



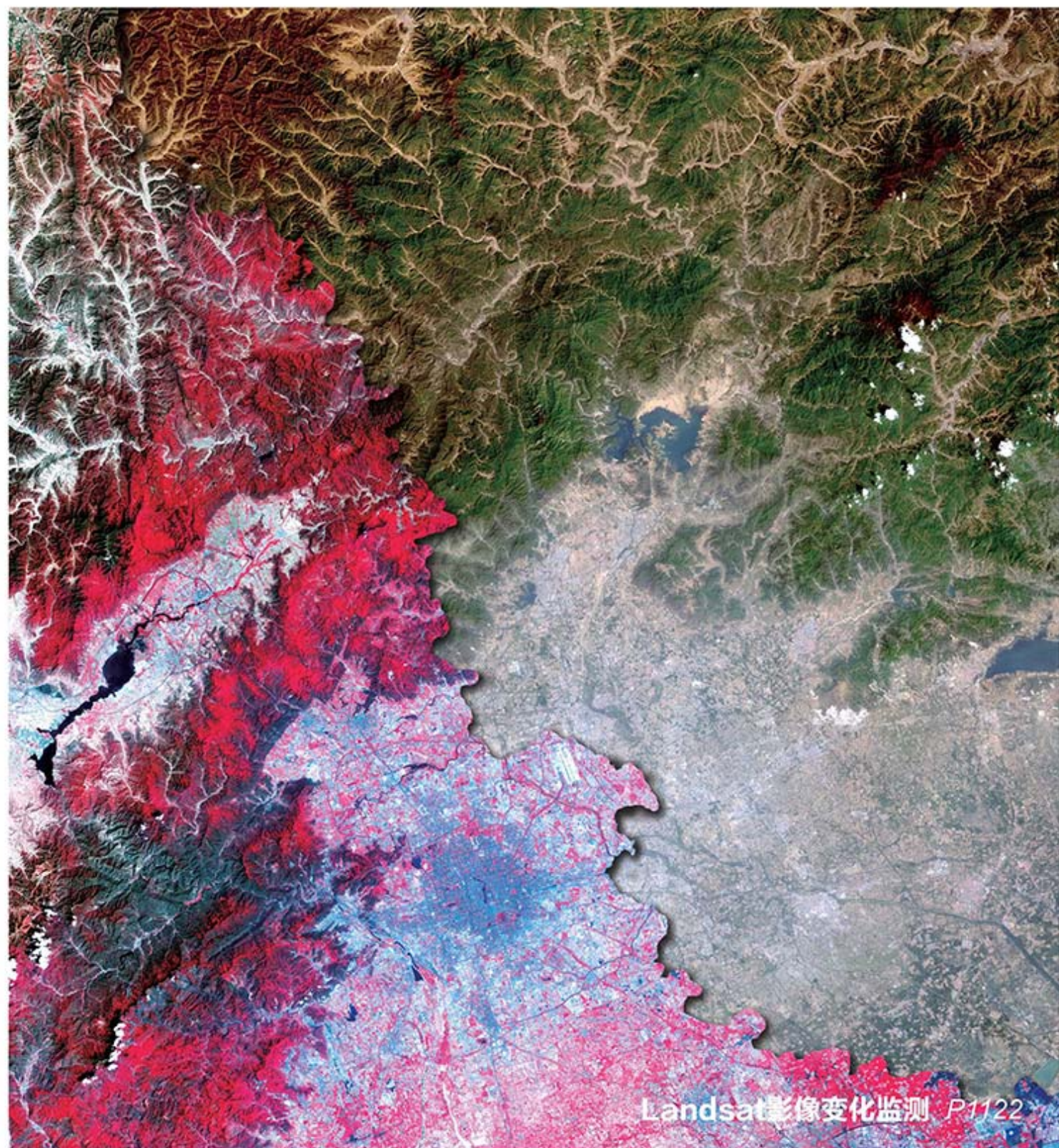
JOURNAL OF IMAGE AND GRAPHICS

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

2015
08
VOL.20

ISSN1006-8961
CN11-3758/TB



Landsat影像变化监测 P1122



第20卷第8期（总第232期）

2015年8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

中国图象图形学学会

北京应用物理与计算数学研究所

主 编 李小文

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 LI Xiaowen

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

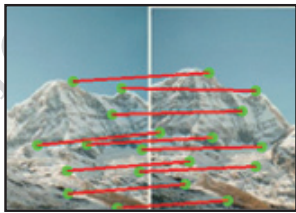
ISSN 1006-8961

CODEN ZTTXFZ

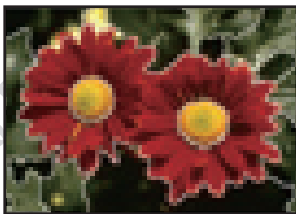
国外发行代号 M1406

国内邮发代号 82-831

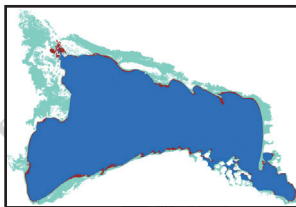
国内定价 50.00元



利用Hough变换的匹配对提纯
(第1017页)



结合格式塔完形规则的自然图像分割(第1026页)



利用长时间序列Landsat分析
博斯腾湖面积变化(第1122页)

图像处理 and 编码

CRT显示图像的色彩管理新模型

黎新伍, 王根生 0993

邻域窗口权重变分的图像修复

王猛, 翟东海, 聂洪玉, 王佳君 1000

自适应滤波窗实现距离加权图像椒盐噪声滤除

林亚明, 李佐勇, 林叶郁, 徐戈 1008

利用Hough变换的匹配对提纯

谢亮, 陈姝, 张钧, 田金文 1017

图像分析和识别

结合格式塔完形规则的自然图像分割

曾接贤, 王玉 1026

融合Kernel PCA形状先验信息的变分图像分割模型

杨建功, 汪西莉, 李虎 1035

基于2维灰度熵阈值选取快速迭代的图像分割

吴诗姘, 吴一全, 周建江, 孟天亮, 戴一冕 1042

基于图像线特征和点云面特征的水下人造目标定位

解则晓, 潘成成, 迟书凯, 魏征 1051

熵能量堆煤图像的定位与识别

袁姮, 王志宏, 姜文涛 1062

时空运动显著性的目标跟踪

谢昭, 刘玉敏, 张骏, 段士雷 1070

结合整体一致性和局部差异性的图像目标显著性检测

张小祥, 马儒宁 1083

基于Radon-经验模式分析的纹理分类

徐卓飞, 张海燕, 刘凯, 侯和平, 徐倩倩, 李利峰 1091

遥感图像处理

四叉树分块的高光谱图像分布式无损编码

王相海, 张智迪, 宋传鸣 1102

基于流形学习的彩色遥感图像分维数估算

王洪波, 罗贺, 王晓佳 1110

利用长时间序列Landsat分析博斯腾湖面积变化

孙爱民, 冯钟葵, 葛小青, 罗宇, 李山山, 闫芳, 冯旭祥 1122

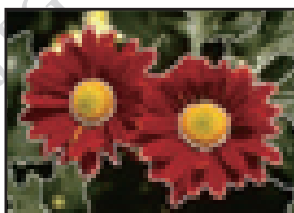
封面图片由中国科学院遥感与数字地球研究所卫星地面系统工程部李山山副研究员提供

CONTENTS

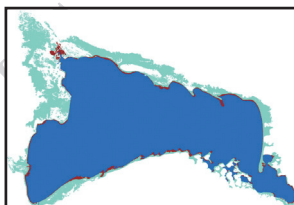
JOURNAL OF IMAGE AND GRAPHICS



Purifying algorithm for rough matched pairs using Hough transform (P1017)



