

文章编号: 2095-2163(2019)02-0063-05

中图分类号: TP391.41

文献标志码: A

一种基于模糊 C 均值聚类的多阈值苹果病害图像分割方法

贾庆节, 齐国红, 忽晓伟

(郑州大学 西亚斯国际学院, 河南 新郑 451150)

摘要: 针对苹果病害叶片的图像分割问题,以斑点落叶病、花叶病和褐斑病为研究对象,提出一种基于模糊 C 均值聚类的多阈值分割算法。首先,将苹果病害叶片图像进行滤波;然后,利用模糊 C 均值进行病斑图像的聚类,再滤除病斑图像中的非病斑像素,根据分类结果获得分割阈值;最后,利用多阈值算法对苹果病害叶片图像进行分割,得到病斑图像。通过与其他分割方法进行比较表明,本方法分割出来的苹果病斑,分割准确率达到 94% 以上,分割效果明显。

关键词: 模糊 C 均值聚类; 苹果病害叶片; 多阈值

A multi-threshold segmentation method based on fuzzy C-means clustering for the apple disease leaf images

JIA Qingjie, QI Guohong, HU Xiaowei

(SIAS International University, Zhengzhou University, Xinzheng Zhengzhou 451150, China)

[Abstract] According to the properties of apple disease, using apple leaf's spot disease, mosaic disease and leaf spot as the research object, the paper proposes a multi-threshold segmentation algorithm based on Fuzzy C-means(FCM) clustering. Firstly, preprocess the image by filter to reduce the effect of noise; secondly, use FCM clustering method to divide the disease spot, and then eliminate some fake disease spot, achieve the thresholds according to the segmentation of the results; finally, use multi-threshold segmentation algorithm to segment the apple disease spot. Three kinds of apple disease leaf images are segmented effectively using the algorithm. The results of experiment indicate that apple disease spots can be separated precisely from the apple leaf images, the correct extraction rate can reach 94%. The research shows that the method is more valid than other method shown in this paper.

[Key words] Fuzzy C-means clustering; apple disease leaf; multi-threshold

0 引言

苹果一直是人们广为喜爱的食物之一,全国种植面积达 8 000 余万亩。随着种植面积的日益扩大,苹果生长过程中发生大规模病害的可能性也在增加,如果不能及时发现病害并进行准确诊断和治疗,可能造成严重的经济损失,这就需要对苹果进行实时监测,以便及时发现病害。计算机图像处理技术以及监控技术的发展,使得苹果病害智能检查和诊断成为可能。研究发现^[1-3],作物的病害一般都会在叶片上表现出来,不同的病害种类引起的叶部症状也不同^[4]。利用现代图像采集技术和图像处理方法对作物叶片病害进行诊断和识别是目前的一个主流研究方向^[5-7]。利用计算机图像处理技术处理病害叶片在节约人工成本的同时,极大地提高了经济效益。

阈值分割算法简单、快速、有效,是图像分割的基本技术之一^[8-10],如基于 Otsu 的图像分割算法^[11]、基于局部熵阈值和 Otsu 的分割算法^[12]、基于自适应全局阈值融合的分割方法^[13]、基于自适应局部阈值的分割方法^[14]、基于遗传算法的最大熵阈值的图像分割方法^[15]等。单阈值分割算法将图像分为 2 类之后,对于这 2 个类内的目标将不再进行区分,而如果在这 2 个类内继续使用阈值分割,能够将目标更为精确地分割出来,由此,将单阈值算法推广到多阈值。多阈值分割算法有 Otsu 多阈值快速分割算法^[16]、最大熵多阈值算法^[17-18]、基于谱聚类的多阈值图像分割算法^[19]等。多阈值分割图像时,计算量大,且有冗余的不需要分割的背景信息,如果能直接在类似目标信息的图像上进行多阈值分割,可以提高运算速度和目标的准确性。经过对比分析,发现文献[20]采用的模糊 C 均值(Fuzzy C-means,

基金项目: 河南省科技厅项目(182102210546); 河南省高等学校重点科研项目基础研究计划(16A510034, 17A520017); 河南省教育厅 2016 年度民办高校品牌专业(自动化)建设项目(教政法[2016]896 号); 河南省教育厅第九批河南省重点学科(检测技术与自动化装置)建设项目(教高[2018]119 号)。

