

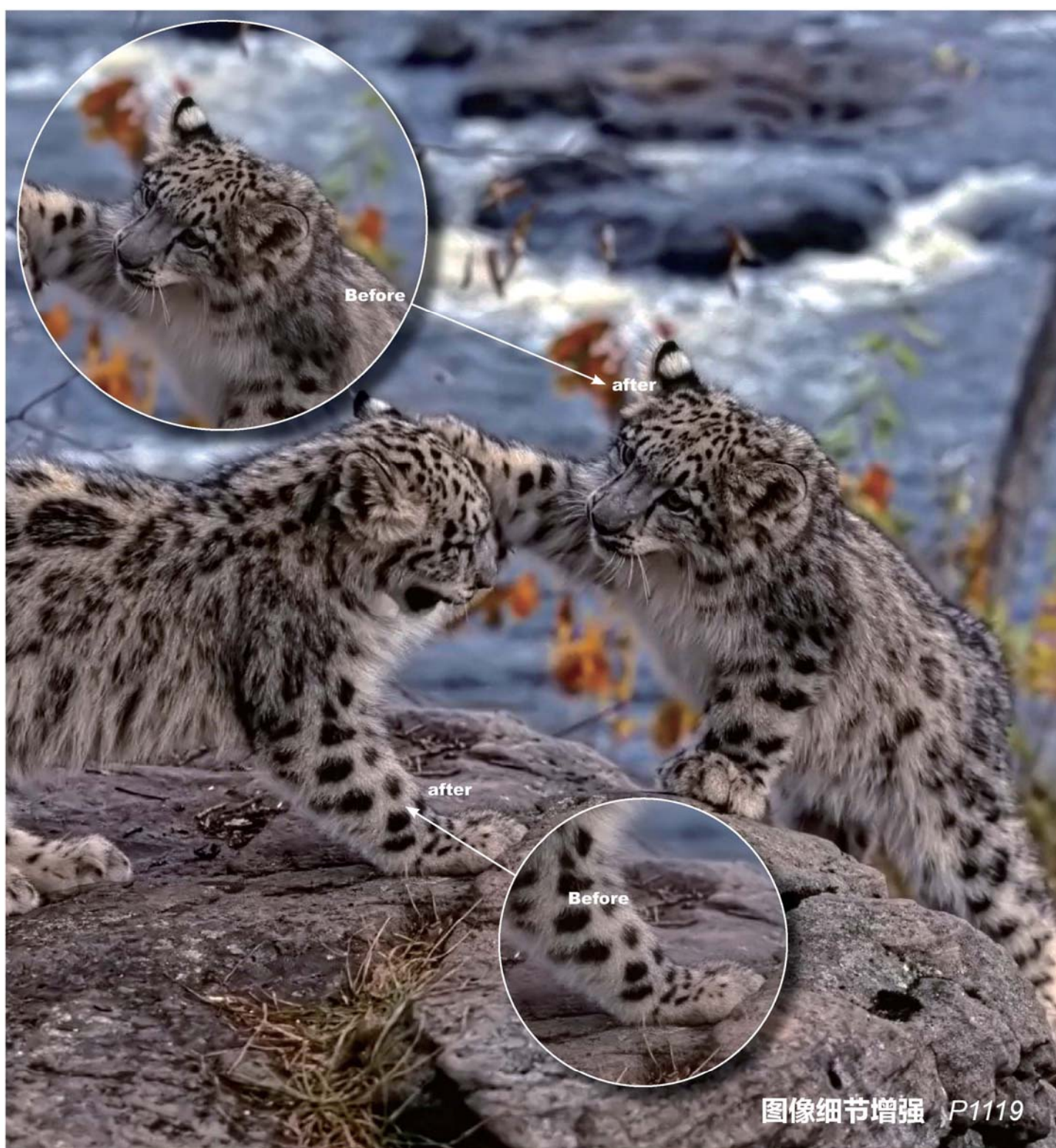


主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2016
09
VOL.21

ISSN1006-8961
CN11-3758/TB



图像细节增强 P1119



第21卷第9期（总第245期）
2016年9月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院
主办单位 中国科学院遥感与数字地球研究所
中国图象图形学会
北京应用物理与计算数学研究所

主 编 顾行发
编辑出版《中国图象图形学报》编辑出版委员会
邮政信箱 北京9718信箱
邮 编 100101
电子信箱 jig@radi.ac.cn
电 话 010-64807995
网 址 www.cjig.cn

广告经营许可证 京朝工商广字第0361号
总 发 行 北京报刊发行局
订 购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
(邮政信箱: 北京399信箱 邮编: 100048)
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

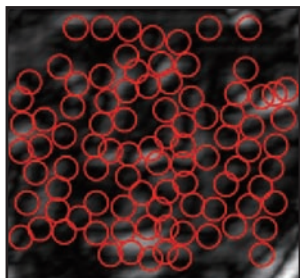
Title inscription: Song Jian Monthly, Started in 1996

Superintended by Chinese Academy of Sciences
Sponsored by Institute of Remote Sensing and Digital Earth, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

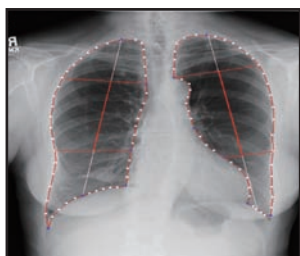
Editor-in-Chief Gu Xingfa
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics
P.O.Box 9718, Beijing, P.R.China
Zip code 100101
E-mail jig@radi.ac.cn
Telephone 010-64807995
Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ
国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



基于频率簇模型的人脸识别
(第1166页)



胸部解剖结构回归模型的虚拟双能量X线减影方法(第1247页)



参考1维光谱差异的区域生长种子点选取方法(第1256页)

图像处理和编码

融合梯度信息的改进引导滤波

谢伟, 周玉钦, 游敏 1119

用于屏幕图像编码的索引图快速预测算法

陈规胜, 宋传鸣, 王相海, 刘丹 1127

图像分析和识别

哈希编码结合空间金字塔的图像分类

彭天强, 栗芳 1138

自上而下注意力分割的细粒度图像分类

冯语姝, 王子磊 1147

结合类内和类间距离的可能聚类分割算法

刘璐, 吴成茂 1155

基于频率簇模型的人脸识别

袁姮, 王志宏, 姜文涛 1166

图像理解和计算机视觉

不同池化模型的卷积神经网络学习性能研究

刘万军, 梁雪剑, 曲海成 1178

惰性随机游走视觉显著性检测算法

李波, 卢春园, 金连宝, 冷成财 1191

自适应邻域相关性的背景建模

万剑, 洪明坚, 赵晨丘 1202

实时鲁棒的特征点匹配算法

陈天华, 王福龙 1213

暗通道先验图像去雾的大气光校验和光晕消除

赵锦威, 沈逸云, 刘春晓, 欧阳毅 1221

计算机图形学

定点容量限制质心Power图生成

郑利平, 郜文灿, 李尚林, 曹力 1229

医学图像处理

多相期增强CT中肝内血管的自动分割

袁戎, 石姝玥, 谢庆国 1238

胸部解剖结构回归模型的虚拟双能量X线减影方法

陈胜, 张茗屋 1247

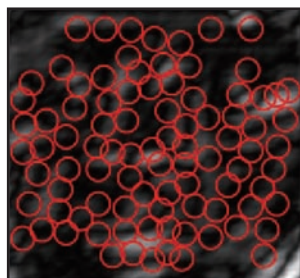
遥感图像处理

参考1维光谱差异的区域生长种子点选取方法

李修霞, 荆林海, 李慧, 唐韵玮, 戈文艳 1256

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Face recognition method based on frequency cluster (P1166)



Virtual dual-energy subtraction method for X-ray radiographs by using regression model based on chest anatomical structure(P1247)



Seed extraction method for seeded region growing based on one-dimensional spectral differences(P1256)

Image Processing and Coding

Improved guided image filtering integrated with gradient information

Xie Wei, Zhou Yuqin, You Min 1119

Fast prediction algorithm of index maps for screen image coding

Chen Guisheng, Song Chuanming, Wang Xianghai, Liu Dan 1127

Image Analysis and Recognition

Image classification algorithm based on hash codes and space pyramid

Peng Tianqiang, Li Fang 1138

Fine-grained image categorization with segmentation based on top-down attention map