Natural image segmentation method based on Gestalt rules (P1026)



Bosten Lake surface area changes analysis using long temporal Landsat image series (P1122)

Image Processing and Coding

- Color management model for CRT color image
Li Xinwu, Wang Gensheng 0993
- Image inpainting with weight variation of neighborhood window
Wang Meng, Zhai Donghai, Nie Hongyu, Wang Jiajun 1000
- Distance-weighted approach based on self-adaptive windows to remove salt and pepper noise in images
Lin Yaming, Li Zuoyong, Lin Yeyu, Xu Ge 1008
- Purifying algorithm for rough matched pairs using Hough transform
Xie Liang, Chen Shu, Zhang Jun, Tian Jinwen 1017

Image Analysis and Recognition

- Natural image segmentation method based on Gestalt rules
Zeng Jiexian, Wang Yu 1026
- Variational image segmentation incorporating Kernel PCA-based shape priors
Yang Jiangong, Wang Xili, Li Hu 1035
- Image segmentation based on the fast iteration for two-dimensional gray entropy threshold selection
Wu Shihua, Wu Yiquan, Zhou Jianjiang, Meng Tianliang, Dai Yimian 1042
- Underwater man-made object localization based on line feature of image and surface feature of point cloud
Xie Zexiao, Pan Chengcheng, Chi Shukai, Wei Zheng 1051
- Coal pile location and recognition method based on entropy energy
Yuan Heng, Wang Zhihong, Jiang Wentao 1062
- Spatio-temporal motion saliency for object tracking
Xie Zhao, Liu Yumin, Zhang Jun, Duan Shilei 1070
- Image target saliency detection by combining the global consistency and local difference
Zhang Xiaoxiang, Ma Runing 1083
- Texture classification based on Radon-empirical mode decomposition analysis
Xu Zhuofei, Zhang Haiyan, Liu Kai, Hou Heping, Xu Qianqian, Li Lifeng 1091

Remote Sensing Image Processing

- Lossless distributed source coding of hyperspectral images based on quadtree segmentation
Wang Xianghai, Zhang Zhidi, Song Chuanming 1102
- Estimation research for fractal dimension of color remote sensing image based on manifold learning
Wang Hongbo, Luo He, Wang Xiaojia 1110
- Bosten Lake surface area changes analysis using long temporal Landsat image series
Sun Aimin, Feng Zhongkui, Ge Xiaoqing, Luo Yu, Li Shanshan, Yan Fang, Feng Xuxiang 1122

中图法分类号: TP301 文献标识码: A 文章编号: 1006-8961(2015)08-1026-09

论文引用格式: Zeng J X, Wang Y. Natural image segmentation method based on Gestalt rules[J]. Journal of Image and Graphics, 2015, 20(8): 1026-1034. [曾接贤, 王玉. 结合格式塔完形规则的自然图像分割[J]. 中国图象图形学报, 2015, 20(8): 1026-1034.] [DOI: 10. 11834/jig. 20150805]

结合格式塔完形规则的自然图像分割

曾接贤, 王玉

南昌航空大学计算机视觉研究所, 南昌 330063

摘要: **目的** 由于自然图像容易受到光照等因素的影响, 其分割精度往往达不到人类视觉感知的需求, 为此提出了一种新的结合格式塔完形规则的自然图像分割方法。**方法** 首先采用 Neut 算法对原图像进行过分割得到若干个子区域, 这些局部子区域能弱化光照、背景模糊等自然因素的影响; 然后引入格式塔完形规则对区域进行度量, 提出了基于区域的量化计算模型, 进一步弱化了自然因素的影响, 而且所得的区域率更加符合人的视觉感知; 最后在区域率的基础上提出了新的合并算法, 该算法简单且执行效率高, 通过区域合并得到最终的分割结果。**结果** 30幅图像的定量和目视对比实验表明, 本文算法不仅能够很好地将格式塔完形规则应用到图像分割上来, 而且对比实验表明, 本文算法在评价指数 PRI、VOI、GCE 上总体性能要优于其他算法, 与人工标注的结果比较接近。**结论** 提出了一种结合格式塔完形规则的自然图像分割方法, 该方法在过分割的基础上, 采用格式塔完形规则对区域进行度量, 有效降低了自然图像易受自然因素的影响, 提高了分割精度。实验结果表明, 本文提出的结合格式塔完形规则的图像分割算法高效性而准确, 但不适合于尖细物体的自然图像的分割。

关键词: 自然图像; 过分割; 格式塔完形规则; 区域合并

Natural image segmentation method based on Gestalt rules

Zeng Jiexian, Wang Yu

Institute of Computer Vision, Nanchang Hangkong University, Nanchang 330063, China

Abstract: **Objective** In this paper, a new natural image segmentation method based on Gestalt rules is proposed to reduce the effect of illumination and other factors, as well as improve segmentation accuracy. **Method** The novelty of the proposed method lies in three aspects: first, the original image is segmented into several subregions with the use of normalized cut algorithm to reduce the effects of illumination, background smooth, and other natural factors. Second, the Gestalt rules are introduced to measure the regions, and a quantitative calculation model based on the regions, which can obtain the ratio region and agrees with human visual perception, is proposed. Third, a new merge algorithm based on the ratio region is proposed. The final segmentation result is obtained by merging the regions through the simple and highly efficient merge algorithm. **Result** Quantitative and visual inspections of 30 images show the effectiveness of the Gestalt rules on image segmentation. The results of the comparative experiments show that the effect of the algorithm is more aligned with human visual perception and that our algorithm performs better than the comparative experiments overall in the evaluation index of probabilistic Rand index, variation of information, and global consistency error. The results of our algorithm are closer to the results of artificial labels. **Conclusion** A natural image segmentation method based on Gestalt rules is proposed. The method

收稿日期: 2015-01-20; 修回日期: 2015-05-08

基金项目: 国家自然科学基金项目(61165011, 61263046); 航天科技创新基金项目(CASC201102); 江西省教育厅科研项目(GJJ12427); 江西省青年科学基金项目(20132BAB211021)

第一作者简介: 曾接贤(1958—), 男, 教授, 1997年于西北工业大学获工学硕士学位, 主要研究方向为图像处理和计算机视觉。

E-mail: zengjx58@163.com

Supported by: National Natural Science Foundation of China (61165011, 61263046)

adopts oversegmentation and measures the region by applying the Gestalt rules that can effectively reduce the effect of natural factors in the segmentation process and improve segmentation accuracy. Experimental results show that the proposed image segmentation algorithm based on Gestalt rules exhibits better efficiency and accuracy than other algorithms, but the effect is not good for natural image with some sharp and thin object.

