

基于区域中心的交互式图像前景提取方法

汪 昊,刘向阳

(河海大学理学院,江苏南京 211100)

摘要:图像分割是从图像中提取有意义的区域,是图像处理和计算机视觉中的关键技术。而自动分割方法不能很好地处理前景复杂的图像,对此提出一种基于区域中心的交互式图像前景提取算法。针对图像前景的复杂度,很难用单一的相似区域描述前景,文中采用多个区域中心来刻画目标区域。为提升图像分割的稳定性,给出基于超像素颜色、空间位置和纹理信息的相似性度量方法;为确保图像分割区域的连通性和准确性,定义了基于超像素的测地距离计算方法。使用基于测地距离的超像素局部密度,来分析图像的若干区域中心;基于用户交互的方式来分析前景的区域中心,得到图像前景。经过大量彩色图像的仿真表明,在分割过程中利用少量的用户交互信息,可有效提升图像分割的稳定性和准确性。

关键词:交互式图像分割;局部密度;区域中心;超像素

中图分类号:TP391.41

文献标识码:A

文章编号:1673-629X(2020)02-0012-05

doi:10.3969/j.issn.1673-629X.2020.02.003

An Interactive Image Foreground Extraction Method Based on Region Center

WANG Hao, LIU Xiang-yang

(School of Science, Hohai University, Nanjing 211100, China)

Abstract: Image segmentation is the extraction of meaningful regions from images and is a key technology in image processing and computer vision. The automatic segmentation method can't handle the complex foreground image well, so we propose an interactive image foreground extraction algorithm based on region center. For the complexity of image foreground, it is difficult to describe the foreground with a single similar area. We use multiple regional centers to describe the target area. In order to improve the stability of image segmentation, a similarity measure method based on superpixel color, spatial position and texture information is given. To ensure the connectivity and accuracy of image segmentation region, a geodesic distance calculation method based on superpixel is defined. The local density of the superpixel based on the geodesic distance is used to analyze several regional centers of the image; the regional center of the foreground is analyzed based on the user interaction to obtain the image foreground. The simulation of a large number of color images shows that a small amount of user interaction information is used in the segmentation process, which can effectively improve the stability and accuracy of image segmentation.

Key words: interactive image segmentation; local density; regional center; superpixel

0 引言

图像分割是指从图像中提取有意义的对象或者感兴趣的区域,它是图像处理和计算机视觉的一种关键技术^[1-2]。由于图像的类型丰富多样,并且图像的复杂度不一,自动分割方法针对前景或背景复杂的图像往往没有好的分割效果,同时计算机提取大量轮廓没有优势,而手工分割人为操作量大,并且定位边界时有误差,因此需要少量的人工干预来提取前景目标或者

感兴趣的区域,使得分割更加容易^[3]。交互式图像分割是指,首先用户确定感兴趣的目标或者前景并做上记号(一般用鼠标在前景和背景处各画几笔或者用一个围盒框住目标或前景区域),然后算法将交互信息作为条件,最终分割出满足用户交互的目标区域^[4]。交互式分割的目的是能够更好地提取感兴趣的对象,其目标区域基于用户的交互指示。经典方法和技术有Livewire^[5]、Grab Cut^[6]、Lazy snapping^[7]、Geodesic^[8]等。

收稿日期:2019-03-15

修回日期:2019-07-15

网络出版时间:2019-09-25

基金项目:国家自然科学基金(61001139)

作者简介:汪 昊(1993-),男,硕士研究生,CCF会员(B1074G),研究方向为图像处理;刘向阳,通讯作者,副教授,硕导,研究方向为图像与视频分析、数据分析和机器学习。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20190925.1525.076.html>

Livewire 方法^[5]是用户在边界上选择一个初始点,在通道中移动光标,计算机基于初始点到当前点的最佳路径自动选择后续点。但是该方法依赖目标物体的边界,处理弱边界物体过程繁琐,并且处理高分率图像效率低。Grab Cut^[6]方法, Rother 等^[9]在 Boykov 等^[10]的基础上提出的通过迭代过程扩展交互式图切割算法,处理 RGB 空间中的彩色图像,使用 GMM^[6]分别模拟对象和背景的颜色分布。用户需要用个包围盒将目标区域框住,但如果围盒中的区域不足以表达目标区域特性时,分割结果不准确。Lazy snapping^[7]技术,是在最大流/最小割算法基础上提出的。首先用分水岭算法对图像进行预分割,图像被过分割成很多小区域,然后通过最大化对象内的颜色相似性和跨越对象和背景之间的边界的梯度量来优化对象边界。但是该方法容易分割出小目标。Geodesic^[8]方法,首先用户提供交互的自由点,然后基于测地距离计算属于前景和背景的概率。但是该方法在边界上没有很好的模型。这些交互式算法都是基于单个像素进行图像分割的,可能导致计算结果不稳定,并且传统的交互式分割图像方法需要画刷在背景与前景处画多笔或者画一个方框,用户交互较多,操作起来不简便。

