连续色调块和连续色调块, 再为不同类型的图像块选择合适的编码算法实现压缩. 由于本文已经在第 2~4 节详细阐述了不连续色调图像的编码方法, 本节将重点关注块类型的分类方法和各类算法的主要混合框架.

文献[17]首先提出了基于块的复合图像编码方法, 通过统计一个图像块内部颜色的数量来区分不连续色调和连续色调图像块, 并将处于 2 种类型区域交界的像素块定义为边界块. 随后, 不连续色调图像块采用无损/近无损的 JPEG-LS 标准算法编码, 连续色调图像块则采用 JPEG 标准算法编码, 而边界块则需通过设置不同的量化参数以便在有损和无损压缩区域之间实现主客观解码质量的过渡. 该方法的计算复杂度较低, 但由于对不连续色调图像块采用 JPEG-LS 无损编码, 其压缩效率比较有限. 文献[6]利用相邻像素的最大误差将像素块划分为自然图像块和文本/边缘块, 并借助率失真模型和人眼主观质量权重为不同类型的像素块计算合适的量化

步长, 进而采用基线版本的 JPEG 有损压缩完成编码, 其编码效率较之文献[17]有一定提高. 不过, 利用调节量化因子来控制边界块编码质量的手段仍无法有效保持文本内容的锐利边界.

考虑到多幅连续的屏幕内容图像之间存在沿着时间维的强相关性, 文献[71]设置了帧内和帧间 2 种编码模式. 若某个图像块在相邻帧间的变化量低于某个阈值, 则采用传统的帧间编码进行压缩; 否则, 采用帧内编码进行处理. 在帧内模式下, 如果图像块包含的颜色数量小于某个自定义阈值, 则将其作为文本\图形块进行无损压缩; 否则, 对其采用有损方法编码. 并且, 在有损编码过程中, 为了保证不同图像块具有相近的解码质量, 若连续色调图像块内相邻像素差值的熵低于某个阈值, 则认为该块包含一定的文本/图形内容(类似于文献[17]中的边界块), 采用精细量化矩阵完成有损压缩, 否则采用低质量量化矩阵进行有损压缩. 由于采用了与文献[17]相近的控制方法, 该算法对部分边界块的文本边缘保持得亦有不足.

文献[9]同样也采用了颜色数量来区别连续色调图像块和不连续色调图像块. 对于不连续色调图像块, 文献[9]采用了基于形状表示或基于调色板+索引图的编码+LZW (Lempel-Ziv-Welch) 编码, 并通过基于率失真约束的优化算法在两者之间自适应

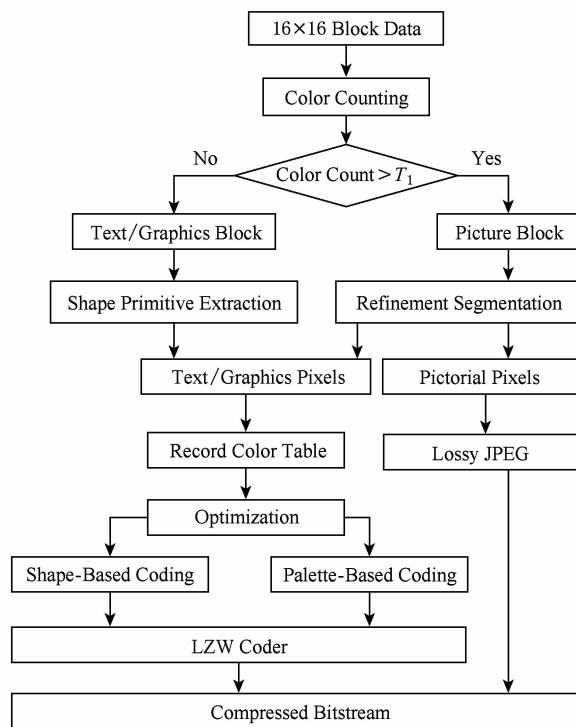


Fig. 15 General framework of SPEC algorithm^[9]

图 15 SPEC 算法的总体框架^[9]

选取. 对于连续色调图像块, 首先提取出其中的不连续色调像素并采用基于形状表示的方法进行编码, 然后用周围像素的平均值填充这些像素得到较为平滑的图像块, 再用 JPEG 进行编码(详细流程见图 15). 该算法对不连续色调图像块的压缩效果较好, 但处理包含大量孤立不连续色调像素的混合图像块时则会由于边信息过多影响编码效率, 编码复杂度也较高.