作者简介: 贾庆节(1966-),女,副教授,主要研究方向:图像处理、模式识别。

收稿日期: 2018-12-10

FCM)聚类算法具有较强的适应性,可以对图像进行模糊聚类分割,以减少冗余背景信息,非常适合将其应用于多阈值分割。为此,本文提出一种基于FCM的多阈值图像分割算法,用于对苹果病害叶片的分割。本方法首先将采集到的苹果病害叶部的图像进行中值滤波,去除噪声的同时较好地保持病害部位的边缘;然后,利用FCM对苹果病斑图像进行聚类,再通过病斑部位和正常部位的红色信息所占比例的不同,将病斑图像中非病斑像素滤除,根据聚类和非病斑像素滤除的结果,获得分割阈值;最后,利用多阈值算法分割出苹果病斑图像。

1 算法研究

FCM方法根据图像中的像素与 c 个聚类中心间的加权相似度测度,对目标函数进行迭代最小化计算,以确定像素点属于某个类别的最佳度^[21]。设 $X = \{x_1, x_2, \dots, x_n\}$ 为样本集, n 为 X 元素总数目, c 为聚类中心数,聚类就是要将 x_1, x_2, \dots, x_n 区分为 X 中的 c 个子集,要求性质相近或相同的样本最大程度地聚类在同一子集内。那么关于 X 的一个模糊 C 划分是一个 $c \times n$ 矩阵 $U = [u_{ij}]$ ($0 \leq u_{ij} \leq 1$), u_{ij} 是样本 x_j 对第 i 类的隶属度,则该矩阵称为模糊聚类矩阵,该矩阵具有以下性质:

$$\sum_{i=1}^c u_{ij} = 1, 0 \leq u_{ij} \leq 1, \sum_{j=1}^n u_{ij} > 0, j = 1, 2, \dots, n; i = 1, 2, \dots, c, \quad (1)$$

则FCM聚类的目标函数如式(2)所示:

$$J_m(U, V) = \sum_{j=1}^n \sum_{i=1}^c u_{ij}^m d_{ij}^2, \quad (2)$$

其中:

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}}{d_{kj}}\right)^{\frac{2}{m-1}}}, \quad (3)$$

$$d_{ij}^2 = \|x_j - v_i\|^2, \quad (4)$$

$$v_i = \frac{1}{\sum_{j=1}^n u_{ij}^m} \sum_{j=1}^n u_{ij}^m x_j, i = 1, \dots, c. \quad (5)$$

其中, V 是 c 个聚类中心组成的集合, $V = [v_i]$; d_{ij} 为第 j 个样本到第 i 类的距离,用来度量数据点与聚类中心有多么相似; $m \in (1, +\infty)$ 是加权指数,求出式(2)中目标函数 $J_m(U, V)$ 的最小值,就是要求的最佳的分类效果。

与其它颜色空间模型相比,RGB彩色模型能够更加准确地表现出病斑部位和正常部位的差异。苹

果叶片的病害部位和正常部位的颜色差异较大,正常部位的颜色表现为类似红色,病害部位的颜色表现为类似黄色,RGB颜色空间中这2部分最大的差异就是 R 分量所占的比例不同,病害部位 R 分量所占的比例是最大的。利用如下算法将病害部位图像中的正常像素滤除。将此方法可解析为如下公式:

$$I_{gm}(x, y) = \begin{cases} 0, & I(x, y).G = \max\{I(x, y).R, I(x, y).G, I(x, y).B\}, \\ 1, & \text{其它}. \end{cases} \quad (6)$$

其中, $I(x, y)$ 为苹果叶部图像; $I_{gm}(x, y)$ 为病斑部位图像; $I(x, y).R$ 、 $I(x, y).G$ 和 $I(x, y).B$ 表示原图 $I(x, y)$ 中像素点 (x, y) 处的RGB的分量值。

通过公式(6)可以将FCM聚类后的病斑部位中非病斑像素点滤除,此后,根据处理后的图像求得分割阈值,利用多阈值算法完成苹果病斑图像的分割。多阈值算法是在单阈值的基础上发展起来的。多阈值分割算法是先将图像划分为多个区域,分别在各个区域求得最佳阈值。如图1所示,如果只考虑局部区域A-B、或B-C、或C-D等等,可以很好地用Otsu法找到(A, B)或(B, C)之间的分割阈值,即局部最佳分割阈值。找到图1中各个局部的最佳阈值,就可以分割开各个独立的波峰,从而实现图像的多阈值分割。