针对上述问题,文中提出了一种简单快速的基于区域中心的交互式图像前景提取方法。因为图像的复杂度不一,按照某种相似特征(例如颜色、亮度、纹理等)对图像进行划分,图像会被分成若干个区域,所以会出现前景或者目标区域被分成了若干个小区域的情况,从而很难用单一的区域形成前景或者目标区域。而子区域是根据某种相似特征形成的,单一的子区域都有相应的子区域中心,类似于区域聚类中心。因此文中采用基于超像素颜色、空间位置和纹理信息的相似性度量方法,使用基于测地距离的超像素局部密度和距离,分析计算图像的所有子区域中心;其次利用用户交互的信息分析属于前景或者目标区域的所有子区域中心,最后形成前景子区域并合并所有子区域得到分割结果。实验结果表明该算法简单可行,稳定性好。

1 算法思路

该方法大致分为四步进行,流程如图1所示。因为超像素可以保留大部分兴趣范围内分割所需的大部分结构^[11],同时降低后续处理数据的复杂度^[12],所以第一步先对图像进行超像素分析。然后提取超像素的五维特征属性,即颜色均值、空间位置、纹理信息,建立相邻超像素的距离公式,构建近邻超像素距离矩阵。再基于近邻距离矩阵,采用测地距离,构建任意两超像素之间的距离矩阵。第二步基于局部密度,分析距离矩阵确定图像的区域中心。第三步根据用户对前景或

者感兴趣的目标区域的交互信息,得到交互像素所在的超像素,确定这些超像素所属的区域中心。第四步将属于这些区域中心的所有超像素形成子区域,最后合并所有子区域,形成前景或者感兴趣的目标。

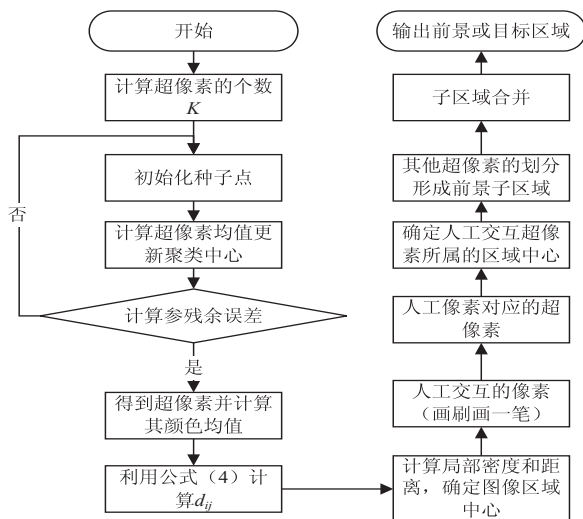


图1 算法流程

2 算法描述

2.1 生成超像素

超像素通常作为图像分割的预处理过程。超像素定义为在某种特征属性下具有很大相似性的像素点的集合,并且这些像素点通常在空间位置中是近邻像素点。其优势是可以相对较少的数量来代替大量像素点所具有的图片特征,保留了在兴趣范围内分割所需的大部分结构^[11],降低了后续图像计算数据的复杂度^[12]。2012年 Achanta 等^[13]提出了简单的线性迭代聚类算法(simple linear iterative clustering, SLIC)。不管是分割的速度,还是分割效果,SLIC 算法都优于其他的超像素算法。首先将彩色图像数值化,转换为 CIELAB 颜色空间下的数值 $[Lab]$,然后结合空间坐标位置 $[xy]$,用5维特征向量 $[Labxy]^T$ 进行聚类,形成超像素。大致步骤如下:

Step1:初始化种子点,即在图像内均匀分配种子点。

Step2:在每个超像素的中心(即初始种子点)的 3×3 邻域内找到一个颜色梯度最小的种子位置,将种子点移到该位置上。

Step3:分配种子点邻域内像素点的标签,形成超像素。

2.2 构建超像素距离矩阵

经过图像预处理后,以超像素为基本单位^[14],构建近邻超像素距离矩阵。首先建立相邻超像素的距离公式,将超像素的特征信息进行数值化,即颜色、空间、纹理特征的数值化。颜色信息用超像素在彩色空间下

的颜色均值表示,记为 l_i, a_i, b_i ; 空间信息用笛卡尔坐标系下超像素聚类中心的位置坐标表示,记为 x_i, y_i ; 纹理信息用超像素的颜色均值的方差来表示,记为 dl_i, da_i, db_i 。综上,每个超像素用数值化的特征属性向量表示,记为 $[l_i a_i b_i x_i y_i dl_i da_i db_i]^T$ 。两个超像素越相似,它们之间的颜色信息值、纹理信息值的差值越小,距离就越小。空间坐标信息的作用是保证两超像素相邻,构建超像素近邻图。得到每个超像素的特征向量 $[l_i a_i b_i x_i y_i dl_i da_i db_i]^T$ 后,计算两两相邻超像素之间的距离,形成近邻矩阵 $d = (d_{ij})_{m \times m}$,其中 d_{ij} 定义如下:

$$d_1 = \sqrt{(l_i - l_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2} \quad (1)$$

$$d_2 = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (2)$$

$$d_3 = \sqrt{(d_{l_i} - d_{l_j})^2 + (d_{a_i} - d_{a_j})^2 + (d_{b_i} - d_{b_j})^2} \quad (3)$$

$$d_{ij} = d_1 + \gamma d_2 + \alpha d_3 \quad (4)$$

然后基于近邻矩阵,进一步构造所有超像素之间的距离矩阵。在近邻图中,每两个相邻元素的距离为 $d = (d_{ij})_{m \times m}$,不相邻的两个元素之间的距离为无穷大。因为测地距离刻画了任意两超像素之间的真实不相似性,所以在距离矩阵中采用测地距离近似两不相邻元素之间的距离。测地距离的使用间接保证了图的区域连通性,而不是只使用超像素间的欧氏距离。其中 γ 为中心位置权重, α 为纹理特征权重。

2.3 确定图像区域中心

2014 年 Alex Rodriguez 等^[15]提出了基于密度峰值的聚类算法。算法(DPCA)的主要思想是先确定聚类中心点(文中的图像区域中心超像素),然后对非聚类中心数据点(文中的非区域中心的超像素)进行归类,文中将超像素当作数据点。一个数据点可以成为聚类中心需要满足两个特性:首先,这个数据点在固定的范围内具有较大的密度;其次,相对于其他局部密度更大的数据点之间的距离相对较大。第二个特性保证了在同范围内局部聚类中心的唯一性,区别了在同范围内其他只是高密度的数据点而不是聚类中心点。基于聚类中心的特性,算法遍历所有的超像素,得到相应的局部密度 ρ_i 值和“距离” δ_i 值(即指本数据点与具有更高局部密度点的距离,也可叫位移偏量)。

为了计算局部密度 ρ_i 值和“距离” δ_i 值,先建立超像素数集合 $S = \{C_i\}_{i=1}^m$,其中 $I_s = \{1, 2, \dots, m\}$ 为相应的指标集, C_i 表示相应的超像素的特征属性, D_{ij} 表示超像素 C_i 和超像素 C_j 之间的距离,即距离矩阵中的距离(相邻的超像素其定义如式 4,不相邻超像素之间的距离为它们之间的测地距离)。

对于每个超像素 C_i ,可以为其定义 ρ_i 和 δ_i :

$$\rho_i = \sum_{j \in I_s \setminus \{i\}} \chi(d_{ij} - d_c) \quad (5)$$

其中函数

$$\chi(x) = \begin{cases} 1, & x < 0 \\ 0, & x \geq 0 \end{cases} \quad (6)$$

参数 d_c 表示截断距离,截断距离的设置可以改变一个数据点在范围内包含的数据点占总数据点的百分比。实验表明百分比设置为 1% 到 2%,聚类效果最佳。即局部密度的含义为距离小于截断距离的数据点的个数,这里为超像素的个数。

$$\delta_i = \begin{cases} \min_j(d_{ij}), \rho_j > \rho_i \\ \max_j(d_{ij}), \forall j \in I_s, \rho_i \geq \rho_j \end{cases} \quad (7)$$

