

# 基于颜色均值显著点聚类的作物病害叶片图像分割方法

张善文, 张晴晴, 齐国红 (郑州大学西亚斯国际学院, 河南郑州 451150)

**摘要** 作物病害叶片图像分割是病害类型识别方法的一个重要步骤, 其分割效果直接影响后续的认识结果。病害叶片图像的复杂多样性使得很多现有的图像分割方法不能有效应用于作物病害叶片图像分割中。针对复杂的自然病害叶片图像分割难题, 提出一种基于颜色均值显著点聚类的作物病害叶片图像分割方法。该方法建立在 HIS 颜色空间, 首先构造基于像素点 HIS 模型的带权无向图, 然后计算病害叶片图像像素点的邻域的颜色均值, 再计算该点前后两个邻域的颜色均值差作为该点的颜色跳跃度, 当跳跃度大于设置的一个阈值时, 该像素点为病斑点。结果表明, 该算法具有较高的分割精确度和较好的抗噪声性能。

**关键词** 病害叶片图像分割; 显著点; 颜色均值显著点聚类; 颜色跳跃度

中图分类号 S126; TP391 文献标识码 A

文章编号 0517-6611(2019)10-0228-03

doi: 10.3969/j.issn.0517-6611.2019.10.067

开放科学(资源服务)标识码(OSID):



## Segmentation Method of Crop Disease Leaf Image Based on Color Mean Significant Point Clustering

ZHANG Shan-wen, ZHANG Qing-qing, QI Guo-hong (SIAS International University, Zhengzhou University, Zhnegzhou, Henan 451150)

**Abstract** The segmentation of crop disease leaf image is an important step in disease type recognition method, and its segmentation effect directly affects the subsequent recognition results. Because of the complexity and diversity of disease leaf images, many existing image segmentation methods cannot be effectively applied to crop disease leaf image segmentation. Aiming at the difficult problem of leaf image segmentation in complex natural environment, a method of diseased leaf image segmentation based on color mean salient point clustering is proposed. This method is based on HIS color space. Firstly, weighted undirected graph based on pixel HSI model was constructed. Then, the color mean of the neighborhood of the pixel of the diseased leaf image was calculated. The difference of the color mean of the two neighborhoods before and after the point was calculated as the color leaping degree of the point. When the leaping degree was larger than a threshold, the pixel was a spot pixel. Experimental results showed that the proposed algorithm had higher segmentation accuracy and better noise immunity.

**Key words** Disease leaf image segmentation; Salient points; Color mean significant point clustering; Color leaping degree

作物病害叶片图像分割是作物病害图像分析处理和病害类型识别的基础工作和重要环节, 也是计算机图像处理和机器视觉领域的研究热点之一。已经有很多复杂图像分割方法和技术, 包括基于图论<sup>[1-2]</sup>、分水岭方法<sup>[3-5]</sup>、模糊聚类<sup>[6-7]</sup>、灰度阈值的<sup>[8]</sup>、边缘检测<sup>[9]</sup>和混合算法的分割方法<sup>[10]</sup>等。很多研究成果已经成功应用于病害检查和识别中。作物病害叶片图像的复杂多样性使得作物病害图像的自动分割目前尚有很多困难和问题<sup>[11-13]</sup>。在彩色病害叶片图像检测中, 由于彩色叶片图像的病斑复杂性和特殊性, 对应的灰度图像边缘检测算子一般不能直接应用于彩色图像。目前彩色病害叶片图像边缘检测问题还没有较成熟的解决方法, 因此研究彩色病害叶片图像边缘检测方法还处于初始阶段。作物病害叶片图像作为特殊目标, 即要求能够更全面、更准确地分割出来的病害区域。在分割中, 当病斑区域与周边区域的对比度不明显时, 它将导致误判, 其结果表现为非病斑区域被保留下来。由于叶脉部分与其他部位的对比度较低, 提取的病害区域会包含一些不相关的部分, 从而影响病害分割的准确性。不准确的分割将会影响到后续病害识别的精确性。因此, 分割后的病害区域需要校正。病害区域的颜色是非常复杂的, 难以用 RGB 色彩空间中的颜色来描述。

