

水稻害虫图像识别技术研究

李文斌 (杭州电子科技大学通信工程学院,浙江杭州 310018)

摘要 [目的]解决水稻害虫传统识别方法的低时效性问题。[方法]采用数字图像处理方法对水稻害虫进行图像识别和分类,对水稻害虫的虫体面积、虫体周长、偏心率、形状参数、似圆度、叶状性、球形性等几何形状特征进行提取和研究,并采用支持向量机(SVM)分类器对水稻害虫二化螟、三化螟、稻飞虱、卷叶螟进行分类。[结果]利用所建立的6个特征判别函数对4种水稻害虫进行判别分类,识别率达到96.67%,说明这6个经过筛选的特征具有很强的判别性。[结论]支持向量机分类器的识别方法很好地解决水稻害虫传统识别方法的低时效性问题。支持向量机以风险最小化为原则,兼顾训练误差与测试误差的最小化,具体体现在分类模型的选择和模型参数的选择上。

关键词 图像处理;特征提取;识别分类
中图分类号 S126 **文献标识码** A **文章编号** 0517-6611(2014)23-08043-03

Research for Rice Pests Image Recognition Technology
LI Wen-bing (College of Communication Engineering, Hangzhou University of Electronic Science and Technology, Hangzhou, Zhejiang 310018)

Abstract [Objective] The aim was to solve the low timeliness issue of the traditional identification methods. [Method] Digital image processing method was used to identify and class the rice pests images, and to extract and research their geometrical configuration features, such as pests area, perimeter, eccentricity, shape parameters, roundness and the other geometric shape features, then support vector machine (SVM) classifier was used to class rice pests, such as Chilo suppressalis Walker, Tryporyza incertulas Walker, Nilaparvata lugens Stal. Le-rodea eufala Edwards. [Result] The recognition rate of 4 kinds of pests was high and reached 96.67% by using established 6 kinds of characteristic critical functions, showing this screened 6 kinds of characteristics had strong discrimination. [Conclusion] The recognition method based on SVM better solves the low timeliness issue of the traditional identification methods. SVM is based on the principle of risk minimization, considers both training error and test error, it is embodied by both selection of classification model and selection of model parameters.

Key words Image process; Feature extraction; Identification and classification

水稻是我国最重要的粮食作物之一,因此提高水稻产量是当前水稻生产的重要目标。但是,由于生产设备和预防监控措施落后,尤其是农业生物灾害又频频发生,对水稻生产造成了非常严重的影响。近年来,水稻害虫危害逐年加重。因此,害虫准确、及时地预测预报成为了水稻害虫预防和治理的前提,而水稻害虫预测预报的核心内容就是图像识别和分类。

然而,以往的水稻害虫识别方法都是依靠人为完成,通过大量的工作人员实地考查水稻害虫的形状、颜色等外部特征进行,需要较长的周期,并且实时性差,严重影响了水稻害虫识别速度。基于此,笔者对水稻害虫图像自动识别技术进行研究。

1 系统流程

图像自动识别技术包括两个方面:水稻害虫图像几何特征提取和图像识别分类。图像几何特征反映了水稻害虫的一些基本特征,合理的几何特征便于更好地获得有效的害虫形态特征信息,进而用于图像识别分类。该研究通过对多种水稻害虫的面积、周长、偏心率、形状参数、似圆度、叶状性、球形性等几何形状特征进行了提取和研究,并采用支持向量机分类器(SVM)对多种水稻害虫进行了分类识别处理,解决了以往识别技术中时效性差的问题。系统流程图如图1所示。

2 图像预处理

2.1 图像去噪及其灰度化

基于 ARM 系统的 COMS 摄像头

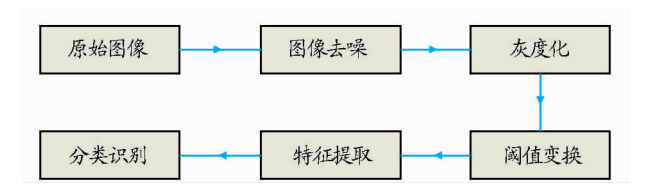


图1 系统流程

头获得的图像由于受到自身像素及其各种自然环境和拍摄角度的影响,往往图片质量会较差,从而给后面的识别工作带来困难,因此,需要对水稻害虫初始图像进行必要的预处理,以提高图像质量。为了提高图像质量,减少初始图像中的噪声,采取了图像平滑处理来去除噪声。常用的图像平滑方法有均值滤波和中值滤波,该系统采用中值滤波的方法,这是因为在同等尺寸大小下,中值滤波具有更好的去噪能力和较低的模糊度^[1]。