由上述的 3 种编码框架可见, 将图像简单地分为不连续色调像素块和连续色调像素块后, 虽然编码过程较为便捷, 可是对那些同时包含不连续色调像素和连续色调类型像素的混合块, 其编码效率则不令人满意. 为解决这一不足, 文献[72]利用梯度-直方图分布特性对图像块进行了更加细致的类型划分, 分别是平滑图像块、不连续色调图像块、混合图

像块和连续色调图像块. 如图 16 所示, 首先将待分类块内的每个像素划分为低梯度像素、中梯度像素和高梯度像素, 然后分析其灰度直方图, 若该块含有大量中梯度像素, 则归为连续色调图像块; 若该块含有大量低梯度像素且直方图仅有 1 个主要峰值, 则归为平滑图像块; 若该块包含大量高梯度像素并且直方图有若干主要峰值, 则归为不连续色调图像块; 若该块包含大量高梯度像素且直方图无明显多峰, 则归为混合图像块. 其中, 连续色调图像块采用 JPEG 进行编码; 平滑图像块采用标量量化和算术编码实现压缩; 不连续色调图像块采用调色板-索引图方法完成编码; 混合图像块则利用 1 级 Haar 小波和算术编码进行处理. 较之通过颜色数量来分类的方法, 该算法的分类更加准确, 对混合块的编码效率和质量更高, 编码复杂度也较为合理.

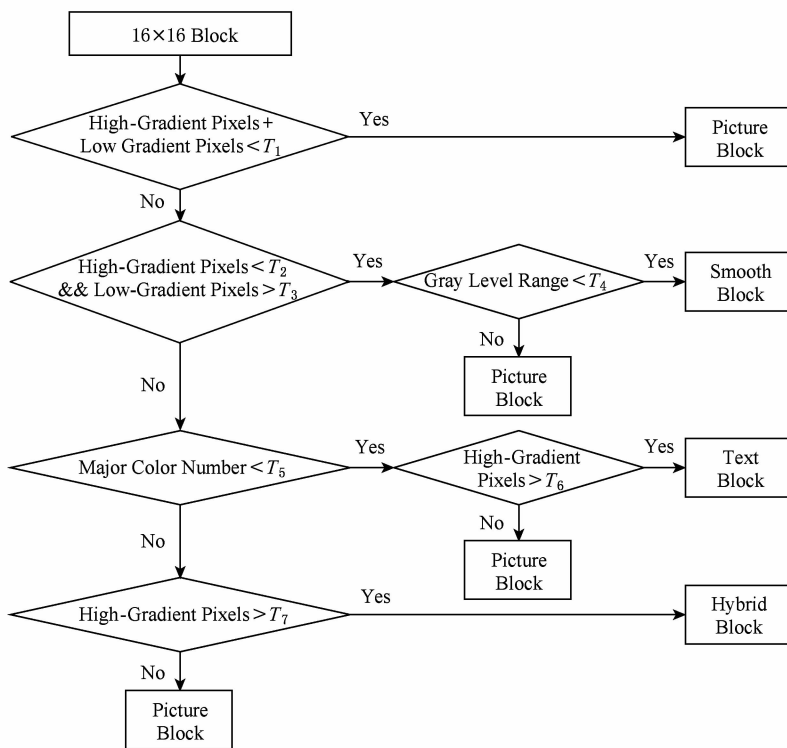


Fig. 16 Flowchart of block classification of BFC algorithm^[72]

图 16 BFC 算法的图像块分类流程图^[72]

文献[16]也采用了基于像素梯度和颜色直方图思路将像素块划分为文本块和连续色调块, 不同之处在于, 文献[16]分别采用 PNG 和 JPEG 形成 2 个独立码流: PNG 码流包含所有文本块信息, 图像块用同一颜色填充; 而 JPEG 码流则包含所有图像块数据, 文本块用同一颜色填充. 由于 PNG 支持透明区域, 在解码端重构出 2 个码流后, 只需将 PNG 图像覆盖至 JPEG 图像之上, 并使 PNG 中对应连续

色调块的位置具有透明效果即可构成混合图像. 该编码方法巧妙地利用了 PNG 算法, 使得编码器能很好地与浏览器兼容, 而且其复杂度也低于文献[72]. 后来, 其作者在文献[24]中对文献[18]的文本块编码算法进行了改进, 提出了一种 2 级层次预测的调色板-索引图编码(详见 2.1.3 节)部分, 不仅计算复杂度较之文献[18]降低了约 96%, 解码帧的峰值信噪比也平均提高了 1 dB 左右.