## 1 均值显著点聚类方法

假设观测到一组样本  $Y_1, Y_2, \dots, Y_n$ , 通常观测值按时间出

现先后排列。对于第  $m$  个点  $Y_m$  的期望有  $EY_1 = \dots = EY_{m-1} = a_1$ ,  $EY_m = \dots = EY_n = a_2$ , 且  $1 < m \leq n$ , 则当  $EY_1 = EY_n = \dots = EY_n$  时, 无显著点。当  $a_1 \neq a_2$  时,  $Y_m$  称为显著点,  $|a_1 - a_2|$  为  $Y_m$  的跳跃度。

在像素均值显著点聚类模型中, 假设像素点为  $Y_1, Y_2, \dots, Y_n$  为该点的灰度值, 且服从独立正态分布, 则显著点的个数至多为 1 个的均值显著点模型可表述为:

$$Y_i = \begin{cases} \mu_0 + \varepsilon_i, & 1 \leq i \leq k^* \\ \mu_0 + \Delta + \varepsilon_i, & k^* + 1 \leq i \leq n \end{cases} \quad (1)$$

式中,  $k^*$  为显著点,  $\mu_0$  为前  $k^*$  个像素的均值,  $\Delta$  为像素集  $Y_1, Y_2, \dots, Y_n$  在  $k^*$  的均值变化幅度, 称为跃度, 假定随机误差  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$  为期望值 0 方差为  $\sigma^2$  的独立正态分布。

在均值显著点模型中定义 3 种均值:

$$\hat{\mu}^{(k)} = \frac{1}{k} \sum_{i=1}^k Y_i \quad (2)$$

$$\hat{\mu}_{(k)} = \frac{1}{n - k + 1} \sum_{i=k}^n Y_i \quad (3)$$

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n Y_i \quad (4)$$

式中,  $\hat{\mu}^{(k)}$  为前  $k$  个像素的均值, 简称前邻域均值,  $\hat{\mu}_{(k)}$  为后  $n - k$  个像素的均值, 简称后邻域均值,  $\hat{\mu}$  为所有  $n$  个像素的均值。

计算  $k$  处的均值跳跃度  $\delta(k)$ ,

$$\delta(k) = |\hat{\mu}^{(k)} - \hat{\mu}_{(k)}| \quad (5)$$

若有  $k^*$  满足  $\delta(k^*) = \max |\delta(k)|$ , 则  $k^*$  为该区域内的第 1 个显著点。

病害叶片图像是由很多像素点组成的 3 个通道的 3 个矩阵。在图像病斑检测中, 可以认为在一定区域内颜色变化

**基金项目** 国家自然科学基金项目(6147237); 河南省教育厅科技攻关项目(182102210544, 19B520029, 182102311094)。

**作者简介** 张善文(1965—), 男, 陕西西安人, 教授, 博士, 博士生导师, 从事复杂图像处理及其应用研究。

**收稿日期** 2018-11-11

的点就是显著点,即为图像的病斑像素点。观测值为图像像素点的 RGB 颜色灰度值,从水平方向、垂直方向、主对角线方向、反对角线方向取观测值并计算跳跃度,如图 1 所示。

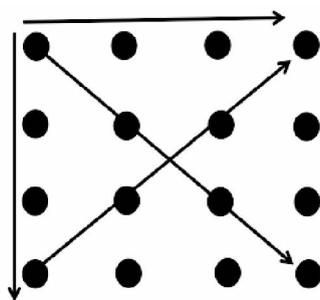


图 1 扫描方向示意图

Fig.1 Sketch map of scanning direction

假设图像  $X$  为  $n * m$  个像素点,其中  $X_{i,j}$  为第  $i$  行第  $j$  个像素点。由式(5)计算  $X_{i,j}$  点在 HSI 3 个分量上的 3 个跳跃度,分别记为  $\delta_H(X_{ij})$ 、 $\delta_S(X_{ij})$  和  $\delta_I(X_{ij})$ 。水平方向的检测算法具体步骤如下:

①初始变量  $a$  为第  $i$  行的第 1 个像素点的位置偏移量, $b$  为该行最后 1 个像素点的位置偏移量。由式(6)计算  $X_{ij}$  的跳跃度  $\delta(x_{ij})$  的 3 个分量:

$$\delta(x_{ij}) = \left[ \left( 1 - \frac{j}{m} \right) \frac{j}{m} \right]^{1/2+r} \cdot [\delta_H(x_{ij}) + \delta_S(x_{ij}) + \delta_I(x_{ij})] \quad (6)$$

②计算第  $i$  行中所有像素点的跳跃度  $\delta(x_{ij})$ ,  $j=1,2,\dots,m$  由式(7)检测出  $[a,b]$  邻域内的第 1 个显著点  $x_{ij}^*$ ,

$$\delta(x_{ij}^*) = \max[\delta(x_{ij})] \quad (7)$$

③计算得到显著点  $x_{ij}^*$  在  $[a,b]$  邻域中相对于  $a$  点的位置偏移  $k$  个像素,则当  $\delta(x_{ij}^*) > h_1$  时,将检测邻域分成  $[a,k-1]$  和  $[k,b]$  2 个邻域进行检测。重复执行以上步骤①

和②,得到第  $i$  行中的所有病斑像素点。

④依次计算图像中的所有行,即得到水平方向的所有显著点。

⑤去除非病斑像素点。检测到显著点集合为  $\Phi = \{X_{ij}^* \}$ ,若有  $X_{ij} \in \Phi$ ,记  $X_{ij} = 1$ 。对于检测出的显著点,若其相邻的像素点均不是显著点时,判断该点为非病斑像素点,则即设置为 0。

⑥利用膨胀腐蚀运算对得到的病斑像素点集处理,得到病斑图像。

## 2 性能分析

为了说明该算法的有效性,使用图像分割的类别平均准确率 (Meanacc) 和平均区域重合度 (MeanIU) 为分割精度指标<sup>[14]</sup>。在分割时应考虑算法的分割时间。

$$\text{Meanacc} = \frac{1}{n_{cl}} \sum_p \frac{n_{pp}}{t_p} \quad (8)$$

$$\text{MeanIU} = \frac{1}{n_{cl}} \sum_p \frac{n_{pp}}{t_p + \sum_q n_{pq} - n_{pp}} \quad (9)$$

式中, $n_{cl}$  为分割图像像素所属类别, $p$  表示图像中像素正确对应的像素类别, $q$  表示图像中像素错误分类后所属像素类别, $t_p$  为分割中  $p$  类别的像素总数  $t_p = \sum_j n_{pj}$ , $n_{pp}$  为分割中正确分割为类别  $p$  的像素总数, $n_{pq}$  为分割中属于类别  $p$  但被误分为  $q$  的像素数量。

## 3 试验与分析

为了测试该方法的有效性,将其应用于黄瓜病害叶片图像病斑分割中,并与基于 SVM 分割方法<sup>[11]</sup>、基于 K-means 聚类法<sup>[12]</sup> 和基于改进最大类间方差 (Otsu) 的方法<sup>[13]</sup> 进行了对比。