Key words: natural image; over-segmentation; Gestalt rules; region merging

0 引言

自然图像作为各类图像中最贴近实际应用的一类图像,在交通监控、军事安全、社交网络等领域都有广泛的应用。自然图像不同于人造图像与医学图像,它的构成规律性差,背景、纹理复杂,很容易受到光照、亮度变化等噪声的影响,使得它的分割随之变得越加困难。传统的自然图像分割方法根据策略不同可分为基于边缘、区域及特定理论等的分割方法。基于边缘的分割方法依赖于边缘检测算子提取的边缘,这些不连续边缘反应了图像在灰度、纹理、噪声等方面的位置,再将这些不连续的边缘连接成完整的边界,实现对图像的分割,如微分算子法,曲面拟合法, Hough 变换法等。基于区域的分割方法是把图像分裂成若干个像素的集合,主要是按照像素的相似性把像素划分到各个区域,如区域生长及分裂合并法、阈值法、聚类等。由于自然图像的分割技术与信息领域的其他学科,如数学、计算机科学、人工智能、模式识别等密切相关,因此随着这些学科中的新理论、新技术的产生,出现了不少结合特定理论的自然图像分割技术,如基于 Graph Cut、水平集和视觉显著性等的分割方法。基于 Graph Cut 的分割方法主要把图像看做是有向图来进行分割的方法,如 Shi 等人^[1]提出的 Ncut 算法,该算法融合了全局与局部信息,虽然消除了 cut 的孤立点,但随着图中节点数的增大,问题的求解也变得异常复杂,目前为止还没有精确求出分割结果的方法,实际应用中常采用近似的求解方法; Felzenszwalb 等人^[2]提出的一种高效的基于图的分割算法,与 Ncut 算法一样,这种算法也融合了非局部信息,使得图像在特征变化小的区域很好地保留图像的细节,而变化大的区域并不保留细节。基于水平集的分割方法是利用水平集函数的偏微分方程求解来获得分割的结果,如 Chunming 等人^[3]提出的距离正规化水平集算法,该算法由于引入了距离正规化的特性,使得水平集的

计算准确且稳定,但水平集应用在医学图像居多。基于视觉显著性的分割方法是根据模拟生物视觉注意机制的选择性来进行分割的方法,如 Cheng 等人^[4]从像素点和区域出发,提出了基于直方图对比度和区域对比度的显著性检测的方法,该方法是基于全局的显著性检测方法,克服了局部算法对显著目标较大时效果差的缺点; Perazzi 等人^[5]在对比度的基础上,设计了一个显著性过滤器,该过滤器能有效地计算和检测每个像素的显著图; Keyang 等人^[6]在空间先验性的基础上,提出了基于颜色和结构的显著性检测方法,该方法克服了以往只使用超像素和颜色对比度进行显著性检测的局限性。

这些基于视觉显著性的方法虽然取得了较好的效果,但视觉显著性是生物注意机制选择结果所产生的现象,它忽略了预注意阶段,也即注意机制是如何通过知觉组织来进行选择的。

知觉组织是指人对视觉输入信息进行组织的过程,它是人类获取环境信息的一个显著特性,人类在面对复杂场景时能迅速判断哪些信息组织在一起,哪些信息分开,而计算机视觉却很难判断,所以模拟人的知觉组织处理计算机视觉中的问题显得尤为重要。格式塔心理学家^[7]针对这一问题,提出了格式塔完形规则。将格式塔完形规则应用到计算机视觉,通过模拟生物知觉组织机理来处理计算机视觉中的问题成为了一种重要的手段,目前已有一些学者将该方法应用到自然图像分割中。如 Wang 等人^[8]提出了基于图模型的 Ratio-contour 算法,图的顶点用边缘中简化的边缘段表示,根据格式塔完形规则得到顶点之间的 Ratio-contour,再通过 Ratio-contour 进行闭合得到分割的结果。邓强等人^[9]在 Ratio-contour 的基础上将全局显著信息融合到轮廓编组中,提高了轮廓提取的准确性。

目前,这些基于格式塔完形规则进行图像处理的方法针对的是灰度图像,而且以检测到的边缘为基元进行后续处理,无法解决基于边缘的编组算法

所固有的问题,也即边缘极易受到图像中的纹理、噪声的影响,特别是背景上边缘像素较多时会影响分割的准确度。本文针对彩色图像,以预分割的区域为基元,提出了一种新的结合格式塔完形规则的自然图像分割方法。首先对图像进行预分割,此时的分割形成了若干个大小均匀的子区域,解决基于像素点或边缘片段易受光照、亮度等因素影响的问题。然后,引入格式塔完形规则对区域进行度量,提出基于区域的量化计算模型,得到区域率。最后,在区域率的基础上提出基于最大区域率的合并方法,通过区域合并得到最终的分割结果。算法流程图如图1所示。

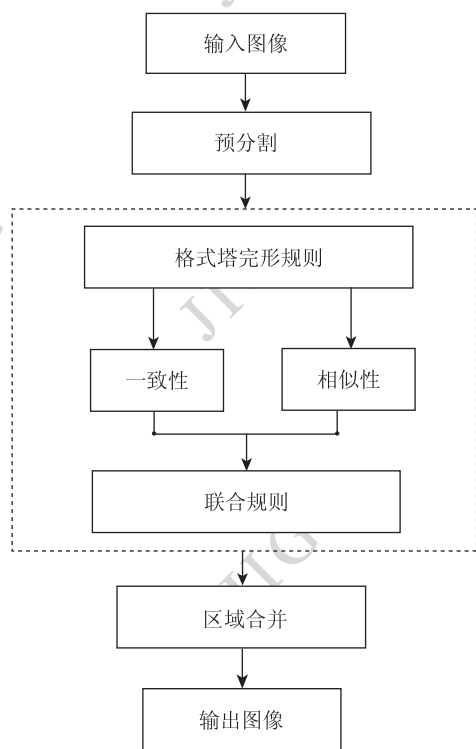


图1 算法流程图

Fig. 1 Flowchart of the proposed algorithm

本文算法的主要贡献如下:

1) 由于格式塔完形规则是模拟人类视觉感知得到的一系列规则,比较稳定,不易受光照、背景等因素的影响,而且将它应用到图像处理中,所得的结果也更加满足人类的视觉感知。所以本文引入格式塔完形规则对区域进行度量,提出了基于区域的量化计算模型,解决了传统的基于边缘的量化计算模型易受光照、背景等因素影响问题。