与上述算法均不同,文献[44-45]认为采用简单的块分类算法并不能保证所有像素块都能得到最恰当的处理,故该文不再对像素块进行分类,而是利用率失真准则在基于字典的编码 gzip 和 H.264 帧内编码算法中择优选取效率较高的方法完成压缩.由于能够取得最优的率失真性能,其后出现的大多数屏幕内容编码器(如文献[8,14,48]等)均采用了这种办法动态决定编码模式.虽然该算法获得了较好的压缩性能,但编码复杂度较高,尚需进一步简化其编码模式的选取.

5.3 HEVC-SCC 编码国际标准进展

基于块的混合编码框架已被 JCT-VC 制定的 HEVC-SCC 草案所采纳,本节简要介绍标准的主要进展,其详细情况可参见文献[14].

HEVC-SCC 标准的制定开始于 2014 年,JCT-VC 公开征集屏幕内容编码方案^[73],并于同年公布了第 1 版和第 2 版草案^[74],至今已经发布第 6 版草案^[75].该草案在最新一代视频编码标准 HEVC^[76]及其扩展标准 HEVC-RExt^[77]的基础上,引进了帧内块拷贝、调色板编码、自适应色彩空间变换、自适应运动向量分辨率等多项新技术(本文第 2~4 节已经详细阐述这些方法的主要思想),并陆续发布了 SCM-1.0 至 SCM-8.3 等若干版本的校验模型软件,其每个版本的改进情况可详见文献[78].根据 2016 年 2 月公布的一项对比实验结果显示,SCM-6.0 针对屏幕内容的压缩效率较之 AVC 高 4:4:4 类校验模型 JM-19.0 平均高出 81% 以上^[79],表明 HEVC-SCC 达到了较高的编码性能.2017 年,HEVC-SCC 已作为扩展内容正式加入到 HEVC 标准中,但距离成为国际标准还需进一步的工作.

6 屏幕内容编码研究展望

目前,屏幕内容编码方兴未艾,是图像和视频编码领域的热点研究领域之一,其基本方法和技术虽然日臻完善,但尚未完备,许多问题还处于不断探索中.鉴于此,我们认为未来屏幕内容编码将有望在以下 3 个方面取得进展:

1) 建立多种编码方法的标准框架及其快速决策方案

由于屏幕内容的统计多样性,混合编码方案的有效性已经为研究人员所认可,但是其混合编码框架尚未统一.基于块分类的混合编码受制于分类的精确性,基于率失真决策的编码方法则又具有较高

的复杂度.而由于屏幕内容编码主要面向实时性要求较高的应用场合,编码模式的快速决策就显得格外重要.故此,兼顾编码效率和计算复杂度的自适应混合编码统一构架需要深入研究和解决.

2) 建模屏幕内容的空频域统计规律

众所周知,自然图像和视频编码效率的不断提高应归功于人们对其空间域和频率域统计规律,尤其是频率域统计规律认知程度的逐渐深入.然而,现有文献(如文献[8,11-13,80]等)对屏幕内容数据分布规律的研究尚不够深入,相关报道也较少.而且,目前几乎所有编码算法都仅利用了屏幕内容图像的空间域相关性,包括局部相关性、非局部相关性和局部方向相关性^[81],其编码效率仍有很大的提升空间.若能找到一种有效的数学变换及其可资利用的系数分布规律,则有望进一步改善屏幕内容的压缩比.

3) 探索新型图像质量评价方法和编码方法

现有编码框架的率失真模型大多追求信号意义下的保真度,而峰值信噪比和均方误差等函数并不能很好地度量解码图像的人眼主观质量.关于结构相似度(structural similarity, SSIM)指标的研究表明^[82],人眼视觉系统(human vision system, HVS)对于图像中的结构差异较之亮度差异更加敏感.故此,合理利用 HVS 的特点,建立符合人眼主观评价的屏幕图像质量评价方法可在保持主观解码质量的前提下有效提高其编码效率^[83-85].另外,目前所提出的基于调色板-索引图的编码、基于模板匹配的编码、基于块匹配的编码和基于字典的编码等方法大多是在对屏幕内容数据特点的初步认识基础上,受启发自经典的编码技术.相信新型图像质量评价方法(如文献[83-85])和更准确的空频域统计能为屏幕内容编码指明一个改进方向,乃至促进屏幕内容编码新思路的涌现.