由定义可知,当点的局部密度不是最大的时候,说明该点并不是中心点,即将距离设定为与离它最近点的距离;当点的局部密度是最大的时候,说明该点是中心点,即将该点所对应的距离设为与离它最远点的距离。

根据上文提及聚类中心的特性,聚类中心的局部密度较大和“距离”较大这两个特征需要同时满足。综合考虑 ρ 值和 δ 值,重新定义新的变量 γ ,定义如下:

$$\gamma_i = \rho_i \delta_i, i \in I_s \quad (8)$$

这意味着聚类中心点具有较大的 γ 值。针对 γ 进行排序选取前 k 个大的 γ 值,并将其对应的超像素进行标号,记为 $\{n_j\}_{j=1}^k$, C_{n_j} 表示为第 j 个类的聚类中心(即图像的区域中心超像素)。

2.4 确定图像区域聚类中心

首先人工用画刷在前景或者目标区域上画一笔,分析交互的像素,找到对应 SLIC 算法中包含交互像素的超像素,并对记录标记的超像素编号,最后基于密度峰值算法找到这些超像素所属的区域中心。

在 2.3 节中,针对 γ 值已经确定了图像的所有区域中心点。对非区域中心数据点的划分,基于密度峰值聚类算法,遍历所有数据点的 ρ_i 值。如果数据点的局部密度 ρ_i 值比它近邻的数据点的 ρ_i 值小,它们同属于一个类,即在同一个区域。定义为:

$$c_i = \begin{cases} k, & C_k \text{ 为聚类中心} \\ -1, & \text{否则} \end{cases} \quad (9)$$

其中, k 为第 k 类聚类中心, -1 表示近邻数据点的 ρ_i 值大于所在的类簇。如果这个数据点是聚类中心点,那么比它的 ρ_i 值小的近邻数据点同属于这个类簇;如果近邻数据点的 ρ_i 值大于所在的类簇,则把它归类到其他具有更高 ρ_i 值的类簇。所有的非数据点完成分类后,分析标记的数据点所属的聚类中心。如果标记的数据点是聚类中心点,则记下对应数据点的编号;如果标记的数据点是非聚类中心,则记录这个数

据点所属的聚类中心数据点的编号。

2.5 前景子区域合并

基于局部密度找到人工交互的区域中心,同时也完成了非区域中心点(即非聚类中心点)的归类,即完成对所有超像素所属的区域中心的归类。分别将属于不同人工交互区域中心的超像素合并形成一个子区域,最后将这些子区域合并,形成较大的区域。子区域一般表达式为 subregion,用符号 SR 表示。若前景或者目标区域由 N 个子区域合并而成,则前景或者目标区域可以表示成:

$$\text{Region} = \bigcup_{i=1}^N \text{SR}_i (\text{SR}_i \cap \text{SR}_j = \emptyset, \forall i \neq j) \quad (10)$$

其中,Region 为人工交互的前景或者感兴趣的目标区域。

3 实验结果分析

3.1 算法过程及结果分析

实验选取了一张纹理图像来演示文中算法过程并且得到最终的分割结果。目标是提取人工交互部分的纹理图像区域,最终的分割结果如图 2(f)所示,其中白色区域为目标区域。结果显示算法稳定,分割准确。算法过程:将这张纹理图像超像素化,然后确定若干个图像区域中心。再基于用户的交互找到目标所属的区域中心,形成子区域得到分割结果。

图 2 超像素的大小设置为 $K = 100$;图像区域中心设置 $C = 6$,其他参数设置为 $\gamma = 0.01, \alpha = 0.05$ 。