图像处理环境选用 Intel(R) Pentium(R) i5 处理器、内存 4GB 的 PC 机,软件使用 matlab 7.0 进行试验。

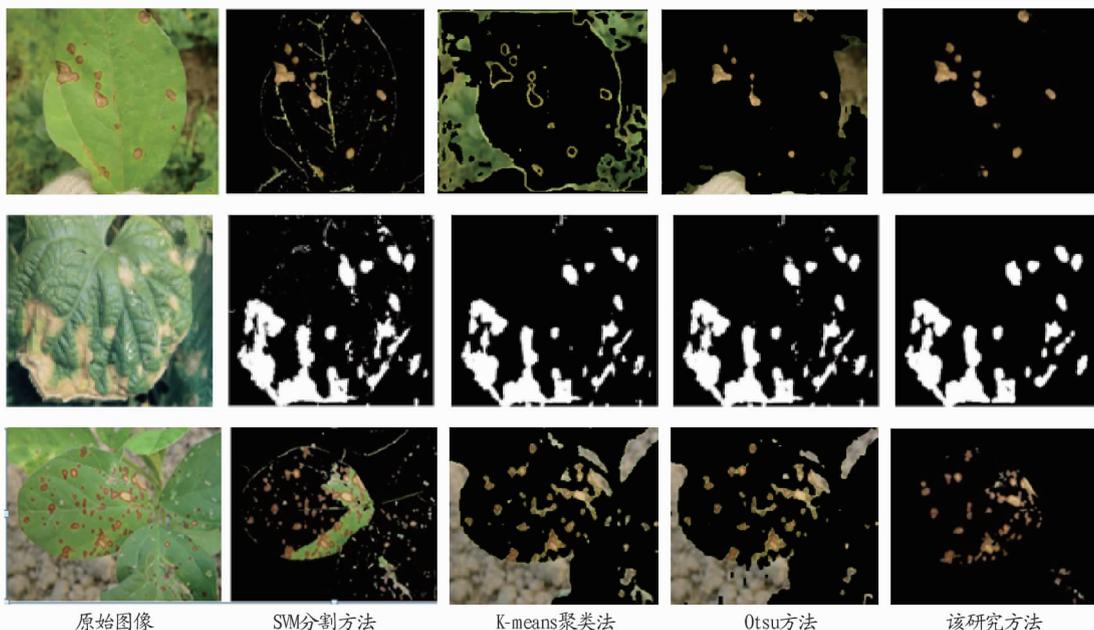


图 2 5 种不同分割方法的分割病斑图像比较

Fig.2 Comparison of the segmentation disease spots by five segmentation methods

将上述方法在 50 幅病害叶片图像上的类别平均准确率和平均区域重合度作为评价指标。图 2 为利用该方法和 4 种现有的分割方法对 3 种病害叶片图像的分割结果。表 1 为利用 4 种方法对 50 幅图像进行分割的平均准确率和平均区域重合度以及分割时间。

从图 2 和表 1 可以看出,该方法优于其他 3 种方法,而且能够减少复杂背景信息对病斑的干扰。其原因是,该方法充分利用了病害叶片图像中的病斑像素的显著性特性,以彩色图像的 HSI 颜色值为特征值,在水平、垂直、对角线方向的区域上建立均值变点模型,最后通过叠加得到检测结果。

表 1 不同方法分割精度比较

Table 1 Comparison of the segmentation accuracy of different methods

分割方法 Segmentation methods	类别平均 准确率 Class average accuracy %	平均区域 重合度 Average area overlap %	单个图像分 割时间 Single image segmentation time//s
SVM	96.103	95.150	0.45
K-means	95.142	94.243	0.36
Otsu	95.172	93.296	0.29
该研究方法 Method of this research	98.019	97.204	0.24

#### 4 结论

该研究将统计中的显著点思想应用于病害叶片图像分割中,提出了一种基于均值显著性聚类的作物病害叶片图像分割方法。该方法利用 HIS 颜色空间的 3 个分量,定义了跳跃度,由此进行图像分割,得到病斑图像。实验结果表明,提出的分割算法分割速度快、分割准确率高,并且能够减少背景干扰。该方法为后续进一步提取特征参数、识别和诊断病