7 结束语

云计算时代下,屏幕内容的快速编码和传输得到了学术界和工业界的广泛关注.本文重点阐述了屏幕内容编码方法的研究进展:首先,从屏幕内容图像的数据统计特性切入,以屏幕内容编码方法的演进为主线,将现有方法分为基于调色板-索引图的编码算法、基于模板匹配的编码算法、基于块匹配的编码算法、基于字典的编码算法、基于形状表示的编码算法、时间域编码方法和色度分量编码方法 7 类.一方面,经过对各类方法发展脉络及其基本思想的详

细梳理和比较,本文讨论了不同算法的优势、不足和适用范围;另一方面,发现目前屏幕内容的混合编码框架尚不成熟,屏幕内容的空频域统计规律需进一步明确,而且仍待发掘新型图像质量评价方法及其适用的编码方法.据此,对屏幕内容编码方法的未来发展进行了展望.

参 考 文 献

- [1] de Queiroz R L, Buckley R R, Xu Ming. Mixed raster content (MRC) model for compound image compression [C] //Proc of SPIE VCIP'99. Bellingham, WA: SPIE, 1999; 1106-1117
- [2] Lu Yan, Li Shipeng, Shen Huifeng. Virtualized screen: A third element for cloud-mobile convergence [J]. IEEE Multimedia, 2011, 18(2): 4-11
- [3] JCT-VC. Use cases and requirements for lossless and screen content coding, JCTVC-M0172 [R]. Geneva, Switzerland: ITU-T, 2013
- [4] Konstantinide K, Tretter D. A method for variable quantization in JPEG for improved text quality in compound documents [C] //Proc of IEEE ICIP'98. Piscataway, NJ: IEEE, 1998; 565-568
- [5] Konstantinide K, Tretter D. A JPEG variable quantization method for compound documents [J]. IEEE Trans on Image Processing, 2000, 9(7): 1282-1287
- [6] Ramos M, de Queiroz R. Classified JPEG coding of mixed documents [J]. IEEE Trans on Image Processing, 2000, 9(4): 716-720
- [7] Zaghetto A, de Queiroz R L. Segmentation-driven compound document coding based on H. 264/AVC-intra [J]. IEEE Trans on Image Processing, 2007, 16(7): 1755-1760
- [8] Lan Cuiling, Shi Guangming, Wu Feng. Compress compound images in H. 264/MPGE-4 AVC by exploiting spatial correlation [J]. IEEE Trans on Image Processing, 2010, 19(4): 946-957
- [9] Lin Tao, Hao Pengwei. Compound image compression for real-time computer screen image transmission [J]. IEEE Trans on Image Processing, 2005, 14(8): 993-1005
- [10] Peng W H, Xu Jizheng, Ostermann J, et al. Call for papers: Screen content video coding and applications [EB/OL]. [2016-08-15]. http://jetcas.polito.it/CFP-Screen_Content_Video_Coding_and_Applications.pdf
- [11] Ding Wenpeng, Lu Yan, Wu Feng. Enable efficient compound image compression in H. 264/AVC intra coding [C] //Proc of IEEE ICIP'07. Piscataway, NJ: IEEE, 2007; II-337-II-340
- [12] Ma Zhan, Wang Wei, Xu Meng, et al. Advanced screen content coding using color table and index map [J]. IEEE Trans on Image Processing, 2014, 23(10): 4399-4412
- [13] Mrak M, Grgic S, Grgic M. Picture quality measures in compression systems [C] //Proc of IEEE EUROCON'03. Piscataway, NJ: IEEE, 2003; 233-236