2) 通过格式塔完形规则对区域度量得到了一组区域率,在此基础上,提出了基于最大区域率合

并算法,该算法理论简单,且获得了较好的实验效果。

1 相关知识介绍

格式塔心理学的研究是以知觉心理为主。格式塔心理学^[7]认为,每一种心理现象都是一个被分离的格式塔整体,整体并不是各个部分的简单相加或组合,整体先于部分而存在,并制约着部分的性质和意义。其目的是将视觉场域中的组织因素或原理发掘出来,格式塔心理学家提出了许多知觉方法和规律。其中,最主要的就是格式塔的完形规则。

格式塔完形规则 (law of organization) 的主要一些规则如下:

1) 邻近性规则 (proximity), 即距离较近的各部分趋向于组成一个整体,如图2(a)所示。

2) 相似性规则 (similarity), 即某一方面相似的各部分趋向于组成一个整体,如图2(b)所示。

3) 平行性规则 (parallelism), 即物体平行的部分趋向于组成一个整体。

4) 连续性规则 (continuity), 即以一定规律呈现出连续性的部分趋向于组成一个整体。

5) 封闭性规则 (closure), 即彼此相属、构成封闭实体的各部分趋向于组成一个整体,如图2(c)所示。

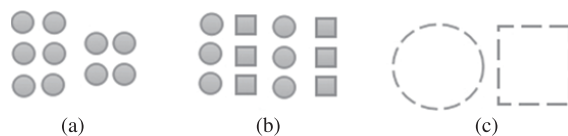


图2 格式塔完形规则

Fig. 2 Gestalt rules

2 本文算法

本文算法主要分3个步骤:预分割,基于格式塔完形规则的区域度量和基于最大区域率的合并算法。

2.1 预分割

由于自然图像易受到光照等因素的影响,因此,在对图像处理之前先进行预分割。目前,广泛应用于预分割的方法是超像素计算方法^[10]和分水岭分

割算法^[11]。由于分水岭算法的鲁棒性较差, 选用超像素的计算方法。超像素的计算方法有多种, 如基于图论的方法和基于梯度上升的方法。采用基于图论方法中的 Ncut 算法, 通过颜色和纹理的相似性构建一个加权图进行预分割。原因是 Ncut 算法能控制超像素的数目, 所得的超像素大小比较一致, 且能很好地保留边界信息。采用超像素方法将图像分割成若干个子区域, 由于这些子区域是由一系列位置相邻、颜色和纹理等特征相似的像素组成, 因此相比于边缘与像素点, 这些子区域不易受光照等外在因素的影响, 有利于获得更加精确的分割结果。

2.2 基于格式塔完形规则的区域度量

Ncut 预分割后形成了若干个子区域, 也即超像素, 接下来均采用区域来代替超像素。为了后面的区域合并, 需要对区域进行度量。区域度量是进行区域合并的一个重要步骤, 度量方法直接决定着区域合并的结果和图像的最终分割结果。传统的区域度量是通过距离来衡量区域之间存在的某种关系程度, 直接利用距离来衡量忽略了待分割的视觉特征, 使得区域的描述能力有限。因此, 从视觉特性出发, 引入格式塔完形规则对区域进行度量。格式塔完形规则是模拟人类视觉感知得到的一些规则, 包括邻近性、连续性、一致性、闭合性等规则。采用哪种规则以及怎样的量化计算对区域进行度量是本节要考虑的内容。

区域的合并是指将两个部分区域组成一个整体区域。对于区域合并的必要条件是: 两个区域在空间上相邻、颜色上相似、梯度方向上一致。而完形规则中的相似性是指某一个方面相似的部分趋于组成一个整体, 一致性指一定规律呈现出一致性的部分趋于组成一个整体。所以, 本文融合颜色的相似性、以及梯度方向的一致性来度量区域, 也即采用完形规则中的相似性和一致性对区域进行度量。将颜色作为相似性的特征进行描述, 灰度梯度方向作为一致性的特征进行描述。

2.2.1 颜色特征的计算

由于 HSV 颜色模型^[12]是一种直观的颜色模型, 和人类视觉感知很相似, 所以采用 HSV 颜色模型。在 HSV 颜色模型中, H 色调分量表示视觉感知的主导色模型, 可以模拟人对颜色的识别。因此, 选择了两区域的 H 分量颜色差作为相似性的特征描

述。具体计算如下:

1) 转换颜色空间。将彩色图像由 RGB 空间转换到 HSV 空间: 设在 RGB 空间中 (r, g, b) 分别是一个颜色的红、绿和蓝坐标, 它们的值是在 0 到 1 的实数。另外设 max 是 r, g 和 b 中的最大值, min 是这些值中的最小值。则在 HSV 空间中的 (h, s, v) 为

$$h = \begin{cases} 0^\circ & max = min \\ 60^\circ \times \frac{g-b}{max-min} + 0^\circ & max = r \text{ 且 } g \geq b \\ 60^\circ \times \frac{g-b}{max-min} + 360^\circ & max = r \text{ 且 } g < b \\ 60^\circ \times \frac{b-r}{max-min} + 120^\circ & max = g \\ 60^\circ \times \frac{r-g}{max-min} + 240^\circ & max = b \end{cases}$$

$$s = \begin{cases} 0 & max = 0 \\ \frac{max-min}{max} = 1 - \frac{min}{max} & \text{其他} \end{cases}$$

$$v = max \quad (1)$$

式中, $h \in [0, 360^\circ]$ 是角度的色相角, $s \in [0, 1]$, $v \in [0, 1]$ 分别是饱和度和亮度。

2) 计算颜色特征。在 H 通道统计每个区域的各个像素点的取值频率, 得到每个区域内各个像素点的期望值。将两区域的期望值差的绝对值作为颜色特征差。也即 $c^{sim} = |E_i - E_j|$, 其中, E_i 表示区域 i 像素颜色的期望值, E_j 表示区域 j 像素颜色的期望值, c^{sim} 表示颜色特征差。

2.2.2 梯度特征的计算

由于梯度方向不受光照影响, 且属于同一个对象的像素, 其梯度方向特征近似的相同, 而不同对象的梯度方向特征则会存在较大差异。因此, 选取区域的梯度差作为一致性的特征描述。具体计算如下:

1) 计算像素点的梯度角。使用 1 维离散微分模板 $[-1, 0, 1]$ 及其转置 $[-1, 0, 1]^T$ 对原始图像对应的灰度图像做卷积, 分别计算各像素点在水平方向和垂直方向的梯度分量 g_x, g_y , 利用公式 $\theta = \arctan(g_y/g_x)$ 计算各个像素点的梯度角。

2) 计算每个区域的梯度方向直方图。将 $0^\circ \sim 360^\circ$ 的梯度方向平均划分为 15 (一般是 9 ~ 18) 个直方图通道, 根据区域内像素点的梯度角确定该像素所属的直方图通道, 以属于该通道的像素点数量