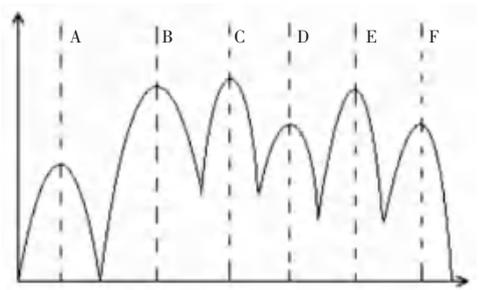


图1 多峰值图像

Fig. 1 Multi-peak image

假设图像的直方图分割成 m 类,各类均分割成 C_{i0}, C_{i1} ,文中利用Otsu法求得局部最佳阈值,分割2组的类间方差最大时,即求得最佳分割阈值。图像中存在 m 个待分割的类,所有类的类间方差定义为:

$$\sigma_i^2 = \omega_0 (\mu_0 - \mu_r)^2 + \omega_1 (\mu_1 - \mu_r)^2 + \dots + \omega_{m-1} (\mu_{m-1} - \mu_r)^2 (i = 1, \dots, m). \quad (7)$$

其中, σ_i^2 为所有类的类间方差; ω_i 和 μ_i 分别是第 i 类的比例和均值; μ_r 是所有类的总均值;最优阈值 $(T_1^*, \dots, T_m^*) = \arg \max \sigma_i^2$ 。

2 算法验证

在图像采集过程中,不可避免地会受到噪声的干扰,为了后期的图像分割,就需要对图像进行平滑滤波处理以消除图像的干扰信息,保留并增强有用信息,提高后续图像处理的速度和可靠性。常用的去噪方法有平滑滤波、直方图均衡化、均值滤波和中值滤波法等,由于中值滤波在去除噪声的同时可以很好地保留图像的边缘信息,文中选择中值滤波对

苹果病害叶部进行滤波处理。

采用本文算法在 Matlab 编程环境下对 3 种苹果病害叶片(斑点落叶病、花叶病和褐斑病)进行分割,如图 2(a)~(d)所示。通过图像的直方图得到最佳阈值组,其中本文图像的最佳阈值组分别为(50,190,230)、(50,80,112,230)和(78,202)。为了验证本文算法的有效性,与多阈值分割出来的病斑图像进行比较,如图 2(e)所示。



(a) 3 种苹果病害叶片
(a) Three apple disease leaves



(b) 3 种苹果病害叶片中值滤波图像
(b) Median filtered image of three apple disease leaves



(c) 基于 FCM 分割结果
(c) Segmentation results based on FCM



(d) 基于本文算法病斑图像分割结果
(d) Image segmentation result based on the algorithm of this paper



(e) 基于多阈值分割病斑图像分割结果
(e) Results of lesion image segmentation based on multi threshold

图 2 3 种苹果病害叶片病斑图像分割结果比较

Fig. 2 Comparison of image segmentation results of three apple disease leaf algorithm

多阈值算法根据直方图选取最佳阈值,不考虑非病斑像素的干扰。图 2(e)是基于多阈值的 3 种苹果病斑图像的分割结果,可以看出,分割的病斑图像区域不连续,且由于病斑图像和正常部位的颜色像素相近,叶片正常部位和叶脉部位误作为苹果病斑图像分割出来。而本文算法首先使用 FCM 进行病斑图像的模糊聚类,然后再进行非病斑图像的像素滤除,最后利用多阈值算法进行病斑图像分割。多阈值算法根据图像的峰值信息得到不同区域的最佳阈值,将其应用于病斑图像分割中,从而得到分割较为完整的病斑图像。如图 2(d)所示,本文算法分割出来的苹果病斑轮廓清晰,形状较为完整。为了验证本文算法在分割病斑方面的优越性,将以上 3 种苹果病斑图像手动分割的结果作为正确的分割来计算分割正确率,采用如下公式计算分割的正确率:

$$Accuracy = \frac{TB}{TF + TB} \quad (8)$$

其中, TF (True Foreground) 是病斑判别正确的像素点数量; TB (True Background) 是病斑被判错的像素点数量。统计文中 3 种苹果病斑图像的分类正确率见表 1。

表 1 多阈值算法和本文算法的 3 种苹果病斑图像的分割正确率

Tab. 1 Multi-threshold algorithm and segmentation accuracy of three apple lesion images in this paper %