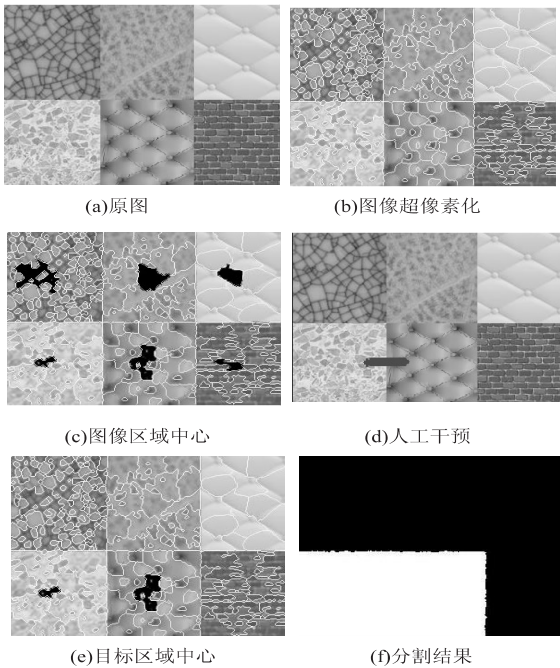


图 2 文中算法中间过程及结果

3.2 超像素参数的设置以及结果分析

在一定范围内增加超像素的个数,会使得分割效果更好。实验选取了一张飞鸟图,目标是从背景中提

取出人工交互的飞鸟,分割结果如图 3 所示。图像的超像素个数 K 分别设置为 200 和 400,得到实验分割结果。区域中心个数设置为 $C = 10$,其他参数设置为 $\gamma = 0.001, \alpha = 0.01$ 。

由实验结果可看出,超像素 $K = 200$ 时,鸟的尾巴与翅膀的连接处的轮廓没有很好地分割出来。增加超像素的个数到 $K = 400$ 时,鸟的尾巴与翅膀的连接处的轮廓很好地分割出来了。适当增加超像素的个数可以更细致地提取目标的轮廓,使得分割更准确。

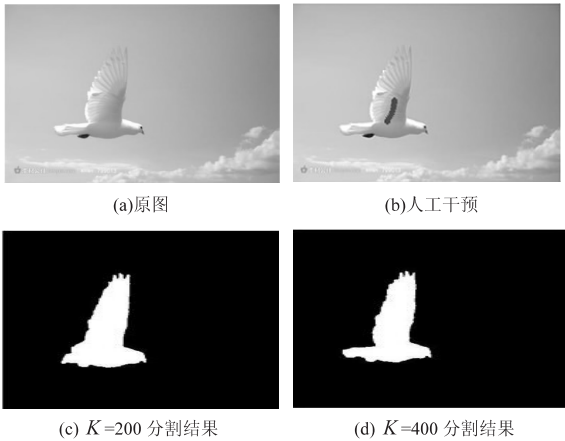


图 3 超像素参数设置分割结果

3.3 区域中心个数的设置以及结果分析

在一定范围内,图像的区域中心个数设定的与图像的复杂度成正相关,即越复杂的图像,区域中心的个数设置越多。实验选取了两张动植物图,一张红花图和海星图。目标是从背景中提取人工交互的红花和海星,分割结果如图 4 所示。两幅图的超像素个数都设置为 $K = 200$,红花图与海星图的区域中心个数分别设置为 $C = 3$ 和 $C = 12$ 。其他参数设置为 $\gamma = 0.05, \alpha = 0.05; \gamma = 0.5, \alpha = 0.01$ 。由仿真实验结果可看出,红花图的复杂度明显低于海星图的复杂度,区域中心的设定跟图像的复杂度相关,复杂度越高的图像,设定的区域中心的个数越多。

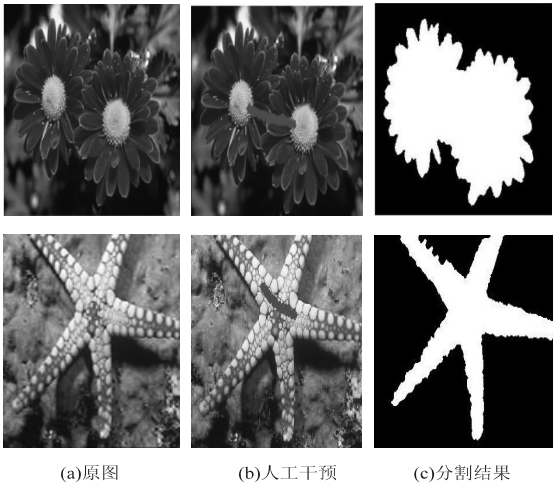


图 4 动植物分割结果

3.4 其他参数设置及结果分析

在一定程度上, γ, α 参数的设定与前景和背景的相似度成负相关, 背景与前景越相似, γ, α 的参数设置越大。实验选取了一些分辨率较高的人物图像, 目标是从背景中提取出人工交互的人物, 结果如图5所示。图5自上而下超像素个数 K 分别设置为 800, 800, 1 500, 1 500; 区域中心个数分别设置为 $C = 8, 12, 15, 50$; 其他参数设置为 $\gamma = 0.15, \alpha = 0.05$; $\gamma = 0.1, \alpha = 0.05$; $\gamma = 0.005, \alpha = 0.05$; $\gamma = 0.001, \alpha = 0.05$ 。