作为该通道的值。

3) 计算梯度的方向特征。将每个区域的直方图中峰值最大的通道作为该区域的梯度角, 即梯度方向。两个区域梯度角差的绝对值作为梯度方向差。也即 $g^{\text{con}} = |\theta_i - \theta_j|$, 其中 θ_i 表示区域 i 的梯度方向, θ_j 表示区域 j 的梯度方向, g^{con} 表示梯度方向差。

2.2.3 基于区域的量化计算模型

根据格式塔完形规则选取了区域特征描述后, 还需要设计这些特征在相应规则下的量化计算模型。一致性和相似性联合作用下的概率为

$$\rho(e_i e_j | r_{\text{con}}, r_{\text{sim}}) = \alpha_1 \rho(e_i e_j | r_{\text{con}}) + \alpha_2 \rho(e_i e_j | r_{\text{sim}}) \quad (2)$$

式中, $\rho(e_i e_j | r_{\text{con}})$ 表示区域 e_i, e_j 在一致性条件下属于同一个区域的概率计算公式, $\rho(e_i e_j | r_{\text{sim}})$ 表示区域 e_i, e_j 在相似性条件下属于同一个区域的概率计算公式。为了描述方便, 本文将两区域属于同一区域的概率简称为区域率。 $e_i e_j, r_{\text{con}}, r_{\text{sim}}$ 分别表示是一个事件, $e_i e_j$ 表示区域 e_i, e_j 属于同一个区域, r_{con} 表示一致性, r_{sim} 表示相似性。 α_1, α_2 表示权值。根据贝叶斯公式得

$$\rho(e_i e_j | r_{\text{con}}) = \frac{\rho(r_{\text{con}} | e_i e_j) \rho(e_i e_j)}{\rho(r_{\text{con}})} \quad (3)$$

$$\rho(e_i e_j | r_{\text{sim}}) = \frac{\rho(r_{\text{sim}} | e_i e_j) \rho(e_i e_j)}{\rho(r_{\text{sim}})} \quad (4)$$

根据全概率公式, $\rho(r_{\text{con}}), \rho(r_{\text{sim}})$ 的概率估计又可以由区域 e_i, e_j 属于同一区域与其对立事件区域 e_i, e_j 不属于同一区域这两种情况考虑, 即

$$\rho(r_{\text{con}}) = \rho(r_{\text{con}} | e_i e_j) \rho(e_i e_j) + \rho(r_{\text{con}} | \overline{e_i e_j}) \rho(\overline{e_i e_j}) \quad (5)$$

$$\rho(r_{\text{sim}}) = \rho(r_{\text{sim}} | e_i e_j) \rho(e_i e_j) + \rho(r_{\text{sim}} | \overline{e_i e_j}) \rho(\overline{e_i e_j}) \quad (6)$$

式中, $\overline{e_i e_j}$ 是 $e_i e_j$ 的对立事件, 表示区域 e_i, e_j 不属于同一区域。将式(5)(6)代入式(3)(4), 再将式(3)(4)代入式(2)可得本文联合概率的计算公式, 为了方便描述接下来将用 $\rho(e_i e_j)$ 表示 $\rho(e_i e_j | r_{\text{con}}, r_{\text{sim}})$, 也即本文的区域特征在一致性和相似性联合作用下区域率的计算公式, 即

$$\rho(e_i e_j) = \alpha_1 \frac{1}{1 + \frac{\rho(r_{\text{con}} | \overline{e_i e_j})}{\rho(r_{\text{con}} | e_i e_j)} \cdot \frac{\rho(\overline{e_i e_j})}{\rho(e_i e_j)}} +$$

$$\alpha_2 \frac{1}{1 + \frac{\rho(r_{\text{sim}} | \overline{e_i e_j})}{\rho(r_{\text{sim}} | e_i e_j)} \cdot \frac{\rho(\overline{e_i e_j})}{\rho(e_i e_j)}} \quad (7)$$

式中, $\frac{\rho(\overline{e_i e_j})}{\rho(e_i e_j)} = \frac{N/10 - 2}{2}$, N 表示预分割得到的区域个数^[13]。而 $\rho(r_{\text{con}} | e_i e_j), \rho(r_{\text{con}} | \overline{e_i e_j})$ 和 $\rho(r_{\text{sim}} | e_i e_j), \rho(r_{\text{sim}} | \overline{e_i e_j})$ 分别由一致性和相似性的量化计算模型所得。也即当区域 e_i, e_j 属于同一区域时, 其一致性与相似性量化计算模型均采用正态分布, 即

$$\rho(r_{\text{con}} | e_i e_j) = \frac{1}{\sqrt{2\pi}\sigma^{\text{con}}} e^{\left\{-\frac{(g^{\text{con}})^2}{2(\sigma^{\text{con}})^2}\right\}} \quad (8)$$

$$\rho(r_{\text{sim}} | e_i e_j) = \frac{1}{\sqrt{2\pi}\sigma^{\text{sim}}} e^{\left\{-\frac{(c^{\text{sim}})^2}{2(\sigma^{\text{sim}})^2}\right\}} \quad (9)$$

式中, $\sigma^{\text{con}} = 1/\pi, \sigma^{\text{sim}} = 1/8, g^{\text{con}}, c^{\text{sim}}$ 分别表示区域 e_i, e_j 的梯度差和颜色差^[14]。参考文献[14], 给出当区域 e_i, e_j 不属于同一区域时, 一致性与相似性的计算模型为

$$\rho(r_{\text{con}} | \overline{e_i e_j}) = 1/\pi \quad (10)$$

$$\rho(r_{\text{sim}} | \overline{e_i e_j}) = c^{\text{sim}}/256 \quad (11)$$

2.3 基于最大区域率的合并算法

为了得到分割的最终结果, 需要对预分割的区域进行合并。本文是通过格式塔完形规则的量化计算模型所得到的区域率, 也即 $\rho(e_i e_j)$ 作为合并的输入。很显然 $\rho(e_i e_j)$ 越大, 两个区域属于同一个区域的可能性就越大。这样就需要构造一个合理的算法进行合并。同时, 希望该算法简单, 执行效率高。参考文献[15]的区域合并算法, 提出基于最大区域率的合并算法。算法步骤如下:

1) 构建相邻区域。对于每个区域 $e_i (i=1, 2, \dots, N)$, 首先构建它的相邻区域 $e_j (j=1, 2, \dots, m)$; 再构建 e_j 相邻区域 $e_z (z=1, 2, \dots, n)$ 。显然 $i \subseteq z$, 其中, N 表示总区域的个数, 由初始分割得到; m, n 表示相邻区域数, 根据不同的区域得到不同的区域个数。

2) 区域合并。根据式(12)判断相邻区域 e_i 与 e_j 是否合并, 根据式(13)判断合并是否停止。其中 $\rho(e_i e_j)$ 与 $\rho(e_j e_z)$ 由式(7)求得; T 为给定的阈值。

$$\begin{cases} e_i = e_i \cup e_j & \rho(e_i e_j) = \max_{z=1, 2, \dots, n} \rho(e_j e_z) \\ e_i \neq e_i \cup e_j & \text{其他} \end{cases} \quad (12)$$

$$\begin{cases} \text{继续合并} & \rho(e_i e_j) \geq T \\ \text{合并停止} & \rho(e_i e_j) < T \end{cases} \quad (13)$$