Feng Yushan, Wang Zilei 1147

Possibilistic clustering segmentation algorithm based on intra-class and inter-class distance

Liu Lu, Wu Chengmao 1155

Face recognition method based on frequency cluster

Yuan Heng, Wang Zhihong, Jiang Wentao 1166

Image Understanding and Computer Vision

Learning performance of convolutional neural networks with different pooling models

Liu Wanjuan, Liang Xuejian, Qu Haicheng 1178

Saliency detection based on lazy random walk

Li Bo, Lu Chunyuan, Jin Lianbao, Leng Chengcai 1191

Background modeling based on adaptive neighborhood correlation

Wan Jian, Hong Mingjian, Zhao Chenqiu 1202

Real-time robust feature-point matching algorithm

Chen Tianhua, Wang Fulong 1213

Dark channel prior-based image dehazing with atmospheric light validation and halo elimination

Zhao Jinwei, Shen Yiyun, Liu Chunxiao, Ouyang Yi 1221

Computer Graphics

Generation method for a centroidal capacity constrained Power diagram with fixed sites

Zheng Liping, Gao Wencan, Li Shanglin, Cao Li 1229

Medical Image Processing

Automatic hepatic vessel segmentation algorithm in multi-phase contrast-enhanced CT

Yuan Rong, Shi Shuyue, Xie Qingguo 1238

Virtual dual-energy subtraction method for X-ray radiographs by using regression model based on chest anatomical structure

Chen Sheng, Zhang Mingwu 1247

Remote Sensing Image Processing

Seed extraction method for seeded region growing based on one-dimensional spectral differences

Li Xiuxia, Jing Linhai, Li Hui, Tang Yunwei, Ge Wenyan 1256

中图法分类号: TP301.6 文献标识码: A 文章编号: 1006-8961(2016)09-1127-11

论文引用格式: Chen G S, Song C M, Wang X H, Liu D. Fast prediction algorithm of index maps for screen image coding[J]. Journal of Image and Graphics, 2016, 21(9): 1127-1137. [陈规胜, 宋传鸣, 王相海, 刘丹. 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.] [DOI:10.11834/jig.20160902]

用于屏幕图像编码的索引图快速预测算法

陈规胜¹, 宋传鸣¹, 王相海¹, 刘丹^{1,2}

1. 辽宁师范大学计算机与信息技术学院, 大连 116029; 2. 大连理工大学信息与通信工程学院, 大连 116024

摘要: 目的 屏幕图像编码要求压缩效率高、实时性好、复杂性适中。调色板编码是屏幕图像编码的主流方法之一, 但其索引图的预测编码效率仍有待提高。为此提出一种基于局部方向相关性的索引图快速预测编码算法。**方法** 实验统计发现两个相邻的索引值在 0.93 的平均概率下具有相同的方向性, 本文称之为索引图的“局部方向相关性”, 进而提出采用一个 2×3 的模板进行初始方向预测, 在其预测失败的情况下再采用一个 3×4 的方向模板进行二次方向预测。**结果** 在 19 个标准测试视频序列和 3 个测试图像上进行了大量实验, 实验结果表明, 本文算法的预测准确率为 95.43%, 比典型的多级预测算法(MSP)平均提高 2.48%, 尤其对于包含文字字符和大量几何图元的复杂场景视频更为有效, 并且计算复杂度显著低于 MSP 算法, 能够较好地满足屏幕图像编码的要求。**结论** 本文提出的索引图的预测算法, 发掘了索引图的局部方向相关性, 加快了预测速度, 可应用在屏幕图像文本/图形块的调色板编码中。

关键词: 图像编码; 屏幕图像; 复合图像; 索引图; 方向预测

Fast prediction algorithm of index maps for screen image coding

Chen Guisheng¹, Song Chuanming¹, Wang Xianghai¹, Liu Dan^{1,2}

1. School of Computer and Information Technology, Liaoning Normal University, Dalian 116029, China;

2. School of Information and Communication Engineering, Dalian University of Technology, Dalian 116024, China

Abstract: **Objective** Screen image coding requires high coding efficiency, real-time performance, and moderate computational complexity. Palette coding is a state-of-art screen content coding method, but its predictive coding efficiency of index maps needs improvement. This study proposes a fast prediction algorithm of index maps based on the local directional correlation. **Method** Experiments show that two neighboring indexes have the same directionality with a probability of 0.93. We define this “the local directional correlation” of an index map. We then use a 2×3 template to perform an initial direction prediction. If the initial prediction fails, we use a 3×4 template to perform the second-round direction prediction. **Result** We conducted extensive experiments on 19 standard test video sequences and 3 test images. Experimental results showed that the prediction accuracy of our algorithm reached 95.43%, which was a 2.48% average increase over typical multi-stage prediction algorithms, and was particularly suitable for videos with text characters, complex scenes, and multiple geometric elements. Moreover, the computational complexity was significantly lower than that of MSP. Thus, this algorithm

收稿日期: 2016-01-17; 修回日期: 2016-04-17

基金项目: 国家自然科学基金项目(61402214, 41271422); 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(L2013406, L2015285); 大连市科学技术基金项目(2013J21DW027)

第一作者简介: 陈规胜(1985—)男, 辽宁师范大学软件工程专业硕士研究生, 主要研究方向为图像和视频编码。E-mail: cgs-10@163.com

通信作者: 宋传鸣, 副教授, E-mail: chmsong@lnnu.edu.cn

Supported by: National Natural Science Foundation of China(61402214, 41271422)

satisfied the requirements of screen image coding. **Conclusion** This study presented a prediction algorithm of index maps, which exploited the local directional correlation of index maps and accelerated the prediction speed. The proposed algorithm was found to be applicable to the palette-based coding of text/graphics blocks in screen images.

Key words: image coding; screen image; compound image; index map; direction prediction

0 引言

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 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^[2]、桌面云等应用,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 1 920 × 1 080 像素分辨率@ 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、JPEG 2000、JPEG-LS 对屏幕视频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^[1,4],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,研究一种压缩效率高、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。

在复合图像编码的研究初期,人们尝试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。文献[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比特。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以自适应地调整或者预先设定,可是由于文本图像和自然图像有着截然不同的统计和频谱特征^[8-9],残差图像的统计特性发生变化,用传统的预测+变换结构的单一编码器压缩屏幕图像等复合图像的效率大多不够理想。于是,就出现了基于对象分割的编码算法^[10]、基于层的编码算法^[4,11]和基于块的编码算法。其中,基于对象分割的编码算法将图像划分成具有一定语义的、形状不规则的区域(如一个图形或字符等)再分别编码;基于层的编码算法将像素划分到不同的层(如前景层、背景层和掩码层^[4]等)并逐层编码;基于块的编码算法将图像分割成一系列大小相同、互不重叠的像素块,并进一步划分为文本块、图形块和图像块等类型,然后对自然图像块采用 PNG、JPEG 或 JPEG-LS 等算法进行编码(如文献[10,12]),文本/图形块则利用专门技术实现压缩。由于前两者的分割过程较为复杂且技术尚不成熟,基于块的编码算法就逐渐发展为屏幕图像等复合图像编码的首选。

考虑到自然图像块的编码技术已趋于成熟,而文本/图形块却缺乏高效率的编码算法,现有研究工作大多集中在后者,如文献[13-17]提出了用于文本/图形块的词典编码和 Lempel-Ziv 编码,文献[18]提出了基于形状原语提取的编码(SPEC)等。尤其近两年来,为了应对愈加迫切的屏幕视频编码需求,国际视频编码联合协作组(JCT-VC)成立了专门工作组负责制定 HEVC 的扩展标准——屏幕视频编码 HEVC-SCC (high efficiency video coding-screen content coding extension),并在世界范围内征集方案^[18]。目前,该标准已采纳了帧内块拷贝(intra block copy)^[19]、调色板^[20]、自适应颜色变换^[21]和自适应运动向量分辨率^[22]等 4 种新的编码模式,