另一方面,通过 COMS 摄像头获取的图像都是彩色图像,然而彩色图像包含较多的像素,因此会占用较大的计算时间,而图像灰度化就是将彩色图像转换为灰度图像的过程,这样可以减少计算时间和减少内存占用,所以在进行图像特征提取之前首先将图像进行灰度化处理,处理结果如图2所示。

2.2 图像灰度阈值变换处理 在图像识别中,对直接拍摄得到的图像直接进行分类是不科学的,这是因为没经过处理的图像数据占用很大的存储空间,如果直接进行识别,则具有大量的计算量,降低效率;另外,直接拍摄得到的图像有很多冗余的信息,如图像的背景等一切与虫体无关的信息量,因此,在提取水稻害虫特征前必须先去除背景等无用信息,



原始图像



灰度图像

图2 原始图像及处理后的灰度图像

突出虫体有效信息。因此,采用灰度阈值变换处理,使得背景图案和虫体能够很好地分离开来。

所谓灰度阈值变换就是将一幅灰度图像转换成黑白的二值图像,设置一个灰度值,该灰度值起到分界线的作用。图像中某像素的灰度值小于该灰度值,则将该像素的灰度值设置为 0,否则设置为 255,这个起到分界线作用的灰度值成为阈值。



灰度图像



二值化图像

图3 灰度图像及处理后的二值化后图像

3 图像特征提取

对于人类视觉而言,人们通常利用水稻害虫的物理结构特征来进行识别处理,但这些特征对于机器识别是具有很大难度的^[2]。所谓特征提取就是从水稻害虫中提取出可以把虫体种类区分开来的并能被机器所直接使用的的数据量。要对水稻害虫进行特征提取,首先要确定水稻害虫有哪些虫体特性参数。其次,要适当地选择提取的特征参数,特征参数选择的好坏直接影响到后面能否正确地识别出害虫种类,因为有些虫体原始特征对于分类器的识别分类影响不大。因此,需要从试验的所有特征参数中挑选出效果最好的特征参数。该系统中,主要对水稻害虫的面积、周长、偏心率、形状参数、似圆度、叶状性、球形性和孔洞数等形态学特征进行识别分类^[3]。

3.1 虫体面积 虫体面积是一个能直接区分虫体种类的特征参数,直接关系到虫体的形状和大小,是用来描述水稻害虫大小的基本特性,图像虫体面积表示该图像中虫体在一定范围内所占的像素点总数。虫体面积表达式如下:

$$A = \sum_{x=1}^M \sum_{y=1}^M f(x, y)$$

3.2 虫体周长 周长和面积一样,是描述虫体外围形状和轮廓的重要参数。水稻害虫种类不同,它们的图像所占区域

在试验中,得出将灰度阈值设置为 180 的时候的二值化图像是最优的,基本上完整地描绘出了整个昆虫图像的轮廓。当设置为大于 180 的时候,图像的轮廓会变得厚重圆滑,从而使得外形失真;当灰度阈值小于 180 的时候,图像轮廓会有不同程度缺失,会使得大量数据丢失。试验结果如图 3 所示。

的大小也相差很大;周长 $P = A - \text{SUM}(in)$,其中, A 表示图像区域面积, $\text{SUM}(in)$ 表示图像的 4 邻域范围内的像素总数。

3.3 偏心率 偏心率的表达式为 $E = p/q$,是一个用来描述虫体紧凑性的一个参数,它在一定程度上表示水稻害虫的虫体形状,体型狭窄度。一般情况下,偏心率越大,则虫体就越狭窄;偏心率越小,则虫体外形就越宽大。

3.4 形状参数 形状参数的表达式为 $C = P^3/4\pi A$,该参数描述了虫体所占图像区域的紧凑性,形状参数这一特征可以有效地识别那些虫体受损的水稻害虫。一般情况下,当虫体受损后,很难将其区分出来,而加入形状参数后,这一缺陷得到了很好的改善。