- [14] Xu Jizheng, Joshi R, Cohen R A. Overview of the emerging HEVC screen content coding extension [J]. IEEE Trans on Circuits Systems for Video Technology, 2016, 26(1): 50-62
- [15] Wang Shiqi, Zhang Xinfeng, Liu Xianming, et al. Utility-driven adaptive preprocessing for screen content video compression [J]. IEEE Trans on Multimedia, 2017, 19(3): 660-667
- [16] Lin Tao, Zhang Peijun, Wang Shuhui, et al. Mixed chroma sampling-rate high efficiency video coding for full-chroma screen content [J]. IEEE Trans on Circuits Systems for Video Technology, 2013, 23(1): 173-185
- [17] Said A, Drukarev A. Simplified segmentation for compound image compression [C] //Proc of IEEE ICIP'99. Piscataway, NJ: IEEE, 1999; 229-233
- [18] Pan Zhaotai, Shen Huifeng, Lu Yan, et al. Browser-friendly hybrid codec for compound image compression [C] //Proc of IEEE ISCAS'11. Piscataway, NJ: IEEE, 2011; 101-104
- [19] Zhu Weijia, Ding Wenpeng, Xiong Ruiqin. Compound image compression by multi-stage prediction [C] //Proc of VCIP'12. Piscataway, NJ: IEEE, 2012; 1-6
- [20] JCT-VC. Screen content coding with multi-stage base color and index map representation, JCTVC-M0330 [R]. Geneva, Switzerland; ITU-T, 2013
- [21] JCT-VC. Palette mode for screen content coding, JCTVC-M0323 [R]. Geneva, Switzerland; ITU-T, 2013
- [22] JCT-VC. AHG10: Simplification of palette based coding, JCTVC-Q0047 [R]. Geneva, Switzerland; ITU-T, 2014
- [23] JCT-VC. Non-RCE3: Modified palette mode for screen content coding, JCTVC-N0249 [R]. Geneva, Switzerland; ITU-T, 2013
- [24] Pan Zhaotai, Shen Huifeng, Lu Yan, et al. A low-complexity screen compression scheme for interactive screen sharing [J]. IEEE Trans on Circuits and Systems for Video Technology, 2013, 23(6): 949-960
- [25] Zhu Wenjing, Au O C, Dai Wei, et al. Palette-based compound image compression in HEVC by exploiting non-local spatial correlation [C] //Proc of IEEE ICASSP'14. Piscataway, NJ: IEEE, 2014; 7348-7352
- [26] JCT-VC. Screen content coding test model 2 encoder description (SCM 2), JCTVC-R1014 [R]. Geneva, Switzerland; ITU-T, 2014
- [27] Zeng Wenjun, Li Jin, Lei Shawmin. An efficient color re-indexing scheme for palette-based compression [C] //Proc of IEEE ICIP'00. Piscataway, NJ: IEEE, 2000; 476-479
- [28] Shen Huifeng, Lu Yan, Wu Feng, et al. Low-cost realtime screen sharing to multiple clients [C] //Proc of IEEE ICME'10. Piscataway, NJ: IEEE, 2010; 980-985
- [29] Zhu Weijia, Ding Wenpeng, Xu Jizheng, et al. Screen content coding based on HEVC framework [J]. IEEE Trans on Multimedia, 2014, 16(5): 1316-1326

