

DOI: 10.7652/xjtuxb201501019

应用视觉显著性的快速有偏聚类超像素算法

李鹏^{1,2}, 杨旻², 方涛¹

(1. 上海交通大学自动化系, 200240, 上海; 2. 西安交通大学电子与信息工程学院, 710049, 西安)

摘要: 针对正则化超像素方法的超像素数随边缘拟合要求迅速增长的问题,提出了一种有偏聚类超像素算法。结合人类视觉对目标专注程度不一的特点,在 SLIC 算法框架下,提出了基于视觉显著性的非均匀初始化方法和有偏聚类距离函数。算法在图像的显著性区域进行密集的过分割,保持目标边缘的细节信息,而在非显著区域仅生成稀疏的超像素,以降低分割块数,再通过一步全局聚类 and 边缘逐步细化过程,有效地保证了图像的边缘拟合,同时提高了算法的速度。实验表明,在相同超像素数下,所提算法在边缘查全率、欠分割错误率以及运行速度方面均优于传统算法。

关键词: 超像素;视觉显著性;有偏聚类;边缘细化

中图分类号: TP391 **文献标志码:** A **文章编号:** 0253-987X(2015)01-0112-06

A Fast Superpixel Algorithm with Biased-Clustering Using Visual Saliency

LI Peng^{1,2}, YANG Yang², FANG Tao¹

(1. Department of Automation, Shanghai Jiaotong University, Shanghai 200240, China;

2. School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: A novel biased-clustering superpixel algorithm is proposed in the framework of SLIC to improve the problem that the conventional superpixel methods have bottleneck in controlling the tradeoff between superpixel number and boundary adherence. The algorithm employs the visual saliency into the non-uniform mesh initialization step and the biased-clustering distance function by noticing that human's visual attentions are distinctive to different salient objects, so that dense over-segmentations are generated in salient regions to keep sufficient information for object's boundary; while sparse segmentations are generated in non-salient regions to reduce the number of segmentation blocks. Moreover, the ideas of one-step global clustering and gradual boundary refining are applied to speed up the algorithm. Experimental results and comparisons with several state-of-the-art superpixel algorithms show that the proposed algorithm reflects the boundary adherence for salient objects well under the same number of segmentation blocks, and has a higher boundary recall and lower under-segmentation, as well as the least time-consuming.

Keywords: superpixel; visual saliency; biased-clustering; boundary refining

超像素是一种基于图像局部颜色、纹理等特征的图像过分割表示,近年来已成为计算机视觉算法,如目标识别^[1]、图像分割^[2]、特征检测等应用中的重要工具。作为预处理步骤,超像素不仅能极大地提高后续算法的计算效率,并且因为分割块之间具有

更高层的上下文信息,在视觉算法中有着重要的利用价值,其对目标边缘的拟合能力、分割块形态的正则化程度以及计算效率是衡量超像素算法好坏的标准。

目前一些图像分割算法,如 Mean shift 算法^[3]、

GS 算法^[4]、Quick shift 算法^[5]等具有较好的图像边缘拟合效果,但由于生成的是非正则化分割块,不利于后续对分割块间空间上下文信息的引入。正则化的超像素方法有 N-cuts 算法^[6]、Lattices 算法^[7]、Vcell 方法^[8]和 SLIC 算法^[9]等,但由于超像素数是事先指定的,当需要更大程度拟合目标的局部边缘细节时,正则化超像素方法需要增加全局的过分割数目,因此算法的复杂度也相应地提高。

考虑到高层视觉感知的特点,人类更希望辨识图像中显著性目标的细节,而由于背景所含信息量较少,稀疏的划分块是可以接受的分割方式。基于上述观察,本文提出了一种新的有偏聚类超像素算法,其特点是结合了高层视觉显著性特征,根据图像内容生成变密度的超像素结果,从而同时达到了在图像显著性区域的边缘细节拟合和局部超像素正则化的要求。实验结果表明,在相同的超像素数下,所提算法在边缘查全率、欠分割错误率以及运行速度方面均优于传统算法。