3.5 似圆度 似圆度 $R = 4A/\pi l^2$,其中, A 是图像面积, l 是图像横轴长度。似圆度描述了图像形状的问题,是一个相对参数,可以用来描述水稻害虫的形态特征。

3.6 叶状性 叶状性反映水稻害虫的边界幅度变化特性,定义为: $B = R_l/W$,式中, B 为叶状性参数, R_l 为区域重心到边界的最短距离, W 为图像的横轴长度。

3.7 球形性 球形性 $SP = R_i/R_e$,其中, R_i 和 R_e 分别表示目标内切圆和外切圆的半径,两个圆的圆心都在区域的重心上。

3.8 孔洞数 将水稻害虫图像区域中无用的小区域定义为孔洞,对小区域进行标记并计数实现孔洞数的计算。

4 图像特征提取结果分析

选取 50 幅二化螟图像、50 幅三化螟图像、50 幅稻飞虱图像、50 幅卷叶螟图像,测得上述 8 个特征数据如表 1 所示。从表 1 可以看出,二化螟图像面积在 3 种水稻害虫中最大,周长也最大;而卷叶螟图像偏心率最大,是卷叶螟体型狭长、细窄的体现;稻飞虱、二化螟和三化螟展翅宽度较大,反映为偏心率较小。叶状性反映了目标边界的曲直变化频率和幅度。在 4 种昆虫中,卷叶螟图像的叶状性最大,是由于卷叶螟向外伸展使得图像边界幅度和方向变化很大,二化螟和稻

飞虱的翅和体之间也有比较大的变化,其叶状性小于卷叶螟而大于三化螟,根据似圆度,也可以判断出卷叶螟似圆度最大。另一方面,根据球形性和孔洞数很难判断出一个昆虫的形状特征和种类。

因此,运用逐个分析方法对上述特征进行筛选,得到 6 个特征,分别为区域面积、偏心率、形状参数、周长、似圆度、叶状性,剔除了孔洞数特征以及球形性。利用所建立的 6 个特征判别函数对以上 4 种昆虫进行了判别分类,结果如表 1 所示,识别率达到 96.67%,说明这 6 个经过筛选的特征具有很强的判别性。

表 1 图像特征提取结果

水稻害虫	虫体面积	虫体周长	形状参数	似圆度	叶状性	偏心率	球形性	孔洞数
二化螟	10 393.54	768.82	4.69	0.37	0.008 46	0.44	0.15	26.44
三化螟	5 196.72	521.32	4.35	0.35	0.001 80	0.31	0.11	26.12
卷叶螟	2 846.29	415.64	7.65	1.04	0.146 00	1.64	0.13	22.04
稻飞虱	4 351.63	498.32	6.59	0.98	0.009 40	0.42	0.17	25.02

5 支持向量机(SVM)分类器的设计

一般情况下,传统意义上的识别技术都是通过分类器对训练样本的拟合情况进行虫体识别,利用最小化训练集上的分类器作为训练目标,通过提供充足的样本来提高分类器的识别率,然而,这个方法存在一个很严重的缺陷。当样本数量严重不足的时候,不能保证一个很好地分类了训练样本的分类器也能够很好地测试样本,在缺乏代表性的小训练集情况下,一味地降低训练集上的分类错误就会导致过度拟合。

针对传统识别模式中存在的缺点,采取支持向量机(SVM)分类器的识别方法,该方法很好地解决了这一问题。支持向量机以风险最小化为原则,即兼顾训练误差与测试误差的最小化,具体体现在分类模型的选择和模型参数的选择上^[4]。

该试验中,将二化螟、三化螟、稻飞虱、卷叶螟 4 类水稻害虫分别标记为 A、B、C、D 4 类样本,将这 4 类样本两类两类地组成训练集,得到(A,B)、(A,C)、(A,D)、(B,C)、(B,D)、(C,D)。对于这 4 类样本中的任意一类,就可以用相对应的分类器来识别,像 A 类样本,可以组合成(A,B)、(A,C)、(A,D)这 3 类分类器来对害虫进行分类识别,因此可以根据这些分类器的置信度来对这些二分类器进行可靠性排列。一般情况下,置信度高的分类器得到的结果可靠性就高,置信度低的分类器出现误判断的机会就比较大。虫体识别判别过程如下所示。