- [30] Xiu Xiaoyu, He Yuwen, Joshi R, et al. Palette-based coding in the screen content coding extension of the HEVC standard [C] //Proc of IEEE DCC'15. Piscataway, NJ: IEEE, 2015; 253-262
- [31] Cherigui S, Guillemot C, Thoreau D, et al. Correspondence map-aided neighbor embedding for image intra prediction [J]. IEEE Trans on Image Processing, 2013, 22(3): 1161-1174
- [32] Xu Yiling, Huang Wei, Wang Wei, et al. 2-D index map coding for HEVC screen content compression [C] //Proc of IEEE DCC'15. Piscataway, NJ: IEEE, 2015; 263-272
- [33] Shen Huifeng, Lu Yan, Wu Feng, et al. Low-cost realtime screen sharing to multiple clients [C] //Proc of IEEE ICME'10. Piscataway, NJ: IEEE, 2010; 980-985
- [34] Tan T K, Boon C S, Suzuki Y. Intra prediction by template matching [C] //Proc of IEEE ICIP'06. Piscataway, NJ: IEEE, 2006; 1693-1696
- [35] Tan T K, Boon C S, Suzuki Y. Intra prediction by averaged template matching predictors [C] //Proc of IEEE CCNC'07. Piscataway, NJ: IEEE, 2007; 405-409
- [36] Tao Pin, Feng Lixin, Song Sichao, et al. Improvement of re-sample template matching for lossless screen content video [C] //Proc of IEEE ICME'05. Piscataway, NJ: IEEE, 2015; 1-6
- [37] Wang Zheng, Tao Pin, Feng Lixin, et al. Palette improvement for template matching intra coding [C] //Proc of CCF HHME'15. Beijing, China Machine Press, 2015; 186-192 (in Chinese)
(王正, 陶品, 冯立新, 等. 基于模板匹配的调色板方法[C] //第11届全国和谐人机环境联合会会议论文集. 北京: 机械工业出版社, 2015; 186-192)
- [38] Wang Zheng, Tao Pin, Feng Lixin, et al. Palette improvement for template matching intra coding [J]. Journal of Computer-Aided Design & Computer Graphics, 2016, 28(7): 1146-1151 (in Chinese)
(王正, 陶品, 冯立新, 等. 基于模板匹配的调色板方法[J]. 计算机辅助设计与图形学学报, 2016, 28(7): 1146-1151)
- [39] Feng Lixin, Tao Pin, Wen Jiangtao, et al. Lossless intra coding on whole frame for screen content video based on template matching [J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(7): 1353-1358 (in Chinese)
(冯立新, 陶品, 温江涛, 等. 基于模板匹配整帧屏幕视频帧内无损编码方法[J]. 北京航空航天大学学报, 2015, 41(7): 1353-1358)
- [40] JCT-VC. New intra prediction using intra-macroblock motion compensation, JVT-C151 [R]. Geneva, Switzerland; ITU-T, 2002
- [41] Zhu Weijia, Ding Wenpeng, Xu Jizheng, et al. Hash-based block matching for screen content coding [J]. IEEE Trans on Multimedia, 2015, 17(7): 935-944
- [42] Chen Chunchi, Xu Xiaozhong, Liao Ruling, et al. Screen content coding using non-square intra block copy for HEVC [C] //Proc of IEEE ICME'14. Piscataway, NJ: IEEE, 2014; 1-6
- [43] Zhao Liping, Lin Tao, Gong Xunwei, et al. Intra mini-block copy algorithm for screen content coding [J]. Journal of Computer Applications, 2016, 26(7): 1938-1943, 1980 (in Chinese)
(赵利平, 林涛, 龚迅炜, 等. 帧内微块复制的屏幕图像编码算法[J]. 计算机应用, 2016, 26(7): 1938-1943, 1980)
- [44] Wang Shuhui, Lin Tao. A unified LZ and hybrid coding for compound image partial-lossless compression [C] //Proc of CISP'09. Piscataway, NJ: IEEE, 2009; 1-5
- [45] Wang Shuhui, Lin Tao. Compound image compression based on unified LZ and hybrid coding [J]. IET Image Processing, 2013, 7(5): 484-499
- [46] JCT-VC. AHG7: Full-chroma (YUV444) dictionary + hybrid dual-coder extension of HEVC, JCTVC-K0133 [R]. Geneva, Switzerland; ITU-T, 2012
- [47] JCT-VC. Improvements on 1D dictionary coding, JCTVC-Q0124 [R]. Geneva, Switzerland; ITU-T, 2014
- [48] Chen Xianyi, Zhao Liping, Lin Tao. A new HEVC intra mode for screen content coding [J]. Journal of Electronics & Information Technology, 2015, 37(11): 2685-2690 (in Chinese)
(陈先义, 赵利平, 林涛. 一种新的用于屏幕图像编码的 HEVC 帧内模式[J]. 电子与信息学报, 2015, 37(11): 2685-2690)
- [49] Li Bin, Xu Jizheng, Wu Feng. 1-D dictionary mode for screen content coding [C] //Proc of IEEE VCIP'14. Piscataway, NJ: IEEE, 2014; 189-192
- [50] Jin Xiaojuan, Zhang Peijun, Lin Tao. Optimization algorithm on hash table based on HEVC screen content coding [J]. Computer Engineering and Applications, 2014, 50(17): 155-159 (in Chinese)
(金小娟, 张培君, 林涛. 基于 HEVC 屏幕图像编码的哈希表的优化算法[J]. 计算机工程与应用, 2014, 50(17): 155-159)
- [51] Zhang Peijun, Jin Xiaojuan, Wang Shuhui, et al. Screen content coding of combining full-chroma HEVC and lossy matching dictionary coder [J]. Computer Science, 2014, 41(3): 286-292 (in Chinese)
(张培君, 金小娟, 王淑慧, 等. 结合全色度 HEVC 和有损字典算法的屏幕图像编码[J]. 计算机科学, 2014, 41(3): 286-292)
- [52] Zhu Weijia, Ding Wenpeng, Xu Jizheng, et al. 2-D dictionary based video coding for screen contents [C] //Proc of IEEE DCC'14. Piscataway, NJ: IEEE, 2014; 43-52
- [53] JCT-VC. Screen content coding using 2-D dictionary mode, JCTVC-O0357 [R]. Geneva, Switzerland; ITU-T, 2013
- [54] Chen Xianyi, Zhao Liping, Chen Zhizhong, et al. 2D intra string copy for screen content coding [J]. Journal of Computer Applications, 2015, 35(9): 2640-2647, 2677 (in Chinese)
(陈先义, 赵利平, 陈治中, 等. 二维帧内串匹配屏幕图像编码算法[J]. 计算机应用, 2015, 35(9): 2640-2647, 2677)