1 传统的 SLIC 超像素算法

超像素算法可以归结为将具有局部相似性的像素划分成色块的聚类问题,其算法包括初始化、距离函数、聚类搜索范围 3 个核心部分。SLIC 算法在初始化阶段,将图像划分为尺寸相等的网格,以网格中心作为初始超像素中心。对于任意的两个像素,定义像素间的距离 D_{ij} ,同时考虑了像素空间位置和颜色信息,即

$$D_{ij} = \left(d_c^2 + \left(\frac{d_s}{S} \right)^2 m^2 \right)^{1/2} \quad (1)$$

式中: d_c 代表两像素在 CIELAB 色彩模型中的颜色距离; d_s 则表示它们在图像上的空间距离; m 为距离权重,一般设定为常量; S 为初始网格的边长。

SLIC 指定了以超像素中心相邻的 2 倍网格边长的矩形搜索范围,迭代计算该范围内的像素与各超像素中心的距离,并将其指派给距离最近的超像素中心。相比于经典的 K-means 算法的全局搜索,SLIC 大大减少了距离计算的次数,但整个聚类过程仍需要十几次迭代才能收敛。

实验发现,SLIC 经过两次迭代搜索后已经生成了接近最终收敛结果的超像素,后续的迭代过程虽然计算了搜索范围内所有像素的距离,但事实上仅对分割边缘上的若干像素进行了调整。算法的绝大部分时间花费在了对超像素块边界的微调。

2 有偏聚类超像素算法

本文提出的有偏聚类超像素算法,针对传统 SLIC 算法的 3 个核心部分进行了改进。算法对显著区域生成密集超像素,保留了边缘细节信息,而在非显著区域生成稀疏的超像素,减少全局超像素数,并采用一步聚类和边缘细化搜索方案,大大提高了聚类过程的速度。算法总体思路如下:

(1)在初始化阶段,根据图像显著性图,生成显著区域不同密度尺寸的初始网格划分;

(2)设计结合图像显著性信息的距离函数,对显著性不同的区域采用不同的空间距离和颜色距离权重;

(3)首先进行一次聚类搜索,随后迭代进行边缘细化,即仅对当前超像素边缘上的像素进行距离计算和超像素划分,以减少距离计算次数,从而提高算法效率。

所提算法分为初始化网格、一步聚类和边缘细化 3 个步骤,如图 1 所示。

2.1 初始化网格

容易看出,当初始超像素中心间的距离越小,超像素就越密集,反之则越稀疏。因此,在显著区域选取更加密集的初始超像素中心,可使图像中显著性分割分得更加细致。目前已有很多较为成熟的显著性算法,本文采用了 SIG 算法^[10]提取图像的显著性,目的是在较低的运算时间下保证生成的显著图的准确性。

实验发现,当像素显著性大于 0.7 时,所得显著区域接近于人类视觉认知的感兴趣区域。本文采用双阈值方法对显著图进行二值化,即取显著度大于 0.7 的区域与显著度最高的占图像面积 15% 的区域的并集定为显著区域 R2,其余为不显著区域 R1。

改进的初始化算法通过两次不同尺寸的网格划分选取初始超像素中心。假设图像像素数为 N ,图像中显著像素数为 c ,参数 t 表示不显著区域与显著区域超像素大小的比值, S_1 为全局图像均匀划分网格 G1 的间距,则局部显著区域 R2 的网格 G2 划分的间距 S_2 满足