第 1 步:对不同分类器的置信度大小进行排列,依次为(A,C)、(A,B)、(A,D)、(B,D)、(C,D)、(B,C),并分别编号为第 1 类分类器、第 2 类分类器、第 3 类分类器、第 4 类分类器、第 5 类分类器、第 6 类分类器。

第 2 步:设被识别对象为未知昆虫 X,首先由第 1 类分类器进行识别操作,根据第 1 类分类器的判别函数的结果来判断。当第 1 类分类器的判别函数的结果为正数时,则结果为类型 A,所有关于类型 C 的判别函数均被淘汰;若结果为负

数时,则结果为 C,所有关于类型 A 的判别函数都被淘汰;若判别函数的结果为 0,则表示拒绝判断,这时选用第 2 类分类器进行识别;如果结果类型为 C,则所剩判别函数为(B,D)、(C,D)、(B,C)。

第 3 步:被识别对象 X 再由第 4 类分类器进行识别,若判别函数结果为正数,淘汰所有关于 D 类的判别函数,则所剩下的判别函数为第 6 类分类器(B,C)。

第 4 步:被识别的样本在由第 6 类判别函数进行识别,若得到结果为正数,则判定最终的分类器结果为 B。

6 分类识别结果

由表 2 可知,4 种水稻害虫的识别率分别为 96%、94%、94%、90%。

表 2 4 种水稻害虫识别率

水稻害虫	识别数目	识别率//%
二化螟	48	96
三化螟	47	94
稻飞虱	47	94
卷叶螟	45	90

注:参试水稻害虫的数目均为 50。

7 结论

该研究初步选择 8 个水稻害虫图像特征,并根据试验过程和结果挑选出 6 个具有实际判别意义的几何特征,这些特征是相互独立的,符合建立分类器要求特征之间具有相互独立性的原则。该研究提取 8 种几何形状特征非常直观且易于提取,能够反映水稻害虫特点,也能反映出昆虫的体型结构以及形态特征。但如何提取更多能够直接反映昆虫形态特征和鉴别特征的特征及其设计更有效的分类器是今后努力的方向,也是完善昆虫数学形态学理论的重要步骤之一。

(下转第 8082 页)

群体大致分为两类:一是退休后的老年人,他们选择 CSA 农场,除了注重养生健康之外,很大原因是为了让自己枯燥的晚年生活变得积极有意义。二是家中有儿童的家庭,这些家庭选择与 CSA 农场建立信任关系,很大程度上是为了发挥农业的教育功能,增加与孩子的互动。

4 社区支持农业项目融资对策建议

4.1 社区支持农业项目自身解决办法

4.1.1 推广微信营销。社区支持农业项目利用微信营销可以首先进行分类的客户管理,对客户进行归类,针对新老客户进行不同的信息推送,定时发送他们需要的信息。还可以和客户互动,及时收到客户对配送菜品及服务的评价。其次可以精准营销定位,根据微信中附近的人这一功能在自己营业可控范围内进行产品推广。

4.1.2 加大媒体宣传。在产品及“社区支持农业”概念的宣传上,寻求通过媒体跟踪报道的形式,记录农场成长的真实历程。采用制作宣传片的方式,加强与消费者的情感互动,提升消费者对健康食品的重视程度。

4.1.3 培养忠实消费者。知识分子和高薪阶层自上而下的推广方式适合中国文化五千年孕育的土壤。食用有机食品潜移默化地成为一种身份的象征,人们在生活观念和方式上都比较容易接受。另外,CSA 农场可以通过不定期的举办消费者座谈会,鼓励更多消费者参与到 CSA 农场日常生产活动中,形成“种在当季,食在当季”的消费者理念。

4.1.4 创建移动互联网农业品牌。社区支持农业项目经营者利用移动互联网进行推广健康农业品牌,利用网络信息技术搭建的农产品产销对接平台,能够打破时空限制,减少中间环节,降低成本,增加利润,为销售者提供了一个销售平台,也为消费者提供了一个选择无异于农产品“淘宝网”。同时,针对现下智能手机的广泛应用,同时开发了软件手机客户端,便于买卖双方实时操作,随时随地进行交易。

4.1.5 O2O 电子商务进入。运用电商的 O2O 功能,建立 CSA 农场和消费者的平台,将本地农场信息汇集起来。在消费端,提供了多种消费者参与农产品生产和销售方式,增强了农产品生产的透明度和信任度。在供应端,为农场提供生产和销售的 IT 专业服务,提高了农业生产效率和销售管理水平。