- [55] Zou Feng, Chen Ying, Karczewicz M, et al. Hash based intra string copy for HEVC based screen content coding [C] //Proc of IEEE ICMEW'15. Piscataway, NJ: IEEE, 2015: 1-4
- [56] Zhao Liping, Lin Tao, Zhou Kailun. An efficient ISC offset parameter coding algorithm in screen content coding [J/OL]. Chinese Journal of Computers, 2016 [2016-12-07]. <http://www.cnki.net/kcms/detail/11.1826.tp.20160328.1604.002.html> (in Chinese)
(赵利平, 林涛, 周开伦. 屏幕图像压缩中申复制位移参数的高效编码算法[J/OL]. 计算机学报, 2016 [2016-12-07]. <http://www.cnki.net/kcms/detail/11.1826.tp.20160328.1604.002.html>)
- [57] Sanchez V. Sample-based edge prediction based on gradients for lossless screen content coding in HEVC [C] //Proc of PCS'15. Piscataway, NJ: IEEE, 2015: 134-138
- [58] Sanchez V. Lossless screen content coding in HEVC based on sample-wise median and edge prediction [C] //Proc of IEEE ICIP'15. Piscataway, NJ: IEEE, 2015: 4604-4608
- [59] JCT-VC. Motion vector resolution control for screen content coding, JCTVC-P0277 [R]. Geneva, Switzerland: ITU-T, 2014
- [60] JCT-VC. Adaptive motion vector resolution for screen content, JCTVC-S0085_r1 [R]. Geneva, Switzerland: ITU-T, 2014
- [61] Purnachand N, Alves L N, Navarro A. Improvements to TZ search motion estimation algorithm for multiview video coding [C] //Proc of IEEE IWSSIP'12. Piscataway, NJ: IEEE, 2012: 388-391
- [62] Sun Ting, Wan Pengfei, Au O C, et al. Fast binary motion estimation for screen content video coding [C] //Proc of IEEE APSIPA'14. Piscataway, NJ: IEEE, 2014: 1-5
- [63] Duvar R, Urhan O. Performance evaluation of low bit-dept based motion estimation approaches on screen content video [C] //Proc of IEEE Signal Processing and Communication Application Conf (SIU). Piscataway, NJ: IEEE, 2016: 2141-2144
- [64] Celebi A, Akbulut O, Urhan O, et al. Truncated gray-coded bit-plane matching based motion estimation and its hardware architecture [J]. IEEE Trans on Consumer Electronics, 2009, 55(3): 1530-1536
- [65] Zhao Liping, Lin Tao, Zhou Kailun, et al. Pseudo 2D string matching technique for high efficiency screen content coding [J]. IEEE Trans on Multimedia, 2016, 18(3): 339-350
- [66] Marpe D, Kirchhoffer H, George V, et al. Macrobloc-adaptive residual color space transforms for 4:4:4 video coding [C] //Proc of IEEE ICIP'06. Piscataway, NJ: IEEE, 2006: 3157-3160
- [67] Zhang Peijun, Wang Shuhui, Zhou Kailun, et al. Screen content coding by combined full-chroma LZMA and subsampled-chroma HEVC [J]. Journal of Electronics & Information Technology, 2013, 35(1): 196-202 (in Chinese)
(张培君, 王淑慧, 周开伦, 等. 融合全色度 LZMA 与色度子采样 HEVC 的屏幕图像编码[J]. 电子与信息学报, 2013, 35(1): 196-202)
- [68] Zhang Xingyu, Gisquet C, Francois E, et al. Chroma intra prediction based on inter-channel correlation for HEVC [J]. IEEE Trans on Image Processing, 2014, 23(1): 274-286
- [69] ITU-T. Recommendation T.44 Mixed raster content (MRC) [S]. Geneva, Switzerland: ITU-T, 1999
- [70] Haffner P, Bottou L, Howard P G, et al. High quality document image compression with DjVu [J]. Journal of Electronic Imaging, 1998, 7(3): 410-425
- [71] Said A. Compression of compound images and video for enabling rich media in embedded systems [C] //Proc of SPIE VCIP'04. Bellingham, WA: SPIE, 2004: 69-82
- [72] Ding Wenpeng, Liu Dong, He Yuwen, et al. Block-based fast compression for compound images [C] //Proc of IEEE ICME'06. Piscataway, NJ: IEEE, 2006: 809-812
- [73] JCT-VC. Joint call for proposals for coding of screen content, MPEG2014/N14175 [R]. Geneva, Switzerland: ITU-T, 2014
- [74] JCT-VC. HEVC screen content coding draft text 2, JCTVC-S1005 [R]. Geneva, Switzerland: ITU-T, 2014
- [75] JCT-VC. HEVC screen content coding draft text 6, JCTVC-W1005 [R]. Geneva, Switzerland: ITU-T, 2016
- [76] Sullivan G J, Ohm J, Han W J, et al. Overview of the high efficiency video coding (HEVC) standard [J]. IEEE Trans on Circuits Systems for Video Technology, 2012, 22(12): 1649-1668
- [77] JCT-VC. Edition 2 draft text of high efficiency video coding (HEVC), including format range (RExt), scalability (SHVC), and multi-view (MV-HEVC) extensions, JCTVC-R1013 [R]. Geneva, Switzerland: ITU-T, 2014
- [78] JCT-VC. HEVC screen content coding test model 7 (SCM 7), JCTVC-W1014 [R]. Geneva, Switzerland: ITU-T, 2016
- [79] JCT-VC. JCT-VC AHG report: SCC coding performance analysis (AHG6), JCTVC-W0006 [R]. Geneva, Switzerland: ITU-T, 2016
- [80] Chen Cheng, Han Jingning, Xu Yaowu, et al. A staircase transform coding scheme for screen content video coding [C] //Proc of IEEE ICIP'16. Piscataway, NJ: IEEE, 2016: 2365-2369
- [81] Chen Guisheng, Song Chuanming, Wang Xianghai, et al. Fast prediction algorithm of index maps for screen image coding [J]. Journal of Image and Graphics, 2016, 21(9): 1127-1137 (in Chinese)
(陈规胜, 宋传鸣, 王相海, 等. 用于屏幕图像编码的索引图快速预测算法[J]. 中国图象图形学报, 2016, 21(9): 1127-1137)
- [82] Wang Zhou, 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