其中,由于调色板编码发掘了文本/图形图像存在的非连续色调、灰度种类有限的数据冗余特点,成为针对文本/图形图像压缩的重要算法之一。

调色板编码算法的基本思想是选取图像中出现次数最多的几种灰度值作为基本颜色,并为其分别指定索引值建立调色板,再将原图像的每个像素值用对应的索引值替代从而形成一张索引图,最后对调色板和索引图进行预测编码和熵编码。由此可见,若要提高调色板编码的效率,主要有3个环节可以考虑:首先,调色板和索引图的生成环节^[23],研究如何确定基本颜色的数量^[24],减少逃逸色和优化索引值的分配^[25];其次,调色板编码环节,研究如何建立全局调色板或局部调色板,以及基本颜色的预测编码方法;最后,索引图编码环节,研究如何高效压缩索引图,如采用预测编码^[26]、词典编码^[17]、模板匹配^[26-27]和算术编码等。索引图编码是调色板编码最主要的组成部分,其编码效率直接影响调色板编码算法的整体压缩性能,故本文的研究目的旨在改善索引图的预测编码效率。

一方面,作为一种无损编码,索引图的词典编码无法高效率地编码内容相似的像素块^[17],在低码率下的编码效率不够高;另一方面,尽管模板匹配具有较高的编码效率,可是其时空复杂度却较高。而且,现有的词典编码和模板匹配均没有很好地发掘复合图像存在的强局部方向相关性,尤其是对于包含汉字等复杂文字复合图像的预测效率尚有待提高。为此,在分析了复合图像的局部方向相关性的基础上,提出一种更加高效率的索引图预测算法。该算法采用两轮方向预测,初始方向预测和文献^[26]的方向预测类似,但方向的生成方式不同,同时舍弃了文献^[26]的模板预测;在初始方向预测失败的情况下,依据像素局部方向相关性进行二次方向预测。实验结果表明,该算法的预测效率较之文献^[26]有明显提高。

1 调色板编码算法及索引图预测

文本/图形图像是典型的非连续色调图像,包含的灰度值较少,针对这样的数据分布特点,研究人员提出了调色板编码算法。将在简述调色板编码算法的基础上,着重介绍索引图编码算法。

1.1 调色板编码算法概述

调色板编码算法的主要过程如图1所示,首先

将图像中出现次数最多的几种灰度值作为基本颜色,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幅索引图;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。

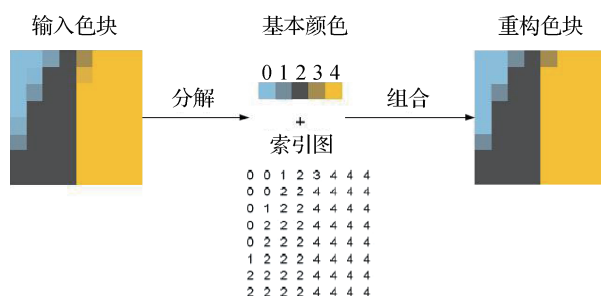


图1 调色板编码算法原理示意图^[26]

Fig. 1 Diagram of palette coding algorithm principle^[26]

调色板的基本颜色数量越多,逃逸色就越少,失真也越小,但索引图却越复杂,不利于提高编码效率;反之,基本颜色数量越少,索引图就越便于压缩,但逃逸色增加,失真也越大。文献^[23]提出了一种基于率失真优化的方法构造调色板和索引图;文献^[24]则通过实验得出选取4种基本颜色为最佳方案的结论,因此在大多算法普遍使用4种基本颜色。

确定了基本颜色数量后,图像块中出现次数最多的4种颜色就被指定为基本颜色,再用某个量化步长将其他颜色量化成基本颜色,不能量化成基本颜色的灰度值作为逃逸色,即

$$Q(x) = \begin{cases} BC_{k^*} & |x - BC_{k^*}| \leq \Delta_1 \\ |x/\Delta_2| & \text{其他} \end{cases} \quad (1)$$

$$k^* = \arg \min_{k=0,1,\dots,b-1} |x - BC_k| \quad (2)$$

式中, x 是被量化的颜色值, b 是基本颜色的数量(取 $b=4$), BC_k 是基本颜色, Δ_1 和 Δ_2 为量化步长,依据编码质量要求而定。如果 $|x - BC_k| > \Delta_1$,则将 x 作为逃逸色。所有的逃逸色使用同一索引值,这样索引图由5种索引值构成(如图1所示)。

接下来,对上述生成的调色板和索引图进行编码。对于调色板的编码,考虑到相邻图像块中包含相似颜色的可能性很大,目前一般采用预测编码实现压缩,即利用已编码图像块的调色板和当前图像块的已编码过的部分调色板来预测当前待编码的调色板,详细过程可参见文献^[28]。而对于索引图编码,由于视频帧的索引图是调色板编码需要处理的

体量最大的一部分数据,对其进行有效编码成为提高调色板算法压缩效率的关键。文献[28]提出采用脉冲编码调制(PCM)对索引图编码;文献[8]采用自适应二元算术编码(CABAC)编码索引图;文献[24]提出了一种2维分层预测模式,文献[9,29]提出了一种1D/2D混合串匹配编码,二者同时兼顾了索引图的局部相关和存在的全局重复模式(pattern);而文献[26]则较早地引入了复合图像的方向预测,提出一种多级预测编码方法(MSP)。由于该方法发掘了复合图像的局部方向相关性,而且能够比上述两种方法更有效地定位全局数据相关,MSP取得了很好的预测效率。

1.2 索引图的多级预测算法

MSP预测方法包括两级预测,第1级为方向预测,第2级为模板预测。前者用于去除局部数据冗余,后者用于去除全局数据冗余(如全局重复模式等)。

1)方向预测。文献[26]认为,文本/图形图像在某个局部区域内的纹理方向大多是一致的。基于这一假设,MSP算法引入了方向预测作为其第1级预测,其基本思路为将已编码索引值的梯度方向作为当前待编码索引值的预测方向。如图2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;其次,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值。

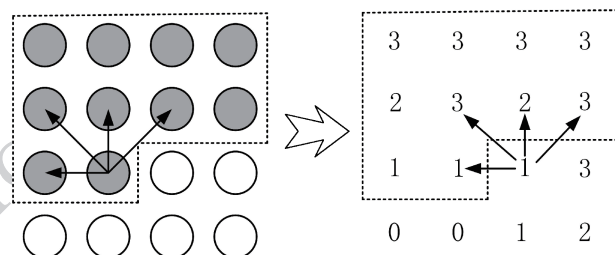


图2 方向预测示意图^[26]

Fig. 2 Diagram of directional prediction^[26]

2)模板预测。为了进一步发掘复合图像存在的全局重复模式,MSP算法又引入了模板预测作为其第2级预测,其基本思路是利用待编码索引的上下文内容统计相关性来计算最佳预测。具体过程如图3所示,实线方框内的索引值是当前待编码的索引,虚线阴影方框内的4个相邻值用来作为预测模

板,在一定范围内的、已经编码过的索引值中查找与预测模板相匹配的索引集合,然后将这些匹配的索引集合作为上下文,统计出现不同索引值的分布情况,例如在图3中,索引值1出现的概率是75%,而索引值2出现的概率是25%,则将出现概率较大的1作为当前待变编码索引的预测值。

0	0	1	2	3	4	4	4	0	0	1	2	3	4	4
0	0	1	2	2	4	4	4	0	0	1	2	2	4	4
0	1	2	2	4	4	4	4	0	1	2	2	4	4	4
0	2	2	2	4	4	4	4	0	0	1	2	2	4	4
0	2	2	2	4	4	4	4	0	2	2	2	4	4	4
1	2	2	2	4	4	4	4	1	2	2	2	4	4	4
2	2	2	2	4	4	4	4	2	2	2	2	4	4	4
2	2	2	2	4	4	4	4	2	2	2	2	4	4	4
0	0	1	2	3	4	4	4	0	0	1	2	3	4	4
0	0	1	2	2	4	4	4	0	0	1	2	2	4	4
0	1	2	2	4	4	4	4	0	1	2	2	4	4	4
0	2	2	2	4	4	4	4	0	2	2	2	4	4	4
0	2	2	2	4	4	4	4	0	2	2	2	4	4	4
1	2	2	2	4	4	4	4	1	2	2	2	4	4	4
2	2	2	2	4	4	4	4	2	2	2	2	4	4	4
2	2	2	2	4	4	4	4	2	2	2	2	4	4	4