$$S_2 = \frac{S_1}{t^{1/2}} \quad (2)$$

定义超像素块总数为 K ,其值等于显著区域网格数目与不显著区域网格数目之和,即

$$\frac{N}{S_1^2} - \frac{c}{S_1^2} + \frac{c}{S_2^2} = K \quad (3)$$

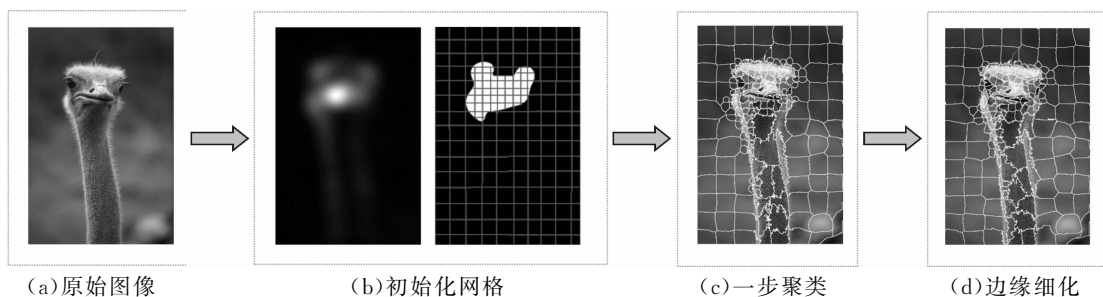


图1 本文算法的步骤

由式(2)、式(3)可得,全局图像的网格 G_1 划分的间距 S_1 为

$$S_1 = \left(\frac{N + c(t-1)}{K} \right)^{1/2} \quad (4)$$

选取划分后的网格中心为初始超像素中心。定义网格 G_1 中心点显著性为 0, 网格 G_2 中心点显著性为 1。

2.2 一步聚类

一步聚类的思想是在定义的距离函数下,对每个超像素中心对应搜索范围内的像素仅进行一次距离计算,将该像素划分到距离最近的超像素块中,其中距离函数的定义和搜索范围的指定都会影响到划分的结果。

图 2 给出了 SLIC 算法取不同权值 m 下的超像素结果。可以看出,随着 m 的减小,色彩在距离计算中所占权重增加,得到的超像素更加拟合图像边缘细节,因而对于图像中的显著区域,应选取更小的权值 m 以保持更多的图像细节。因此,在有偏聚类定义的距离计算公式中,权值 m 取值根据像素的显著性服从高斯分布。像素 P_i 与第 k 个超像素中心的距离定义为

$$D_{ik} = \left(d_c^2 + \left(\frac{d_s}{S} \right)^2 m^2 \right)^{1/2}, \quad \begin{cases} S = S_1, & \text{if } \text{sal}(C_k) = 0 \\ S = S_2, & \text{if } \text{sal}(C_k) = 1 \end{cases} \quad (5)$$

式中: $m = 40 \exp \left(-\frac{(\text{sal}(P_i))^2}{2\sigma^2} \right)$, $\text{sal}(P_i)$ 表示图像中像素 P_i 的显著性; $\text{sal}(C_k)$ 表示第 k 个超像素中

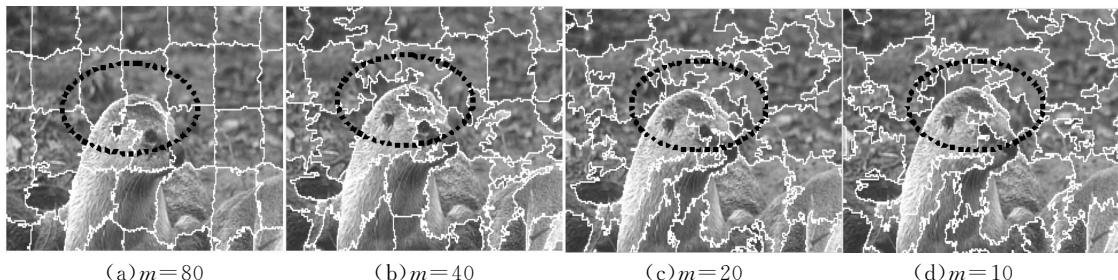
心 C_k 的显著性。根据实验发现,对于大多数图像,当 m 服从峰值为 40 的高斯分布、参数 σ 取 0.5 时效果较好,这样既能保持一定的细节特征,又能使非显著区域超像素形状规则。

有偏聚类超像素算法,每次搜索在以超像素中心 C_k 为中心的矩形邻域内,搜索范围的边长根据超像素中心的显著性自适应调整,若 $\text{sal}(C_k)$ 为 0,则取搜索范围为 $2 \times S_1$, 若为 1 则取 $2 \times S_2$ 。

2.3 边缘细化

从图 1c、图 1d 可以看出,经过一步聚类生成的超像素结果,如果按照聚类算法继续迭代,则只有超像素边缘上像素进行着局部调整。该现象已经被大量实验证实,其原因是在距离函数中考虑了像素的空间距离,因而绝大多数的内点与初始超像素中心较近,经过距离计算判定后不会发生变动。