3)更新图像,重复步骤1)2)直到图像中不再有区域合并为止。

3 算法实现步骤

综上所述,本文算法步骤如下:

1)对原图像采用 Ncut 算法进行预过分割,得到区域 $e_1, e_2, \dots, e_N$ 。这些区域中包含了一些极小的区域,如图 3(b)所示。

2)根据式(14)删除初始分割中的极小区域,如图 3(c)所示,并采用绿色的圈标出图 3(b)(c)中一些对应位置的不同之处。

$$\begin{cases} \text{删除区域 } e_i & \text{Area}_{e_i} \leq \frac{1}{10} \times \frac{m \times m}{N} \\ \text{保留区域 } e_i & \text{其他} \end{cases} \quad (14)$$

式中, m, n 分别表示原图像的长与宽, Area_{e_i} 表示区域 e_i 的像素个数, N 是 Ncut 算法得到的区域总个数。由于初始分割中的极小区域对整体的分割影响不大,删除后可以提高后续步骤的计算效率。

3)根据式(7)分别计算一致性和相似性,即通过灰度通道计算式(7)等号右边的第1项, H 通道计算式(7)等号右边的第2项。

4)通过式(12)判断相邻的两个区域是否合并,再通过式(13)判断合并是否停止。

5)更新图像,重复步骤2)3)4)直到图像中不再有区域合并为止,输出图像,如图 3(d)所示。

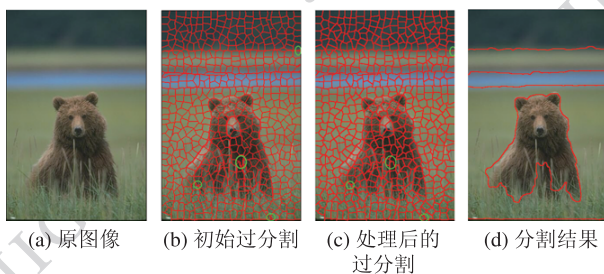


图3 本文算法的分割流程图

Fig. 3 Flowchart of the segmentation algorithm

4 实验结果及分析

为了验证本文算法的有效性,本文用 Matlab 软件进行了实验,并与 Ncut^[1]、文献[16-18]、MeanShift^[19],以及人工分割的结果进行对比。实验环境为 2.59 GHz 的 CPU, 1.99 GB 的 RAM。

4.1 实验参数的选取

4.1.1 对比算法的参数选取

Ncut 算法主要是分割数 K 的选取,选取 $K = 20$ 进行实验对比。因为此时的分割结果与人的主观感知比较一致^[17]。而 MeanShift 算法分割结果主要受带宽(空间带宽 h_s 和色度带宽 h_r)的影响。 (h_s, h_r) 一般在 $([7, 16], [3, 23])$ 范围内选取,研究^[17]发现 Berkeley 分割数据集 $(h_s, h_r) = (13, 19)$ 能够很好地进行对比实验,此外 MeanShift 算法有时候受到 MinimumRegion 影响。选 $(h_s, h_r, M) = (13, 19, 20)$ 。 M 表示 MinimumRegion。文献[16]算法是融合颜色直方图和 K 均值聚类的分割方法,分割的结果主要受到分类数 K_1, K_2 及归一化常数 κ 这 3 个参数的影响, K_1, K_2 的最佳取值区间为 $[6, 13]$, κ 在 $[0.10, 0.15]$ 。文献[16]指出当 $(K_1, K_2, \kappa) = (13, 6, 0.135)$ 时效果最好,因此,选 $(K_1, K_2, \kappa) = (13, 6, 0.135)$ 。文献[17]算法是通过数据压缩的无监督自然图像分割方法,分割的结果主要受给定的阈值 γ 的影响,从文献[17]的分割性能评价指标可知,当 $\gamma = 0.1$ 时,分割效果最好,所以选 $\gamma = 0.1$ 。文献[18]算法是基于多分辨率的非线性降维模型的彩色图像分割方法,该方法主要在 K-means 和 meanshift 这 2 个算法中进行实验,从实验结果看,该方法在 K-means 算法上降维的实验效果较好,在 K-means 算法上降维主要受分类数 K 和阈值 ξ 的影响,实验中选取的 $K = 11, \xi = 0.4$ 效果好,因此针对文献[18]算法选取的参数为 $K = 11, \xi = 0.4$ 。

4.1.2 本文算法参数的选取

如图 4 所示,分别显示了超像素 N 为 300、500、1000 分割图,当 $N = 500$ 时所得的过分割图基本包含了分割所需的信息,而 $N = 300$ 时包含的细节信息太少, $N = 1000$ 虽然能将所有的细节揽括,但大大降低了后续的分割效率。因此,选取 $N = 500$ 。另外,假设每个规则都同等重要,也即取权值 $\alpha_1 = \alpha_2 = 1/2$ 。阈值 T 是通过大量实验求出的经验值。通过对 10 组自然图像进行测试,统计同一阈值 T 时的平均分割区域数。发现当阈值过高时,造成了过分割,这是因为通过 2.2.3 节计算的区域率的值在 $0 \sim 1$ 之间,而 2.3 节合并的条件是相邻两个区域中的最大概率值且大于阈值 T ,当阈值过高时满足合并的条件较少,故造成了过分割;而当阈值过低时,满足的条件又较多,故造成了欠分割,如图 5 所示。

图 6 给出了一组 T 取不同值的结果图,从图 6 可知当阈值 T 偏低时,背景与前景进行了融合,如图 6(a) 所示;当阈值 T 选得偏高时,前景分割成了几个区域块,背景分割成了几个区域,如图 6(c)(d) 所示;实验中的最佳阈值是在 0.45 左右,如图(b)所示,此时的分割结果基本将背景与前景分开,且背景与前景单独成一个区域,所以实验中选取的 T 值为 0.45。