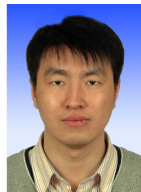
- [83] Gu Ke, Wang Shiqi, Yang Huan, et al. Saliency-guided quality assessment of screen content images [J]. *IEEE Trans on Multimedia*, 2016, 18(6): 1098-1110
- [84] Wang Shiqi, Ma Lin, Fang Yuming, et al. Just noticeable difference estimation for screen content images [J]. *IEEE Trans on Image Processing*, 2016, 25(8): 3838-3851
- [85] Ni Zhangkai, Ma Lin, Zeng Huanqiang, et al. Gradient direction for screen content image quality assessment [J]. *IEEE Signal Processing Letters*, 2016, 23(10): 1394-1398



Liu Dan, born in 1981. Lecturer of the School of Computer and Information Technology of Liaoning Normal University. PhD candidate in communication engineering at the School of Information and Communication Engineering of Dalian University of Technology. Her main research interests include image & video coding, and computer vision.



Chen Guisheng, born in 1985. PhD candidate in computer applied technology from Jilin University. His main research interests include screen content video coding (cgs-10@163.com).



Song Chuanming, born in 1980. Associate professor of the School of Computer and Information Technology of Liaoning Normal University. Received his PhD degree at the Department of Computer Science & Technology of Nanjing University. Member of CCF. His main research interests include image and video coding, and digital watermarking of multimedia.



He Xing, born in 1990. Master in educational technology of the School of Computer and Information Technology of Liaoning Normal University. His main research interests include educational video coding.



Wang Xianghai, born 1965. Professor and PhD supervisor of the School of Computer and Information Technology of Liaoning Normal University. Senior member of CCF. His main research interests include computer graphics and multimedia information processing (xhwang@lnnu.edu.cn).