受 VCell 算法^[8]中对超像素边界处理技巧的启发,本文在后续的迭代过程中进行边缘细化,即只对超像素的边缘像素进行距离计算与指派。边缘细化依旧遵循着 SLIC 算法框架,首先查找超像素中心邻域矩形内的边缘像素,计算边缘像素 P_i 与该超像素中心 C_k 的距离,根据该距离与原距离的大小决定是否重新指派该像素。从生成超像素的过程看,随着边缘像素的重新指派,超像素边缘不断地细化直到超像素的目标误差函数收敛。因此,在后续的迭代过程中,相比于要计算每个超像素邻域内的所有像素,算法的运算次数显著减少。

图2 SLIC 算法在不同权值 m 下的超像素结果

下面比较分析本文算法的时间复杂度。由于选择的 SIG 显著性算法的时间耗费很低,相比于后续的超像素分割过程,运行时间可以忽略。提出的算法在一步聚类过程中,运算的次数为 $4N(N$ 为图像像素数),在边缘细化过程中,每次迭代只计算边缘上像素,以均匀的方格边缘估计,计算次数约为 $4(KN)^{1/2}$ (K 为超像素数,取值远小于 N),因此算法整体时间复杂度为 $O((KN)^{1/2})$ 。传统的 SLIC 超像素算法的时间复杂度为 $O(N)$,GS 算法时间复杂度为 $O(N\log N)$,QS 算法最慢,时间复杂度为 $O(dN^2)$,其中 d 为一个常量,由此得到本文算法在时间复杂度上具有优势。

3 实验结果分析

3.1 实验设计

实验将本文的有偏聚类超像素算法与 SLIC 算法^[9]、QS 算法^[5]、GS 算法^[4]在整体分割效果、算法复杂度与效率、边缘拟合能力以及用于图像的多类目标分割 4 个方面进行了对比。测试图像来源于 Berkeley 数据库^[11],其中包含了 500 张大小为 481×321 的不同场景图像。多类分割训练测试图像来源于 MSRC 数据库^[12],其中包含 21 类不同场景的目标图像,其他 3 种算法参数均采用默认参数。实验计算机 CPU 为 i5-4430,内存为 4 GB。

3.2 超像素结果

超像素作为计算机视觉算法中的重要预处理步骤,其过分割结果是否能很好地拟合目标的边缘,将对后续算法起到关键作用。图 3 为本文算法与最新算法的效果对比。人脸是一般计算机视觉算法的兴趣区域,因此好的超像素算法应该尽可能地表现出人脸的细节信息,比如眉毛、眼睛等。从图 3 中可以看到,本文算法在显著区域人脸上分割的更加密

集细致,将眼睛、眉毛等部位分割为单独超像素,符合人脸视觉感知的特点,并且超像素形状也比较规则,正则化效果好。SLIC 算法、QS 算法在人脸和背景区域具有相同的分割密度,从细节放大图中可以看到,二者都将眼睛、眉毛和之间的皮肤混合为一个超像素,从而丢失了眼睛、眉毛信息。GS 生成非正则化的超像素,其形状和大小都不规则。

图 4 比较了在不同超像素数下,各算法对目标边界细节的表示能力。在相似的显著目标边界的拟合效果上,本文算法需要 100 个超像素,如图 4b 所示,而 SLIC 算法需要 500 个超像素,如图 4d 所示。因此可得,在保持相同的细节信息时,本文算法所需要的超像素数远小于 SLIC 算法所需要的,这将明显减少后续视觉算法需要处理的超像素数,降低时间复杂度。

3.3 运行时间比较

图 5a 为算法对 Berkeley 数据库^[11]中图像的运行时间比较,其中横轴表示了以最大尺寸图像的像素数为基础的归一化图像像素数,超像素数均为 300。可以看出,QS 算法时间耗费最大,曲线基本随图像像素数呈次方倍增长,SLIC 算法时间耗费呈线性增长,GS 算法时间耗费介于 QS 算法、SLIC 算法之间。与文中分析的时间复杂度结果一样,本文算法的运行时间最少,处理最大尺寸图像所需时间不到 5 s。图 5b 为在 Berkeley 数据库^[11]上选取超像素数为 100~500 时的算法时间对比,可以看出,本文算法的运行时间明显低于其他算法的运行时间。