图3 模板预测示意图^[26]

Fig. 3 Diagram of template prediction^[26]

一方面,文献[26]通过实验统计发现方向预测阶段可准确预测83%的索引值,而模板预测阶段可成功预测53%的索引值。另一方面,从时间复杂度来看,方向预测阶段预测一个索引值的复杂度为 $O(1)$,模板预测阶段的复杂度为 $O(WT)$ (W 表示搜索窗口中的待匹配索引数目, T 表示模板包含的索引数目)。可见,模板预测的平均效率并不令人满意。而且,对于包含复杂笔画文字(比如汉字)的文本图像,图案大多不具有特定的全局模式(除非选取较大尺寸的模板),故此模板预测有一定局限性。

2 索引图的局部方向预测方法

2.1 索引图的局部方向相关性

因为在文本/图形的复合图像中存在大量笔画、线条等具有明显几何正则性的元素,所以索引图的相邻索引值之间存在较强的方向相关性。为了证明这一点,利用如图4所示的模板统计不同索引间的方向相关性,结果如表1所示(“ I ”表示索引值,下标表示模板的不同位置),可见索引 x 处的方向与其相邻索引处的方向存在非常高的相关性(条件概率达到0.93),并且随着距离的增加,其相关性逐渐降低,如 $P(I_x = I_a | I_h = I_i) < P(I_x = I_a | I_b = I_c)$ 。表1的统计结果说明在索引值 x 的一定邻域内,若沿着

某方向的两个索引值相等, 则 x 处很可能也存在着沿该方向的几何正则性, 本文将索引图的这种数据相关称为“局部方向相关性”。另外, 由表 1 还可以发现, 相邻索引沿着垂直、水平方向的相关性高于沿着对角线、反对角线方向的相关性, 所以在进行方向预测时, 应优先沿着垂直和水平方向, 再沿着对角线和反对角线方向。

在上述分析基础上, 提出利用待编码索引的局部邻域方向性对其进行预测, 而不采用效率不高且计算

量大的非邻域模板预测方式。本文局部方向预测方法包括两个步骤: 初始方向预测和二次方向预测。

g	h	i	j
f	b	c	d
e	a	x	

图4 局部方向相关性统计模板

Fig. 4 Template for local directional correlation statistics

表1 不同测试序列的局部方向相关性统计

Table 1 Local directional correlation statistics of various test sequences

测试序列	$P(I_x = I_a I_b = I_c)$	$P(I_x = I_c I_b = I_a)$	$P(I_x = I_b I_f = I_a)$	$P(I_x = I_d I_c = I_a)$
ChinaSpeed_1024 × 768_30_1	0.97	0.97	0.96	0.96
programming_1280 × 720_60_1	0.97	0.97	0.94	0.94
wordEditing_1280 × 720_60_1	0.95	0.96	0.94	0.94
SlideEditing_1280 × 720_30_1	0.95	0.95	0.92	0.92
英文文本	0.92	0.94	0.91	0.91
简体中文	0.91	0.91	0.88	0.88
繁体中文	0.89	0.89	0.87	0.87
测试序列	$P(I_x = I_a I_h = I_i)$	$P(I_x = I_c I_e = I_f)$	$P(I_x = I_b I_d = I_i)$	$P(I_x = I_d I_f = I_h)$
ChinaSpeed_1024 × 768_30_1	0.96	0.96	0.95	0.93
programming_1280 × 720_60_1	0.96	0.96	0.91	0.86
wordEditing_1280 × 720_60_1	0.93	0.94	0.92	0.90
SlideEditing_1280 × 720_30_1	0.93	0.94	0.90	0.87
英文文本	0.88	0.91	0.88	0.86
简体中文	0.88	0.88	0.86	0.84
繁体中文	0.86	0.86	0.83	0.82

2.2 初始方向预测

考虑到索引值存在局部方向相关性, 在对待编码索引值 x 进行预测时, 首选其水平方向相邻索引值的预测方向作为初始预测方向。具体地说, 如果前一索引值沿着某个方向 D 预测成功, 则当前索引值 x 仍然沿着方向 D 进行预测; 若前一索引值预测失败, 则采用与 MSP 的方向预测算法进行初始预测, 此时若仍旧预测失败, 则进入二次方向预测。

2.3 二次方向预测

由于初始方向预测失败, 二次方向预测的目的是在待编码索引 x 处判别可能存在的几何正则性。基本思路是首先利用 x 的因果 4-直接邻域索引 a 、 b 、 c 和 d (图 4) 判别 180° 、 135° 、 90° 和 45° 方向的正则性。若仍未能有效判别恰当的预测方向 (如 a 、 b 、

c 和 d 均不相等), 表明 x 可能处于几何正则性更复杂的区域, 则考虑 x 的因果 6-间接邻域索引 e 、 f 、 g 、 h 、 i 和 j (图 4) 来进一步判别沿 60° 、 120° 和 150° 方向的几何正则性, 例如若模板上的 j 点等于 d 点表明很可能存在 60° 的边缘 (图 5(a))。

依据上述思路, 下面分不同情形讨论二次方向预测的过程, 主要原则为优先沿水平、竖直方向预测, 继而沿对角线、反对角线方向预测, 再进行 60° 、 120° 和 150° 方向预测。约定当前待编码的索引值为 x , 其预测值为 P_x , 具体预测算法如下:

情形 1: x 的初始预测方向为水平方向。由于初始预测失败, 表明 x 处不存在水平方向的边缘, 即 $I_x \neq I_a$, 则执行以下判别步骤:

1) 如果 $I_a = I_b$ 并且 $I_a \neq I_c$, 那么说明 x 处很可能

存在竖直边缘,令 $P_x = I_c$ 。

2) 如果 $I_a = I_b$ 并且 $I_a = I_c$, 那么排除 I_b 和 I_c , 令 $P_x = I_d$ 。

3) 如果 $I_a = I_c$ 并且 $I_a \neq I_d$, 那么说明 x 处很可能存在 45° 方向的边缘 (图 5(b)), 令 $P_x = I_d$ 。

4) 如果 $I_a = I_c$ 并且 $I_a = I_d$, 那么排除 I_c 和 I_d , 令 $P_x = I_b$ 。

5) 如果 $I_a \neq I_d$:

(1) 如果 $I_c = I_i$, 那么说明 x 处很可能存在垂直方向的边缘, 令 $P_x = I_c$;

(2) 如果 $I_f = I_b$ 或者 $I_h = I_b$ 或者 $I_g = I_b$, 或者 $I_c = I_h$, 那么说明 x 处很可能存在 150° (图 5(c))、

120° (图 5(d)) 或 135° (图 5(e)) 方向的边缘, 令 $P_x = I_b$;

(3) 如果 $I_j = I_d$ 或者 $I_j = I_c$, 那么说明 x 处很可能存在 60° 或 45° 方向的边缘, 令 $P_x = I_d$ 。

6) 如果 $I_a = I_d$:

(1) 如果 $I_i = I_c$ 或者 $I_j = I_c$, 那么说明 x 处很可能存在垂直或者 60° 方向的线, 令 $P_x = I_c$;

(2) 如果 $I_f = I_b$ 或者 $I_g = I_b$ 或者 $I_h = I_b$, 或者 $I_c = I_h$, 那么说明 x 处很可能存在 120° 、 135° 或 150° 方向的线, 令 $P_x = I_b$ 。

7) 若上述条件不满足, 则默认 $P_x = I_c$, 算法结束。

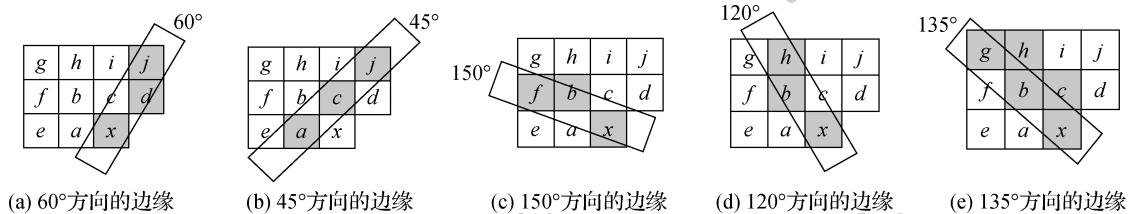


图 5 几何正则性与索引值的关系

Fig. 5 Relationship among geometric regularity and index values ((a) the edge along 60° ; (b) the edge along 45° ; (c) the edge along 150° ; (d) the edge along 120° , and (e) the edge along 135°)