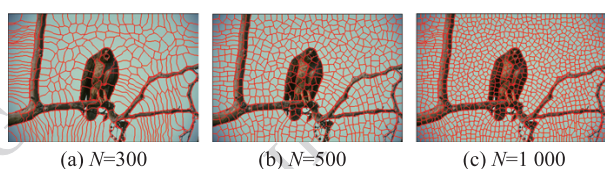


图 4 超像素分割图

Fig. 4 Segmentation map of superpixels

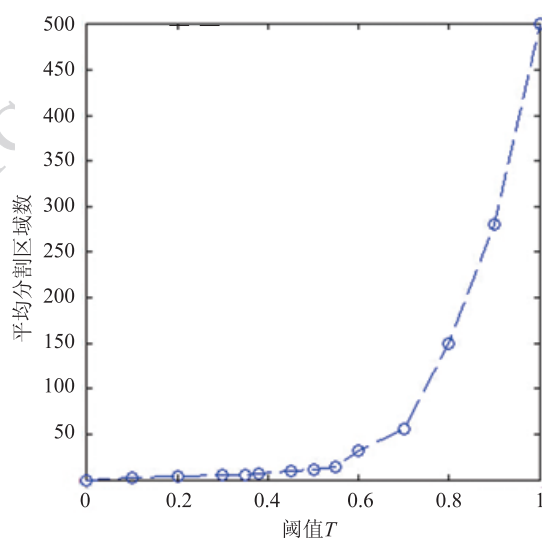
图 5 分割区域数与阈值 T 的关系图

Fig. 5 the diagram of the number of the divided regions and the threshold value

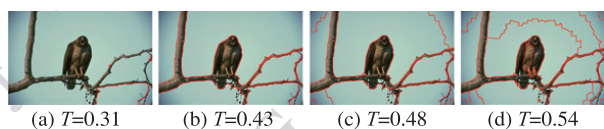
图 6 不同阈值 T 的结果图

Fig. 6 the results of different threshold

4.2 实验对比及结果分析

4.2.1 定性分析

将本文算法与 Ncut^[1] 算法、MeanShift^[19] 算法、以及文献[16-18]算法进行比较,部分结果如图 7 所示。实验图像来自 Berkeley 分割数据库中的 300 幅图像。Berkeley 分割数据库是常用的自然图像分割

的数据库之一;每幅图像均受不同因素的影响,同时通过人工标注出分割的区域。图 7(a) 给出了 5 幅分别受阴影、光照、背景模糊、运动等因素影响的图像。图 7(b) 所示为 MeanShift 算法得到的分割结果,由于算法中选取参数带宽值的难度较大,分割的结果易造成过分割或欠分割;图 7(c) 所示为 Ncut 算法得到的分割结果,分割的结果都不够精确,这是因为 Ncut 算法是个 NP-hard 问题,目前为止还没有精确求出分割结果的方法;图 7(d) 所示为文献[16]算法得到的分割结果,由于算法在融合 K 均值聚类时种子点比较难确定,而且基于颜色的直方图并不能反映图像的空间信息,使得分割的结果不够精确;图 7(e) 所示为文献[17]算法的分割结果,由于该算法受到给定的阈值 γ 影响, γ 取值高,容易造成欠分割或不分割, γ 取值低,又造成过分割,如图 7(e) 所示分割的结果造成了过分割。图 7(f) 所示为文献[18]算法在 K-means 算法上降维的分割实验结果,由于实验结果受到分类数 K 和阈值 ξ 的影响,分割的结果也造成了过分割。图 7(g) 所示为本文算法的分割结果,与 MeanShift 算法、Ncut 算法及文献[16-18]算法相比,由于本文算法预处理采用的是过分割,使得分割的结果不易受光照、阴影、模糊等因素的影响,同时本文采用格式塔完形规则对区域进行度量,使得分割的结果更加满足人的视觉系统,与图 7(h) 中人工标注的结果比较接近。

4.2.2 定量比较

从 Berkeley 分割图像库随机抽取了 30 幅图像进行实验,同时采用 PRI(probabilistic rand index)、VOI(variation of information)和 GCE(global consistency error)评价指数^[16]进行评价。PRI、VOI 和 GCE 均表示待评价的分割结果与每一个手工标注分割结果之间的距离或相似度,取均值作为最终的评价指数。其中, PRI 是计算分割结果的标记与真实标记相一致的像素对的比例,其值越大,表示的分割效果就越好,取值范围区间为 $[0, 1]$ 。VOI 是基于像素点和它的聚类之间的关系,计算像素点从一个聚类变化到另一个聚类丢失或获得了多少信息,其值越小,表示的分割效果越好,取值范围在 $[0, \infty)$ 。GCE 是衡量一个分割结果可以被看做另一个分割结果的子集程度,其值越小,表示分割效果越好,取值范围在 $[0, 1]$ 。

表 1 记录了每种算法的 PRI、VOI 和 GCE 的均

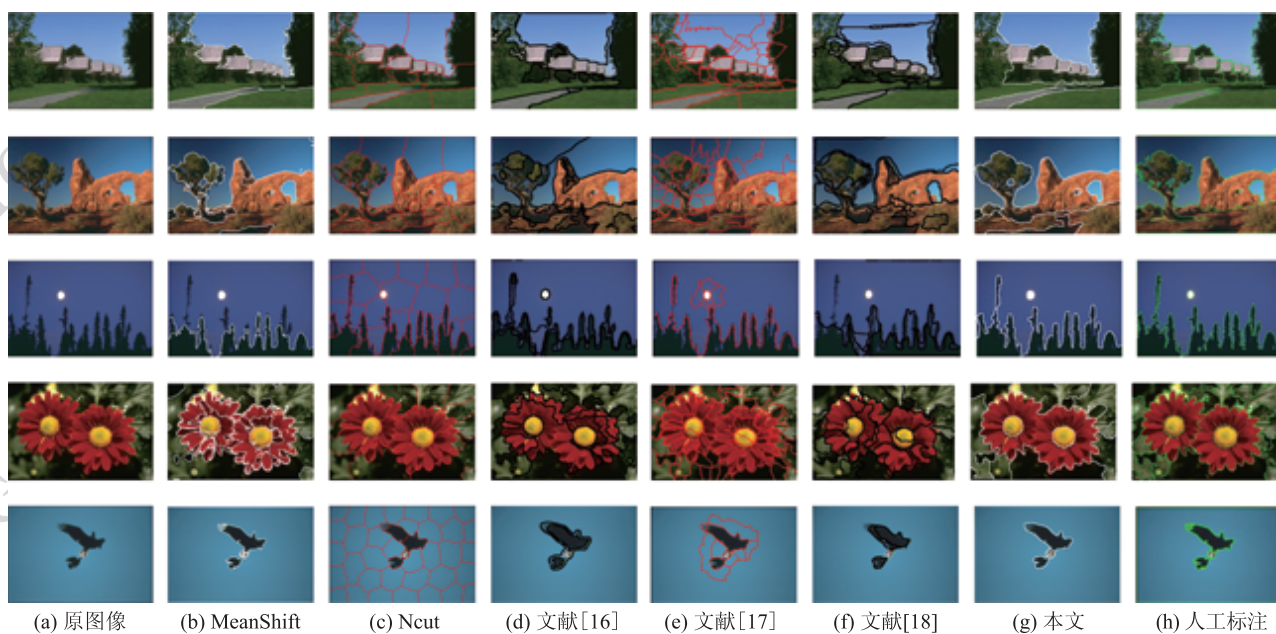


图7 分割结果对比

Fig. 7 Comparing the results of segmentation

值。从表1可以看出本文算法的分割结果在VOI和GCE这2个性能指标上要明显优于对比实验算法。在PRI性能上次之,这是由于PRI是计算分割结果的标记与真实标记相一致的像素对比例,由于本文算法分割结果的区域数较少,可能造成与真实标记相一致的像素对相对较少,使得比例值要低于文献[16,18]算法。但本文算法总体性能指标与手工标记的结果比较接近。