图5 人物分割结果

由这四幅人物图可看出, 前两个人物图与后两个人物的背景与人物的颜色差异不同。前两个人物图的背景与人物的颜色差异性大, 而后两个人物图的背景与人物颜色差异性较小。背景和人物颜色具有较大的相似性时, γ, α 的参数值设定越小; 而背景和人物颜色具有较小的相似性时, γ, α 的参数值设定越大。

当然, 文中算法还存在不足之处。如图4所示的海星图可看出, 海星左上的触手上细小的颗粒没有很好地分割出来; 如图5所示的第四幅图, 人物书包上狭小的书包带没有分割出。所以算法在分离更细小的物体边界处没有过多的优势。

4 结束语

提出了一种快速简单的交互式图像分割算法。针对图像的复杂度, 采用多个区域中心来刻画目标区域, 基于局部密度确定图像区域中心, 再利用用户交互信息分析前景的区域中心, 得到前景的提取。实验结果表明, 该算法有较好的分割结果, 并且用户操作简易。对于未来的工作, 将会在如下两个方面对算法进行改进。一方面是进一步加强对细小边界的处理, 使得分

割更准确; 另一方面是算法某些参数需要人为调控, 会使算法自适应选择参数。

参考文献:

- [1] 王梅, 李玉鑑, 全笑梅. 图像分割的图论方法综述[J]. 计算机应用与软件, 2014, 31(9): 1-12.
- [2] PONT-TUSET J, ARBELÁEZ P, BARRON J T, et al. Multiscale combinatorial grouping for image segmentation and object proposal generation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(1): 128-140.
- [3] 王宁, 黄琳琳. 一种新的交互式图像分割算法[J]. 海军工程大学学报, 2011, 23(2): 42-46.
- [4] HE J, KIM C, KUO C. Interactive segmentation techniques[M]. Berlin: Springer, 2014: 7-16.
- [5] FALCAO A X, UDUPA J K, SAMARASEKARA S, et al. User-steered image segmentation paradigms: live wire and live lane[J]. Graphical Models and Image Processing, 1998, 60(4): 233-260.
- [6] ROTHER C, KOLOGOROV V, BLAKE A. GrabCut: interactive foreground extraction using iterated graph cuts[C]//Proceedings of the ACM SIGGRAPH. Los Angeles, USA: ACM, 2004: 309-314.
- [7] LI Y, SUN J, TANG C K, et al. Lazy snapping[C]//Proceedings of the ACM SIGGRAPH. Los Angeles: ACM, 2004: 303-308.
- [8] BAI X, SAPIRO G. A geodesic framework for fast interactive image and video segmentation and matting[C]//IEEE 11th international conference on computer vision. Rio de Janeiro: IEEE, 2007: 1-8.
- [9] BLAKE A, ROTHER C, BROWN M, et al. Interactive image segmentation using an adaptive GMMRF model[C]//European conference on computer vision. Prague, Czech Republic: Springer, 2004: 428-441.
- [10] BOYKOV Y, KOLMOGOROV V. Computing geodesics and minimal surfaces via graph cuts[C]//Proceedings ninth IEEE international conference on computer vision. Nice, France: IEEE, 2003: 26-33.
- [11] REN X, MALIK J. Learning a classification model for segmentation[C]//Proceedings ninth IEEE international conference on computer vision. Nice, France: IEEE, 2003: 10-17.
- [12] 宋熙煜, 周利莉, 李中国, 等. 图像分割中的超像素方法研究综述[J]. 中国图象图形学报, 2015, 20(5): 599-608.
- [13] ACHANTA R, SHAJI A, SMITH K, et al. SLIC superpixels compared to state-of-the-art superpixel methods[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2012, 34(11): 2274-2282.
- [14] 由柏娇. 一种基于超像素的彩色图像分割算法[D]. 哈尔滨: 哈尔滨工程大学, 2015.
- [15] RODRIGUEZ A, LAIO A. Machine learning. Clustering by fast search and find of density peaks[J]. Science, 2014, 344(6191): 1492-1498.