3.4 边缘拟合效果

一种量化比较超像素边缘拟合能力的标准方法是计算边缘查全率^[8]和欠分割错误率^[8]。边缘查全率计算正确边缘落入以超像素边缘像素为中心、两个像素为半径区域内的部分的比例。另外,假定存

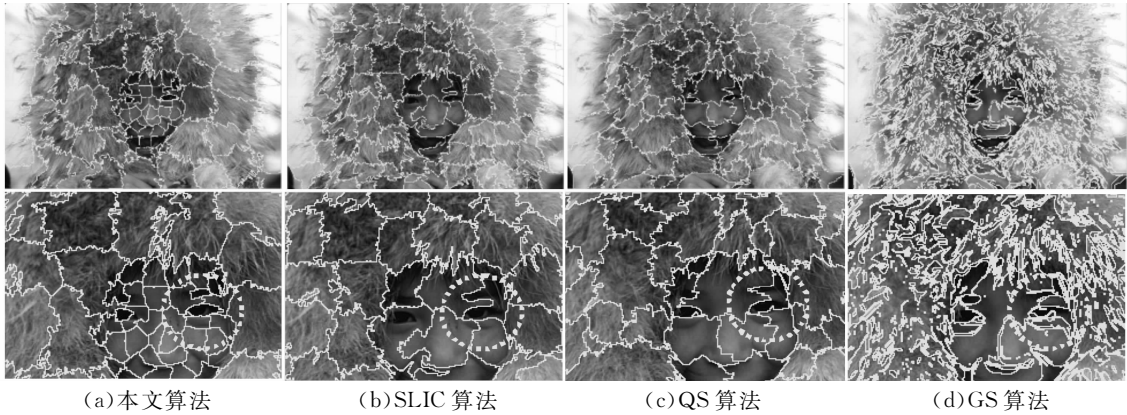


图 3 不同算法在 Berkeley 数据库^[11]中图像上的超像素结果对比(下图是上面结果的局部放大图)

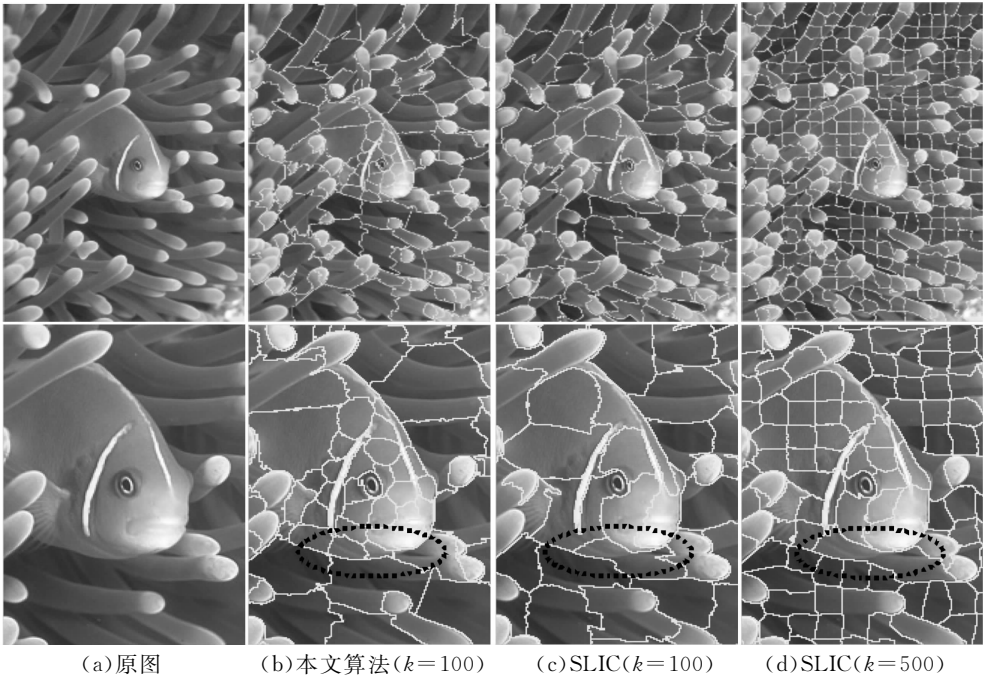


图 4 本文算法与 SLIC 算法在 Berkeley 数据库^[11]中图像上的边界拟合效果对比(下图是上面结果的局部放大图)