表1 几种算法的3个分割性能指标比较

Table 1 Three segmentation performance comparison of several algorithm

分割算法	PRI 均值	VOI 均值	GCE 均值
手工标记分割 ^[18]	0.87	1.10	0.08
MeanShift	0.75	2.48	0.26
Ncut	0.72	2.93	0.22
文献[16]	0.79	2.30	0.21
文献[17]	0.76	2.46	0.18
文献[18]	0.78	2.36	0.23
本文	0.77	1.98	0.18

4.2.3 算法运行时间比较

本文算法运行时间主要包括两部分:一是生成超像素图所运行的时间;二是区域合并所运行的时

间。表2显示了这4种算法对实验图像运行的平均时间。从表2可以可知本文算法总的运行时间虽然高于经典的MeanShift算法和Ncut算法,但低于文献[16]算法。

表2 算法运行时间比较

Table 2 Compares the algorithm running time

	本文算法		文献[16]	MeanShift	Ncut
	超像素图	区域合并	算法	算法	算法
执行时间/s	401.04	29.14	478.54	36.25	386.42

5 结 论

提出了一种新的结合格式塔完形规则的自然图像分割方法。为了弱化光照、背景等因素对自然图像的影响,以区域为基元进行处理,同时采用格式塔完形规则对区域进行度量,提出了基于区域的量化计算模型。格式塔完形规则是模拟人类视觉感知得到,所以所得的区域率更加满足人的视觉感知。为了提高分割的执行效率,提出了基于最大区域率的合并算法。通过实验结果分析与对比,本文算法不仅能够很好地将格式塔完形规则应用到图像分割中来,而且相比于对比算法,效果也更加满足人的视觉感知。当然也存在一些不足,如分割结果依赖于预

分割结果。本文假设区域合并不影响分割结果,则预分割越精确,分割结果就越精确。当预分割精度达到一定值时,即预分割的区域不同时包含前景与背景,此时随着预分割精度增大,对分割结果影响不大,但相应的运行速度也越慢。综合考虑分割精度与运行速度,最终选取了 $N = 500$ 的分割结果作为预分割结果,而此时的分割结果不能分割出一些具有尖细的物体,如鸟的翅膀。另外,本文的区域合并停止采用的是阈值法,而阈值的选取是通过大量实验求得的经验值,未来的工作希望能找到一种高效的区域停止算法。

参考文献 (References)

- [1] Shi J, Malik J. Normalized cuts and image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8): 888-905.
- [2] Felzenszwalb P F, Huttenlocher D P. Efficient graph-based image segmentation[J]. International Journal of Computer Vision, 2004, 59(2): 167-181.
- [3] Li C, Xu C Y, Senior M, et al. Distance regularized level set evolution and its application to image segmentation[J]. IEEE Transactions on Image Processing, 2010, 19(12): 3243-3254. [DOI:10.1109/TIP.2010.2069690]
- [4] Cheng M M, Zhang G X, Mitra N J, et al. Global contrast based salient region detection[J]. Computer Vision and Pattern Recognition, 2011: 409-416.
- [5] Perazzi F, Krahenbuhl P, Pritch Y, et al. Saliency Filters: contrast based filtering for salient region detection[J]. Computer Vision and Pattern Recognition, 2012: 733-740.
- [6] Shi K Y, Wang K Z, Lu J B, et al. PISA: pixelwise image saliency by aggregating complementary appearance contrast measures with Spatial Priors[J]. Computer Vision and Pattern Recognition, 2013: 2115-2122.
- [7] Kurt K. Gestalt Psychology Principle [M]. Li Wei translated. Beijing: Peking University Press, 2010. [库尔特·考夫卡, 格式塔心理学原理[M]. 李维, 译. 北京: 北京大学出版社, 2010.]
- [8] Stahl J S, Wang S. Globally optimal grouping for symmetric closed boundaries by combining boundary and region information[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(3): 395-411. [DOI:10.1109/TPAMI.2007.1186]
- [9] Deng Q, Luo Y P, Ge J F. Put global significant information into contour grouping model[J]. Computer Aided Design and Graphics Journal, 2013, 25(8): 1223-1229. [邓强, 罗予频, 葛俊峰. 融合全局显著信息的轮廓编组模型[J]. 计算机辅助设计与图形学学报, 2013, 25(8): 1223-1229.]
- [10] Dickinson S J, Levinshtein A, Sminchisescu C. Perceptual grouping using superpixels[J]. Pattern Recognition, 2012, 7329: 13-22. [DOI:10.1007/978-3-642-31149-9_2]
- [11] Zhang G F, Wu Z C, Yi L N. Improved marker-based watershed algorithm for segmentation of high spatial resolution remote sensing imagery[J]. Computer Engineering and Applications, 2010, 27(2): 760-763.
- [12] Li S H, Gao G Z. The application of improved HSV color space model in image processing[C]//Proceedings of the 2nd International Conference on IEEE Future Computer and Communication. Wuhan, China: IEEE Press, 2010, 2(V2): 10-13. [DOI:10.1109/ICFCC.2010.5497299]
- [13] Zhou Q. Research on computational model of contour grouping and attention[D]. Beijing: Beijing Jiaotong University, 2006. [邹琪. 轮廓编组的计算模型和注意模型研究[D]. 北京: 北京交通大学, 2006.]
- [14] Bo Y H, Luo S W, Zou Q. Computational model of significant edge detection on the target contour in natural images[J]. Pattern Recognition and Artificial Intelligence, 2010, 23(1): 752-758. [薄一航, 罗四维, 邹琪. 自然图像中目标轮廓上显著边缘检测的计算模型[J]. 模式识别与人工智能, 2010, 23(1): 752-758.]
- [15] Ning J F, Zhang L, Zhang D. Interactive image segmentation by maximal similarity based region merging[J]. Pattern Recognition, 2010, 43: 445-456. [DOI:10.1016/j.patcog.2009.03.004]
- [16] Mignotte M. Segmentation by fusion of histogram-based K-means clusters in different color spaces[J]. IEEE Transactions on Image Processing, 2008, 17(5): 780-787. [DOI:10.1109/TIP.2008.920761]
- [17] Yang A Y, Wright J, Ma Y, et al. Unsupervised segmentation of natural images via lossy data compression[J]. Computer Vision and Image Understanding, 2007, 110(2008): 212-225.
- [18] Mignotte M. Based multiresolution nonLinear dimensionality reduction model for color image segmentation[J]. IEEE Transactions on Neural Networks, 2011, 22(3): 447-460. [DOI:10.1109/TNN.2010.2101614]
- [19] Comaniciu D, Meer P. Mean shift: a robust approach toward feature space analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5): 603-619. [DOI:10.1109/34.1000236]