情形 2: x 的初始预测方向为垂直方向。由于初始预测失败, 表明 x 处不存在垂直方向的边缘, 即 $I_x \neq I_c$, 则按照类似情形 1 的处理思路, 执行以下判别步骤:

1) 如果 $I_c = I_a$ 并且 $I_c \neq I_d$, 那么令 $P_x = I_d$ 。

2) 如果 $I_c = I_a$ 并且 $I_c = I_d$, 那么令 $P_x = I_b$ 。

3) 如果 $I_c = I_b$ 或者 $I_c = I_d$ 并且 $I_c \neq I_a$, 那么令 $P_x = I_a$ 。

4) 如果 $I_c = I_b$ 并且 $I_c = I_a$, 那么令 $P_x = I_d$ 。

5) 如果 $I_c \neq I_d$:

(1) 如果 $I_e = I_a$, 那么令 $P_x = I_a$;

(2) 如果 $I_f = I_a$ 或者 $I_f = I_b$ 或者 $I_g = I_b$ 或者 $I_h = I_b$, 那么令 $P_x = I_b$;

(3) 如果 $I_j = I_d$, 那么令 $P_x = I_d$ 。

6) 若上述条件不满足, 则默认令 $P_x = I_a$, 算法结束。

情形 3: x 的初始预测方向为对角线方向。由于初始预测失败, 表明 x 处不存在 135° 方向的边缘, 即 $I_x \neq I_b$, 则执行以下判别步骤:

1) 如果 $I_b = I_a$ 并且 $I_b \neq I_c$, 那么说明 x 处很可能存在竖直边缘, 令 $P_x = I_c$ 。

2) 如果 $I_b = I_a$ 并且 $I_b = I_c$, 那么排除 I_a 和 I_c , 令 $P_x = I_d$ 。

3) 如果 $I_b = I_c$ 并且 $I_b \neq I_a$, 那么说明 x 处很可能存在水平方向的边缘, 令 $P_x = I_a$ 。

4) 如果 $I_b \neq I_a$ 并且 $I_a = I_c$, 那么说明 x 处很可能存在 45° 方向的边缘, 令 $P_x = I_d$ 。

5) 如果 $I_b \neq I_d$:

(1) 如果 $I_c = I_i$, 那么说明 x 处很可能存在竖直方向的边缘, 令 $P_x = I_c$;

(2) 如果 $I_e = I_a$, 那么说明 x 处很可能存在水平方向的边缘, 令 $P_x = I_a$;

(3) 如果 $I_j = I_d$ 或者 $I_j = I_c$, 那么说明 x 处很可能存在 45° 或 60° 方向的边缘, 令 $P_x = I_d$;

(4) 如果 $I_a = I_f$, 那么说明 x 处很可能存在 150° 方向的边缘, 令 $P_x = I_a$ 。

6) 如果 $I_b = I_d$:

(1) 如果 $I_c = I_i$, 那么说明 x 处很可能存在垂直方向的线, 令 $P_x = I_c$;

(2) 如果 $I_e = I_a$, 那么说明 x 处很可能存在水平方向的边缘, 令 $P_x = I_a$;

(3) 如果 $I_j = I_c$ 或者 $I_c = I_h$, 那么说明 x 处很可

能存在 60° 或 120° 方向的线, 令 $P_x = I_c$;

(4) 如果 $I_a = I_f$, 那么说明 x 处很可能存在 150° 方向的边缘, 令 $P_x = I_a$ 。

7) 若上述条件不满足, 则默认令 $P_x = I_c$, 算法结束。

情形4: x 的初始预测方向为反角线方向。由于初始预测失败, 表明 x 处不存在 45° 方向的边缘, 即 $I_x \neq I_d$, 则按照类似情形3的处理思路, 执行以下判别步骤:

1) 如果 $I_d = I_c$ 并且 $I_d \neq I_a$, 那么令 $P_x = I_a$ 。

2) 如果 $I_d = I_c$ 并且 $I_d = I_a$, 那么令 $P_x = I_b$ 。

3) 如果 $I_d = I_a$ 且 $I_d \neq I_b$:

(1) 如果 $I_i = I_c$ 或者 $I_j = I_c$, 那么令 $P_x = I_c$;

(2) 如果 $I_f = I_b$ 或者 $I_g = I_b$ 或者 $I_h = I_b$ 或者 $I_c = I_h$, 那么令 $P_x = I_b$ 。

4) 如果 $I_d = I_a$ 并且 $I_d = I_b$, 那么令 $P_x = I_c$ 。

5) 如果 $I_d = I_b$:

(1) 如果 $I_i = I_c$ 或者 $I_j = I_c$ 或者 $I_h = I_c$, 那么令 $P_x = I_c$;

(2) 如果 $I_e = I_a$ 或者 $I_f = I_a$, 那么令 $P_x = I_a$ 。

6) 如果 $I_d \neq I_b$:

(1) 如果 $I_c = I_i$ 或者 $I_a = I_b$, 那么令 $P_x = I_c$;

(2) 如果 $I_e = I_a$ 或者 $I_b = I_c$, 那么令 $P_x = I_a$;

(3) 如果 $I_f = I_a$ 或者 $I_f = I_b$ 或者 $I_g = I_b$ 或者 $I_h = I_b$ 或者 $I_c = I_h$, 那么令 $P_x = I_b$ 。

7) 若上述条件不满足, 则默认令 $P_x = I_c$, 算法结束。

需要指出, 图4所示的模板对于预测索引 x 处的几何正则性是完备的, 这主要有两方面原因: 一方面, 如果索引 x 经二次方向预测后仍然预测失败, 表明 x 与其8-邻域内的其他索引值均不相等, 那么索引 x 处不可能存在边或线; 另一方面, 如果索引 x 处存在其他方向的边缘, 那么由光栅图形学中的 Bresenham 算法可知, 索引 x 必定至少与模板上 $a-j$ 点中的两点具有相同的索引值, 这样利用上述过程即可完成索引 x 的预测。

3 实验结果与分析

为了验证本文算法的有效性, 选用 JCT-VC 公布的 19 个标准测试序列(表2)的前 90 帧进行实验。此外, 考虑到这些序列中只含有英文文字, 而英

表2 不同测试序列的预测准确率比较

Table 2 Prediction accuracy comparison of various test sequences

视频序列	预测准确率/%		预测性能增益/%
	文献[26]	本文	
Cg_Twist_Tunne	97.28	98.74	+1.50
PCB_Layout	97.73	98.98	+1.28
Programming	95.88	97.10	+1.28
Video_Conferencing_Doc_Sharing	97.05	98.06	+1.05
WordEditing	95.33	96.60	+1.33
SlideEditing	93.73	95.12	+1.49
FlyingGraphics	95.37	96.51	+1.19
Map	87.40	93.03	+6.45
PPT_Doc_xls	95.93	96.97	+1.08
Robot	88.08	93.66	+6.34
WebBrowsing	95.21	96.21	+1.06
BasketballDrillText	83.37	87.35	+4.77
ChinaSpeed	90.43	93.58	+3.48
MissionControlClip3	94.46	96.28	+1.92
CADWaveform	98.06	98.60	+0.56
Console	97.66	98.56	+0.93
Desktop	95.52	97.13	+1.69
SlideShow	97.39	98.40	+1.03
Venuevu	94.92	97.36	+2.61
英文文本	89.38	92.34	+3.31
繁体中文	86.35	90.44	+4.74
简体中文	83.92	88.50	+5.45

文与其他文字的像素分布存在一定差异, 不利于有效评价现有算法对于包含中文等复杂文字的视频序列的预测效率, 本文又截取了3幅包含英文、简体中文和繁体中文的图像, 分辨率均为 640×640 像素, 如图6所示。