害类别打下良好的基础。

#### 参考文献

- [1] ANGULO J. Geometric algebra colour image representations and derived total orderings for morphological operators—part I: Colour quaternions [J]. Journal of vision and communication image representation, 2010, 21(1): 33–48.
- [2] 朱晓临, 陈漫, 李雪艳, 等. 结合形态学变形虫和复合颜色空间的彩色图像边缘检测模型[J]. 计算机辅助设计与图形学学报, 2014, 26(7): 1060–1066.
- [3] 余旺盛, 侯志强, 宋建军. 基于标记分水岭和区域合并的彩色图像分割[J]. 电子学报, 2011, 39(5): 1007–1012.
- [4] 李占波, 杨二伟, 李进文. 基于改进分水岭和区域合并的彩色图像分割[J]. 计算机工程与设计, 2014, 35(7): 2458–2461.
- [5] 王栋, 于威威, 庄斐弘. 基于标记分水岭和 FLICM 模糊聚类的图像分割方法研究[J]. 微型机与应用, 2017, 36(17): 49–51, 58.
- [6] THEOHARATOS C, ECONOMOU G, FOTOPOULOS S. Color edge detection using the minimal spanning tree[J]. Pattern recognition, 2005, 38(4): 603–606.
- [7] 李丹丹, 史秀璋. 基于 HSI 空间和 K-means 方法的彩色图像分割算法[J]. 微电子学与计算机, 2010, 27(7): 121–124.
- [8] KONISHI S, YUILLE A L, COUGHLAN J, et al. Statistical edge detection: Learning and evaluating edge cues[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 25(1): 57–74.
- [9] EVANS A N, LIU X U. A morphological gradient approach to color edge detection [J]. IEEE Transactions on Image Processing, 2006, 15(6): 1454–1463.
- [10] 周强强, 王志成, 赵卫东, 等. 基于水平集和视觉显著性的植物病害叶片图像分割[J]. 同济大学学报(自然科学版), 2015, 43(9): 1406–1413.
- [11] 白文斌, 白帆, 贺文文, 等. 基于 SVM 的高粱叶片病斑图像自动分割提取方法研究[J]. 农学报, 2014, 4(6): 101–106.
- [12] 邹秋霞, 杨林楠, 彭琳, 等. 基于 Lab 空间和 K-Means 聚类的叶片分割算法研究[J]. 农机化研究, 2015(9): 222–226.
- [13] RADZALI M H, KAMAL N A M, DIAH N M. Measuring leaf area using Otsu Segmentation Method (LAMOS) [J]. Indian journal of science & technology, 2016, 9(48): 1–6.
- [14] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Madison, WI: IEEE, 2015: 3431–3440.

(上接第 186 页)



图 5 蛋壳破碎实物

Fig.5 Eggshell broken solid

#### 5 结论

通过 SolidWorks 模拟仿真破碎辊对鸡蛋冲击力的的大小, 设定不同转速, 发现鸡蛋所受冲击力是不同的, 随着转速的提高, 冲击力增大, 最后通过简化试验模型验证破碎辊转速达到 150 r/min 时, 能够实现鸡蛋壳的破碎。摩擦去皮试验

时, 需要对鸡蛋施加 10 N 左右压力, 才能取得较好的剥皮效果。因此, 通过破碎辊与螺旋机组合机械结构, 可以完成鸡蛋壳破碎和鸡蛋输送, 利用橡胶辊摩擦实现蛋壳去除。该研究可以为企业开发类似产品提供技术参考。

#### 参考文献

- [1] 井美娇, 邓志超, 李在强, 等. 普通鸡蛋、柴鸡蛋和乌鸡蛋中铜、铬、锰含量比较研究[J]. 黑龙江畜牧兽医, 2019(4): 55–57.
- [2] 韩芝洋, 侯进慧, 倪若昕. 富硒鸡蛋的开发与分析[J]. 农产品加工, 2019(3): 6–9.
- [3] 孙恒, 陈作模, 葛文杰. 机械原理[M]. 北京: 高等教育出版社, 2008: 275–276.
- [4] 赵丹, 张伟华. 基于 SolidWorks 千斤顶虚拟装配与运动仿真[J]. 机械研究与应用, 2018, 31(4): 7–9.
- [5] 黄剑锋, 余涛, 陈江义. 刚体碰撞的 Hertz 接触力模型比较分析[J]. 机械设计与制造, 2017(8): 28–30.
- [6] 罗文峰, 刘崇琪, 梁猛, 等. 基于光谱法鸡蛋壳等离子体特征参数诊断[J]. 西安邮电大学学报, 2013, 18(5): 70–72.
- [7] 陈超祥, 胡其登. Solidworks Motion 运动仿真教程[M]. 北京: 机械工业出版社, 2014: 66–67.
- [8] 胡一龙, 杨硕, 蒋娜, 等. 基于 Solidworks 的发动机曲柄连杆机构的运动仿真[J]. 河北建筑工程学院学报, 2018, 36(1): 119–123.
- [9] 胡茂, 苟兴能. 鸡蛋壳的性状[J]. 绵阳经济技术高等专科学校学报, 2002, 19(4): 10–14.
- [10] 小花牛. 10 秒钟剥出整鸡蛋[J]. 医食参考, 2017(6): 38.