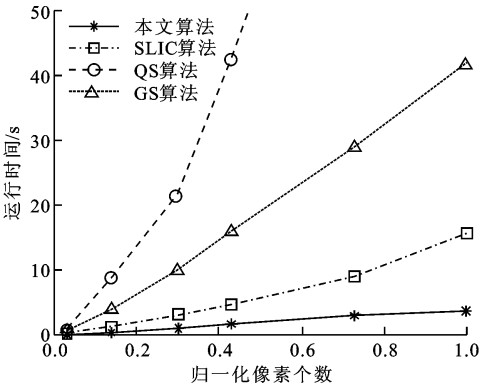
在一个正确的标准分割区域和与该分割区域有交集的超像素,欠分割误差就是计算该超像素超出标准分割区域的比例。整幅图像的欠分割错误率是所有分割区域欠分割误差的平均值。

图 6 分别为本文算法与 SLIC 算法、QS 算法、GS 算法的边界查全率和欠分割错误率的对比。测试集为 Berkeley 数据库^[11]中 200 张测试图像和库中提供的手工目标分割结果。可以看出,在相同的超像素数下,本文算法的边缘查全率最高,欠分割误差率最低。这是因为本文的超像素集中在显著物体上,能保留显著区域更多的边缘信息,且对背景进行稀疏的分割,忽略了背景中无用的边缘信息。例如,本文算法能够更好地拟合人脸中鼻子、眼睛的边缘,而背景中的水面、天空、草原往往不存在边缘或者边缘无用而被忽略。因而,本文算法能够保持图像的

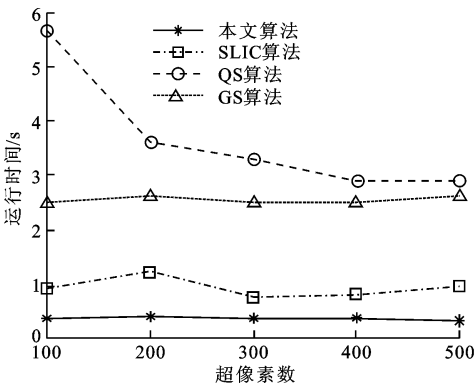
重要边缘信息,有利于提高后续处理的准确性。

3.5 多类目标分割

超像素通常作为目标分割算法的预处理步骤。本文在 MSRC 数据库^[12]上将本文算法与 SLIC 算法、QS 算法、GS 算法用于多类目标分割的效果进行了对比,所有算法的超像素数均取为 200。分割算法采用文献[13]所提算法,首先利用超像素计算图像的颜色、纹理、几何以及位置特征,在获得的特征基础上训练 21 类目标分类器,最后学习 CRF 模型生成图像分割结果。通过与手工分割结果的对比计算算法的分割精度,结果见表 1。可以看出,用于多类目标分割时,本文算法的分割精度均优于其他算法。图 7 为算法与 SLIC 算法用于多类分割的分割效果,可以看出,所提算法能够保留更多的目标细节,分割结果更接近于人工分割结果。