实验环境配置如下: Intel Xeon(R) E5-2609 @ 2.4 GHz CPU, 16 GB 主存, 64 位 Windows7 旗舰版, 开发环境为 Matlab R2013a。

3.1 预测准确率比较

所有实验均在各视频帧/图像的亮度分量进行, 首先利用文献[24]的算法计算得到亮度分量的索引图, 再采用本文算法进行预测并统计预测准确率, 同时将结果与 MSP^[26] 预测结果进行比较, 结果罗列在表2中。从表2可见, 本文算法对所有测试序列的

Abstract The screen image compression as a new branch of image compression, has become the focus of research in recent years. The screen image, fax image and scanned document together to form a class of images: the compound images. Due to the ratio and quality requirements of different types of image compression are different, therefore, the image, which combining with text, graphics and natural images of different content, can't simply compressed by a single compression algorithm. After years of development, there are also some compound image compression standards, such as: DJVU, MRC and so on. Those layer-based algorithm generally segment the image into three layers, include foreground layer, mask layer and background layer and that each layer be compressed by different algorithm respectively. Afterwards, a block-based algorithm be rendered to segment the compound image into different types of image block, include text block, graphic block, pictorial block and so on. The compression algorithm of screen image is based on the general method of compound image compression, but screen image also have its own characteristics. First, Scanned images usually contain some amount of noise, while computer screen images are free of noise. Therefore, for computer screen images, any noise introduced in compression is noticeable in text/graphics regions. Second, Screen images general be used for interactive application, such real-time compression algorithms require very low complexity, whereas scanned image compression does not have such a requirement. Third, Scanned images typically have higher spatial resolution than computer screen images. Finally, Scanned images are captured by an electronic imaging procedure, whereas computer screen images are essentially synthetic images. Ringing artifacts caused by DCT or wavelet transform are not clearly visible around text/graphics edges, because these edges have been blurred in printing or scanning procedures. However, for computer screen images, these ringing artifacts are easily noticeable due to the sharp edges of text/graphics. This paper give a detailed introduction of two essential framework, layer-based and

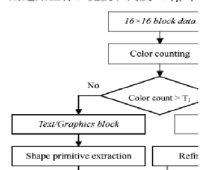
(a) 英文文本测试图像

8]就采用了多种块大小,例如:8x8、16x16、32x32等,这体现了基于块分割方法的灵活性。不同类型的块之间还有着不同的统计特征,例如:颜色数量、直方图、梯度和频谱等,这也使得块分类更容易实现。文献[7]的分类方法是为一块的颜色数量设置一个阈值,小于该阈值的为文本块,否则为图像块。文献[16]分析了文本/图形图像和自然图像的频谱活动 SAM (spectral activity measure) 和空间频率 SFM (spatial frequency measure)。文献[19]利用了像素的梯度值,文本/图形块通常包含较多的高梯度值像素。

文献[6]首先提出了混合图像块的方法,简化了图像的分割问题。还提出了通过图像块内部像素来实现一次遍历 (one pass) 的图像块分类方法,称为块分类,例如:可以通过统计一个像素块内部颜色的数量,如果颜色数量小于一个阈值,则归为文本或图形块,否则归为图像块,这大大地简化了混合图像的像素块分类。由于在一定范围内,相邻的像素块一般属于同一类型,文章还进一步利用了像素块与像素块之间的局部相关性来简化分类。通过分析一个像素块是否处在两种类型的图像的边界,来决定类型的转变。分类之后,文本块采用无损/近无损的 JPEG-LS 标准算法,图像块则采用基于 DCT 的 JPEG 标准算法,处于边界的像素块则再进行特殊处理,通过设置不同的量化参数,增加额外的缓冲等方法,以便在无损和无损压缩之间过渡。文章所采用的方法实现的复杂度比较低,但对文本块的压缩采用的是 JPEG-LS 无损压缩算法,压缩率比较低,而且对包含边界的像素块的处理是通

类块内模式下的有损压缩,代码不能有效地保持锐利的边界。

文献[7]提出形状原语提取和 primitive extraction and coding)。原语:孤立点、水平线、垂直线、像划分为16x16的像素块,再统计量大于一个设定的阈值 (SPEC设为1) 图形块,否则,作为图像块。形:左下角开始按从左到右,从上到下如果当前像素点已包含于前一形状点成为当前像素点,向右、向上、向左或者向右的点,或者是点、或者是矩形,任何一个复杂的形状用这四种原语描述。对形状原语的标注值 (x,y) 表示孤立点;用起始表示水平线;用起始坐标和高度表示起始坐标、宽度和高度 (x,y,w,h)。



(b) 简体中文测试图像

輪到用戶端,在用戶端解釋執行後重新生成螢幕圖像。然而,這種方法文本和圖形資訊佔用比較少的頻寬,並且得到最好的品質,但需要作業系統或者應用軟體的支援,不利於跨平台;並且有些用戶端不支援這種高級圖形語言,致使實現起來非常困難。

(2) 截取螢幕圖像

截取螢幕圖像再通過壓縮編碼後傳輸到用戶端,再由用戶端解碼。這種方法實現起來簡單、可靠,不受具體的軟體平臺的影響;頻寬和品質完全取決於壓縮演算法的有效性。這就對螢幕圖像的壓縮演算法提出了幾點要求:

- 保證較高的保真度
- 實現較高的壓縮率滿足網路傳輸的需要
- 演算法複雜度低,以保留足夠多的 CPU 資源

螢幕圖像混合了文本、圖形和自然圖像等不同內容,和傳真圖像、掃描圖像等一起歸為同一類:混合圖像 (如圖 1 所示)。螢幕圖像作為混合圖像的其中一類,其壓縮演算法在很大程度上繼承了混合圖像壓縮的一般方法,例如: Djvu、MRC 等比較成熟的演算法可以為螢幕圖像壓縮所借鑒和參考。但螢幕圖像和其他類型的混合圖像相比,又有一些獨有特點,例如:由於塵埃污染和掃描器本身的緣故,掃描圖像存在大量固有的雜訊,看起來平滑的區域實際上有大量的起伏;而螢幕圖像則不受雜訊影響。螢幕圖像存在時間相關性,可以像視頻圖像序列一樣採用幀間預測編碼。螢幕圖像還包含更加豐富多樣的內容。此外,螢幕圖像壓縮和掃描圖像壓縮在應用上也不同,掃描圖

目前的一些圖像壓縮標準 (例如 H.264 幀內編碼) 大多是用於自

案嘗試利用這些成熟的壓縮演算法縮混合圖像。文獻[14]就使用 JPEG 調整調整量化來壓縮混合圖像,清晰的邊界,該方法採用較小的量區域。文獻[15]也一樣採用單一像,但和文獻[14]不同的是,壓縮區域分配更多的比特來保持文字,就必須使用閾值,該方法就是分化的 DCT 係數,計算變換之後花算重建圖像有失真的失真,對率變化的係數將被捨棄,這實際上就內容自我調整地調整率失真。H.264/AVC 的幀內編碼模式,只採碼文本內容和圖像內容,其實現不同宏塊之間改變 H.264/AVC 的量和圖像塊的壓縮要求。儘管參數自我調整地調整或者人為事先設定和圖像區域之間有著不同的統計如 2.1 節所描述,使得單一編碼器不夠理想,所以大多數文獻採用了不同的內容,對不同內容採用不同

2.1 不同類型圖像的壓縮要求

不同類型的圖像要求不同的

(c) 繁体中文测试图像

图 6 包含复杂文字的测试图像

Fig. 6 Test images containing complex characters ((a) the test image of English text; (b) the test image of simplified Chinese text; (c) the test image of traditional Chinese text)

预测准确率均高于 MSP,比后者平均提高 2.48%。

一方面,对于纯文本图像,由于英文文本图像只

由 26 个字符的图像组成,每个字符重复出现的概率较高,MSP 算法的模板预测可以发挥较好的效果。然而,从英文文本到简体中文再到繁体中文,随着字符的笔画越来越复杂,模板预测的准确率也随之降低,利用图像的局部相关性进行预测的本文算法就表现出愈加明显的优势,预测准确率较之 MPS 平均提高了 4.5%。