4.2 外部辅助解决方法

4.2.1 政府。政府应对社区支持农业项目融资创造良好的硬件和软件环境。首先,借助政府投资基础设施所引发的连带效应,加大财政资金支持,激发其他投资主体投资农业项目的热情;同时,政府应继续加大“两减免、三补贴”的扶持力度,贯彻惠农政策,激发农民从事农业生产的积极性,对农民

自主投资的项目给予奖励和补助。在投融资法律依据上,政府应尽快制定《农业投资法》,明确各级政府的农业投融资职责,并提供农业投融资保障机制,为农业投融资提供全程服务,通过优惠政策和奖励措施,鼓励农业投融资主体多元化。

4.2.2 其他金融机构。不同的金融机构给不同规模的企业提供金融服务的成本和效率是不尽相同的,大力发展和完善中小金融机构是解决我国中小企业,特别是社区支持农业项目融资难问题的新出路。这些中小金融机构除了在为社区支持农业项目提供服务方面有信息优势,同时还在农村劳动力转移、农业结构的调整和农村的社会保障中扮演着重要的角色。因此,从发展的角度看,社区支持农业项目的发展需要相应的金融体系来配合。我国应尽快建立以中小金融机构为主的金融体系,帮助中小金融机构特别是农村小资规模的信贷机构发展。

4.2.3 民间资本。民间资本具有数量庞大、启用灵活等特点,在投资现代农业中最具活力和生命力,是发展现代农业的基本力量。启动民间资本投资现代农业,已成为破解农业投入不足难题、增加农业投入的最现实、最有效的途径。投资环境是民间资本投资农业的吸引力、推动力和竞争力。要积极清理和修订不利于民间资本投资农业的相关规定,保护民间投资的合法权益,营造公平公正的法规政策环境。完善民间资本投资休闲农业、特色农业的相关制度,尤其是对涉及旅游开发、统合体建设的休闲农业项目进行严格审查,防止部分企业借此占田围地造房。简化审批程序,加大对支持民间资本投资农业的相关政策措施以及民间资本投资农业重要意义的宣传力度,营造民间资本启动、发育和成长的浓厚氛围,不断发展和完善“草根金融”。

参考文献

- [1] 陈卫平,黄娇,刘濛洋.社区支持型农业的发展现况与前景展望[J].农业展望,2011(1):54-58.
- [2] HENDERSON E R V E. Sharing the harvest: a citizen's guide to Community Supported Agriculture [M]. Chelsea Green Publishing, 2007.
- [3] 李良涛,王文惠,王忠义,等.日本和美国社区支持型农业的发展及其启示[J].中国农学通报,2012(2):97-102.
- [4] 于洋.大农业投资热度不减[J].农经,2012(2):40-43.
- [5] 潘洪刚,王礼力,王君萍.中国农业创业投资发展轨迹及趋势[J].电子科技大学学报:社科版,2008(4):24-28.
- [6] 魏喜武.创业投资向农业流动的障碍因素研究[J].科学管理研究,2009(2):111-115.
- [7] 张春燕.石嫣的农业新冒险[N].中国环境报,2013-09-11(008).
- [8] 2014 年中央一号文件解读[J].茶世界,2014(2):16-20.
- [9] 高锁平.农业利用创业投资问题研究[D].北京:中国农业科学院,2006.
- [10] 张毅,黄金生.支农惠农项目经营风险管理探析[J].成都行政学院学报,2010(6):57-60.
- [11] 吴芸.CSA——蕴涵东方哲学的新型农业生产[J].中国图书评论,2012(6):118-120

(上接第 8045 页)

参考文献

- [1] 张铮,王艳萍,薛桂香.数字图像处理与机器视觉[M].北京:人民邮电出版社,2011.
- [2] 李长缨,滕光辉,赵春江,等.利用计算机视觉技术实现对温室植物生

长的无损监测[J].农业工程学报,2003(3):140-143.

- [3] 王琪,万中南,韩俊伟.基于图像中心矩和特征向量的目标识别方法[J].激光与红外,2009(8):895-898.
- [4] 杨树仁,沈洪远.基于相关向量机的机器学习算法研究与应用[J].计算技术与自动化,2010(1):43-47.