叶片序号	多阈值算法	本文算法
1	87.26	94.41
2	86.74	93.07
3	90.53	95.62

从表 1 可以看出,分别用多阈值算法和本文算法对文中 3 种苹果病害图像进行分割,单独使用多阈值算法对文中 3 种苹果病斑图像分割的分割正确率为 88.17%,本文算法分割正确率为 94.37%。

结合图 2 和表 1 的结果可知,相较于多阈值算法,本文算法表现出更好的分割效果。

3 结束语

为了提高苹果病斑图像分割效果,提出一种基于 FCM 的多阈值苹果病害图像分割方法。通过对 3 种苹果病害图像进行中值滤波,以去除图像中噪声的干扰,然后利用 FCM 实现病斑图像的模糊分割,利用病斑部位和正常部位 R 值不同,滤除非病斑像素。最后利用多阈值算法将 3 种病斑图像分割出来,由此完成苹果病斑图像的分割。实验结果表明,与多阈值分割算法相比,本文算法比多阈值算法

平均正确率高 6%,本文算法能够很好地将苹果病斑图像分割出来。

虽然本文算法能够很好地将苹果病斑图像分割出来,但是对于复杂背景下的苹果病斑图像分割,本文算法仍表现出一定的性能局限,因而亟待未来工作的研究改进及优化。下一步拟对复杂背景下的病斑图像进行分割。

参考文献

- [1] 刁智华. 大田小麦叶部病害智能诊断系统研究与应用[D]. 合肥:中国科学技术大学,2010.
- [2] GAVHALE K R, GAWANDE U. An overview of the research on plant leaves disease detection using image processing techniques [J]. IOSR Journal of Computer Engineering, 2014, 16(1): 10-16.
- [3] KULKARNI A H, ASHWIN P R K. Applying image processing technique to detect plant diseases [J]. International Journal of Modern Engineering Research, 2012, 2(5): 3661-3664.
- [4] 姜猛,王军,王志军,等. 苹果树病虫害决策支持系统研究与集成[J]. 山东农业科学, 2015, 47(6): 129-133.
- [5] 李宗儒,何东健. 基于手机拍摄图像分析的苹果病害识别技术研究[J]. 计算机工程与设计, 2010, 31(13): 3051-3053, 3095.
- [6] 樊景超. 苹果果实病害近红外光谱信息获取与识别模型研究[D]. 北京:中国农业科学院,2011.
- [7] 赵芸. 基于高光谱和图像处理技术的油菜病虫害早期监测方法和机理研究[D]. 杭州:浙江大学,2013.
- [8] CAMARGO A, SMITH J S. An image-processing based algorithm to automatically identify plant disease visual symptoms [J]. Biosystems Engineering, 2009, 102(1): 9-21.
- [9] CHEN Wen, CAO Li, QIAN Jingjie, et al. A 2-phase 2-D thresholding algorithm [J]. Digital Signal Processing, 2010, 20(6): 1637-1644.
- [10] HUANG D, WANG C. Optimal multi-level thresholding using a two-stage Otsu optimization approach [J]. Pattern Recognition letters, 2009, 30: 275-284.
- [11] SASAKI Y, OKAMOTO T, IMOU K, et al. Automatic diagnosis of plant disease-recognition between healthy and diseased leaf [J]. Journal of the Japanese Society of Agricultural Machinery, 1999, 61(2): 119-126.
- [12] KURNIAWATI N N, ABDULLAH S N H S, ABDULLAH S. Investigation on image processing techniques for diagnosing paddy diseases [C]//SOCPAR '09: International Conference of Soft Computing and Pattern Recognition. Malacca, Malaysia: IEEE Press, 2009: 272-277.
- [13] 李丽,柴文婷,梅树立. 基于自适应全局阈值融合标记的遥感图像建筑群分割[J]. 农业机械学报, 2013, 44(7): 222-228.
- [14] 夏心怡,苏真伟,李国辉. 基于形状特征的棉花异性纤维图像分割方法[J]. 农机化研究, 2010(7): 13-16.
- [15] 宋家慧. 基于遗传算法的最大熵阈值的图像分割[J]. 电子工程师, 2005, 31(2): 60-63.
- [16] 王祥科,郑志强. Otsu 多阈值快速分割算法及其在彩色图像中的应用[J]. 计算机应用, 2006, 26(S1): 14-15.
- [17] 张俊娜,冯云芝. 基于量子最大熵多阈值算法的图像分割研究[J]. 激光与红外, 2013, 43(5): 578-582.

(下转第 72 页)