另一方面, Map、Robot、BasketballDrillText、ChinaSpeed 和 Venuv 等 5 个视频序列不仅包含文字字符,还包含由计算机生成的大量几何图元和复杂场景,画面细腻、线条不规则。此时, MSP 算法的模板预测存在明显局限,方向预测又只考虑了 45°、90°、135°和 180°方向的索引相关性;本文算法则进一步利用了沿着 60°、120°、135°和 150°方向的索引值依赖,更为充分地发掘了局部方向相关性,于是预测准确率较之 MPS 平均提高了 4.73%。需要特别指出的是,虽然其他视频序列也包含文字字符、线条等,可是还存在大面积的平滑区域 (尤其是 CAD-Waveform 和 Console 序列),字符和线条所占区域比例较低,而 MSP 与本文算法在平滑区的预测性能几乎无差别,这样本文算法的平均预测准确率增益在 1.24%。

此外,为比较 MSP 和本文算法的预测稳定性,绘制了 Map、Robot、Console 和 CADWaveform 等 4 个测试序列的预测准确率曲线,结果如图 7 所示。其中, Map 和 Robot 是两种算法平均预测准确率差异最大的序列, Console 和 CADWaveform 则是平均性能差异最小的序列。从图 7 中可以看出,无论对于 Map、Robot 序列,抑或对于 Console、CADWaveform 序列,本文算法在每一帧上的预测准确率均高于 MSP,且波动幅度小于后者。这说明本文算法的预测性能比 MSP 更加稳定,较后者更利于编码器输出码率更为平稳的编码码流。

3.2 计算复杂度比较

根据 2.3 节的论述,本文算法在预测过程中对待编码索引的局部邻域内的 10 个索引值进行最多 12 次比较,可知其时间和空间复杂度均为 $O(1)$ 。而由 1.2 节, MSP 预测算法的时间复杂度为 $O(WT)$, W 表示搜索窗口中的待匹配索引数目, T 表示模板包含的索引数目;由于需要将待匹配索引进行缓冲存储,其空间复杂度为 $O(W)$ 。故此,本文算法的时间复杂度较之 MSP 降低了 WT 倍,空间复杂度降低了 W 倍。

具体地,表 3 统计了 MSP 预测算法和本文算法在 22 个测试序列上的实际运行时间。从表 3 可以看出,本文算法在每个视频帧上的平均计算时间均

明显少于 MSP 算法,较后者平均节省了 94.39%。

因此,从理论分析和实际运行情况两个方面来看,本文算法的预测效率都取得了有效提升。

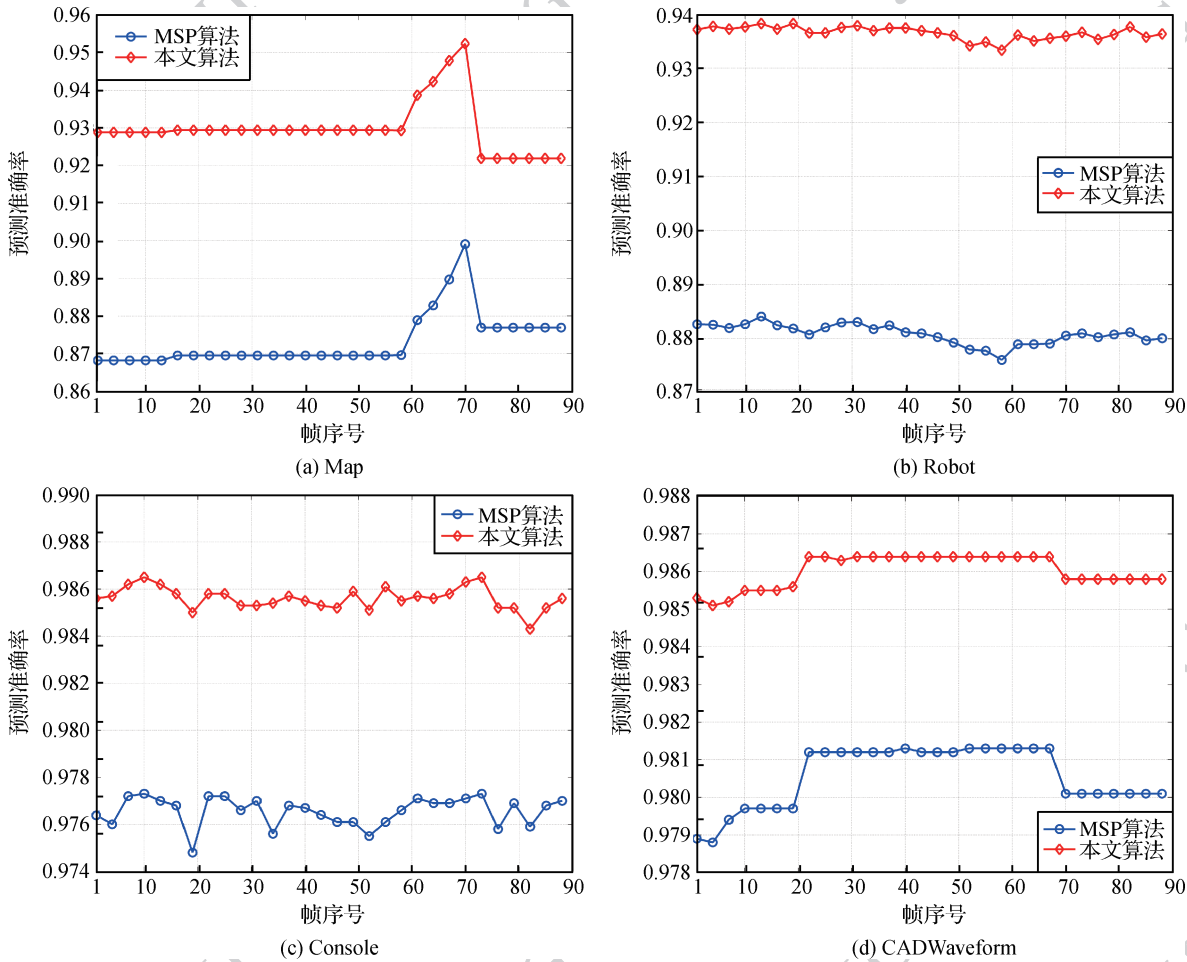


图 7 4 个不同测试序列的预测准确率曲线

Fig. 7 Prediction accuracy curve of four different test sequences ((a) Map; (b) Robot; (c) Console; (d) CADWaveform)

表 3 两种算法在不同测试序列上的运行时间比较

Table 3 Running time comparison of two algorithms on various test sequences

视频序列	运行时间/(s/帧)		预测速度 增益/%	视频序列	运行时间/(s/帧)		预测速度 增益/%
	MSP 算法	本文算法			MSP 算法	本文算法	
Cg_Twist_Tunne	35.66	1.86	+94.78	BasketballDrillText	17.21	1.06	+93.82
PCB_Layout	68.80	3.55	+94.84	ChinaSpeed	29.71	1.65	+94.44
Programming	32.29	1.69	+94.76	MissionControlClip3	71.18	4.06	+94.30
Video_Conferencing_Doc_Sharing	36.78	1.75	+95.24	CADWaveform	80.92	3.91	+95.17
WordEditing	34.92	1.79	+94.88	Console	85.73	4.13	+95.18
SlideEditing	37.04	1.92	+94.81	Desktop	73.49	3.98	+94.59
FlyingGraphics	79.00	4.34	+94.51	SlideShow	21.34	0.99	+95.38
Map	31.47	2.14	+93.20	Venuvu	78.35	4.07	+94.82
PPT_Doc_xls	74.73	3.98	+94.67	英文文本	17.51	1.18	+93.25
Robot	34.39	2.29	+93.33	简体中文	17.07	1.17	+93.17
WebBrowsing	36.62	1.87	+97.89	繁体中文	16.32	1.20	+92.65