(a)不同像素数图像的算法时间



(b)不同超像素数的算法时间

图 5 在 Berkeley 数据库^[11]上不同算法的运行时间对比

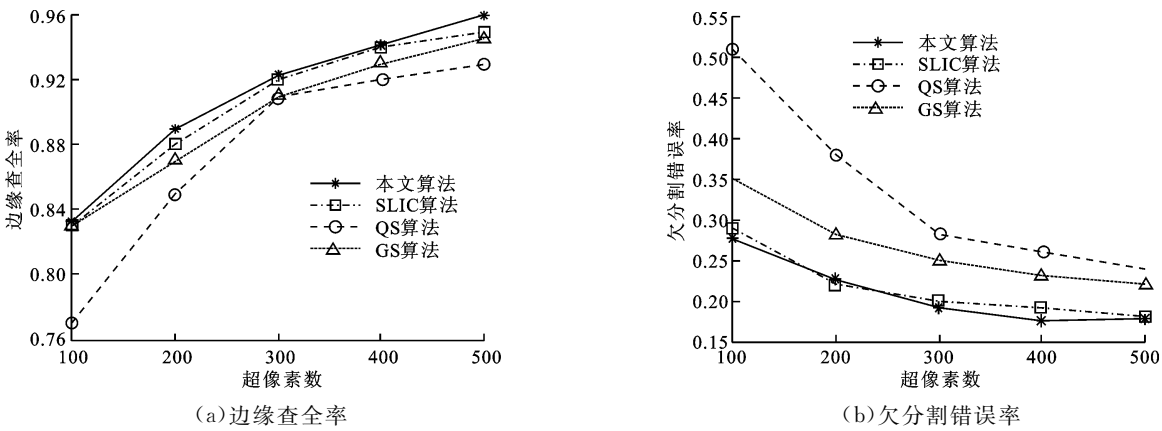


图 6 在 Berkeley 数据库^[11]上不同算法边缘查全率和欠分割错误率对比

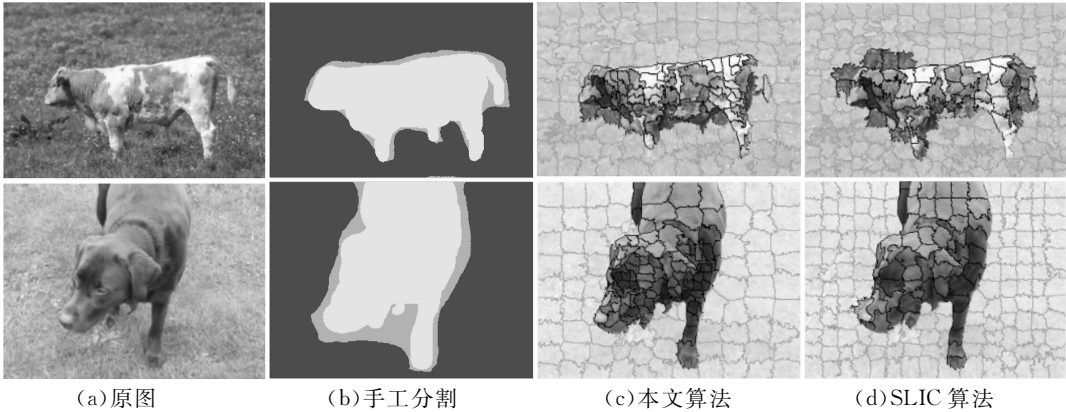


图 7 在 MSRC 数据库^[12]中图像上的多类目标分割结果

表 1 不同超像素算法用于多类目标分割的精度对比

超像素算法	分割精度/%
本文算法	74.9
SLIC 算法 ^[9]	74.3
QS 算法 ^[6]	73.7
GS 算法 ^[4]	73.0

4 结 论

本文提出了一种新的有偏聚类超像素算法,同时考虑了高层的视觉显著性特征和底层的图像颜色特征,引入视觉显著性非均匀初始化方法和有偏聚类距离函数,通过一步全局聚类 and 边缘细化方法形成一种更加符合人类视觉感知特点的变密度局部正则化的超像素结果。实验从分割效果、算法复杂度、边缘拟合能力以及用于多类图像分割等方面均证实了提出算法的有效性和快速性。

参考文献:

[1] MORI G. Guiding model search using segmentation [C] // 10th IEEE International Conference on Computer

Vision. Piscataway, NJ, USA; IEEE, 2005: 1417-1423.

[2] FULKERSON B, VEDALDI A, SOATTO S. Class segmentation and object localization with superpixel neighborhoods [C]//12th IEEE International Conference on Computer Vision. Piscataway, NJ, USA; IEEE, 2009: 670-677.

[3] 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.

[4] FELZENSZWALB P F, HUTTENLOCHER D P. Efficient graph-based image segmentation [J]. International Journal of Computer Vision, 2004, 59(2): 167-181.

[5] VEDALDI A, SOATTO S. Quick shift and kernel methods for mode seeking [C]// 10th European Conference on Computer Vision. Berlin, Germany; Springer, 2008: 705-718.

[6] SHI J, MALIK J. Normalized cuts and image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8): 888-905.

(下转第 138 页)