4 结 论

调色板算法是一种压缩效率较高、实时性好、复杂性适中的复合图像,尤其是屏幕图像的编码方法。作为调色板编码的主要组成部分之一,索引图编码的效率直接影响到调色板编码算法的整体压缩性能。本文分析了复合图像的局部方向相关性,以及多级预测编码 MSP 方法在方向预测和模板预测阶段存在的优势和不足,并在此基础上提出一种更加高效率的索引图预测算法。实验结果表明,该算法在显著降低计算复杂度的情况下,预测效率较之 MSP 方法平均提高了 2.48%,且预测性能更加稳定,可有效应用在面向屏幕图像的文本/图形块的调色板编码过程中。另外,现有算法的索引图生成过程和预测编码过程相互独立,需两次遍历一幅屏幕图像才能完成,而这两个过程本质上是同时进行的,也是紧密相关的。故此,通过一次遍历实现索引图建立与索引图预测的联合算法可望获得更低的编码复杂度和更高的预测效率,这也是进一步努力的方向。

参考文献 (References)

- [1] De Queiroz R L, Buckley R R, Xu M. Mixed raster content (MRC) model for compound image compression[C]//Proceedings of SPIE 3653, Visual Communications and Image Processing. San Jose, CA: SPIE Press, 1999: 1106-1117. [DOI: 10.1117/12.334618]
- [2] Lu Y, Li S P, Shen H F. Virtualized screen: a third element for cloud-mobile convergence [J]. IEEE Multimedia, 2011, 18(2): 4-11. [DOI: 10.1109/MMUL.2011.33]
- [3] Vermeir T. Use cases and requirements for lossless and screen content coding, JCTVC-M0172[R]. Switzerland, Geneva: ITU-T, Joint Collaborative Team on Video Coding, 2013.
- [4] ITU-T. Recommendation T. 44 Mixed raster content (MRC) [S]. Geneva, Switzerland: ITU-T, 1999.
- [5] Konstantinide K, Tretter D. A JPEG variable quantization method for compound documents [J]. IEEE Transactions on Image Processing, 2000, 9(7): 1282-1287. [DOI: 10.1109/83.847840]
- [6] Ramos M G, De Queiroz R L. Classified JPEG coding of mixed document images for printing [J]. IEEE Transactions on Image Processing, 2000, 9(4): 716-720. [DOI: 10.1109/83.841946]
- [7] Zaghetto A, De Queiroz R L. Segmentation-driven compound document coding based on H. 264/AVC-INTRA [J]. IEEE Transactions on Image Processing, 2007, 16(7): 1755-1760. [DOI: 10.1109/TIP.2007.899036]
- [8] Lan C L, Shi G M, Wu F. Compress compound images in H. 264/MPGE-4 AVC by exploiting spatial correlation [J]. IEEE Transactions on Image Processing, 2010, 19(4): 946-957. [DOI: 10.1109/TIP.2009.2038636]
- [9] Ma Z, Wang W, Xu M, et al. Advanced screen content coding using color table and index map [J]. IEEE Transactions on Image Processing, 2014, 23(10): 4399-4412. [DOI: 10.1109/TIP.2014.2346995]
- [10] Said A, Drukarev A. Simplified segmentation for compound image compression [C]//Proceedings of the 1999 International Conference on Image Processing. Kobe: IEEE, 1999: 229-233. [DOI: 10.1109/ICIP.1999.821603]
- [11] Bottou L, Haffner P, Howard P G, et al. High quality document image compression with "DjVu" [J]. Journal of Electronic Imaging, 1998, 7(3): 410-425. [DOI: 10.1117/1.482609]
- [12] Pan Z T, Shen H F, Lu Y, et al. Browser-friendly hybrid codec for compound image compression [C]//The 2011 IEEE International Symposium on Circuits and Systems. Rio de Janeiro: IEEE, 2011: 101-104. [DOI: 10.1109/ISCAS.2011.5937511]
- [13] Wang S H, Lin T. A unified LZ and hybrid coding for compound image partial-lossless compression [C]//Proceedings of The 2nd International Congress on Image and Signal Processing. Tianjin: IEEE, 2009: 1-5. [DOI: 10.1109/CISP.2009.5301019]
- [14] Lan C, Peng X, Xu J, et al. Intra and inter coding tools for screen contents, JCTVC-E145 [R]. Geneva, Switzerland: ITU-T, Joint Collaborative Team on Video Coding, 2011.
- [15] Wang S, Lin T. 4:4:4 screen content coding using macroblock-adaptive mixed chroma-sampling-rate, JCTVC-H0073 [R]. Geneva, Switzerland: ITU-T, Joint Collaborative Team on Video Coding, 2012.
- [16] Lin T, Zhang P, Wang S, et al. 4:4:4 screen content coding using dual-coder mixed chroma-sampling-rate (DMC) techniques, JCTVC-I0272 [R]. Geneva, Switzerland: ITU-T, Joint Collaborative Team on Video Coding, 2012.
- [17] Zhu W J, Ding W P, Xu J Z, et al. 2-D dictionary based video coding for screen contents [C]//Proceedings of 2014 Data Compression Conference. Snowbird, UT: IEEE, 2014: 43-52. [DOI: 10.1109/DCC.2014.11]
- [18] Lin T, Hao P W. Compound image compression for real-time computer screen image transmission [J]. IEEE Transactions on Image Processing, 2005, 14(8): 993-1005. [DOI: 10.1109/TIP.2005.849776]
- [19] Pang C, Sole J, Guo L, et al. Non-RCE3: intra motion compensation with 2-D MVs, JCTVC-N0256 [R]. Geneva, Switzerland: ITU-T, Joint Collaborative Team on Video Coding, 2013.

- [20] Onno P, Xiu X, Huang Y W, et al. Suggested combined software and text for run-based palette mode, JCTVC-R0348 [R]. Geneva, Switzerland: ITU-T, Joint Collaborative Team on Video Coding, 2014.
- [21] Zhang L, Chen J, Sole J, et al. SCCE5 Test 3. 2. 1: In-loop color-space transform, JCTVC-R0147 [R]. Geneva, Switzerland: ITU-T, Joint Collaborative Team on Video Coding, 2014.
- [22] Li B, Xu J G, Sullivan, et al. Adaptive motion vector resolution for screen content, JCTVC-S0085 [R]. Geneva, Switzerland: ITU-T, Joint Collaborative Team on Video Coding, 2014.
- [23] Zhu W J, Au O C, Dai W, et al. Palette-based compound image compression in HEVC by exploiting non-local spatial correlation [C]//Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing. Florence: IEEE, 2014: 7348-7352. [DOI: 10.1109/ICASSP.2014.6855027]
- [24] Pan Z T, Shen H F, Lu Y, et al. A low-complexity screen compression scheme for interactive screen sharing[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2013, 23(6): 949-960. [DOI: 10.1109/TCSVT.2013.2243056]
- [25] Zeng W J, Li J, Lei S. An efficient color re-indexing scheme for palette-based compression[C]//Proceedings of the 2000 International Conference on Image Processing. Vancouver BC: IEEE, 2000: 476-479. [DOI: 10.1109/ICIP.2000.899448]
- [26] Zhu W J, Ding W P, Xiong R Q, et al. Compound image compression by multi-stage prediction[C]//Proceedings of the 2012 IEEE Visual Communications and Image Processing. San Diego, CA: IEEE, 2012: 1-6. [DOI: 10.1109/VCIP.2012.6410758]
- [27] Tao P, Feng L X, Song S C, et al. Improvement of re-sample template matching for lossless screen content video [C]//Proceedings of the 2015 IEEE International Conference on Multimedia and Expo. Turin: IEEE, 2015: 1-6. [DOI: 10.1109/ICME.2015.7177491]
- [28] Shen H F, Lu Y, Wu F, et al. Low-cost realtime screen sharing to multiple clients[C]//Proceedings of the 2010 IEEE International Conference on Multimedia and Expo. Suntec City: IEEE, 2010: 980-985. [DOI: 10.1109/ICME.2010.5583900]
- [29] Xu Y, Huang W, Wang W, et al. 2-D index map coding for HEVC screen content compression [C]//Proceedings of 2015 Data Compression Conference. Snowbird, UT: IEEE, 2015: 263-272. [DOI: 10.1109/DCC.2015.62]
- [30] ITU-T Q6/16 Visual Coding and ISO/IEC JTC1/SC29/WG11 Coding of Moving Pictures and Audio. Joint call for proposals for coding of screen content, MPEG2014/N14175 [R]. Geneva, Switzerland: ITU-T, ISO/IEC, 2014.