

基于机器视觉的胡麻种子自动化考种方法

毛永文, 韩俊英*, 刘成忠

(甘肃农业大学 信息科学技术学院, 甘肃兰州 730000, 中国)

摘要: [目的/意义] 胡麻种子的周长、面积、长短轴和千粒重是胡麻考种过程中常用的参数, 对于胡麻的育种、栽培, 以及种子品质和性状的评估都具有重要的意义。[方法] 针对胡麻种子自动化考种时出现的数据统计错误率高、效率低等问题, 基于机器视觉研究胡麻种子的轮廓特点、探索形态特征的测量方法, 针对籽粒重叠现象提出基于融合角点特征的轮廓拟合图像分割方法, 设计胡麻种子自动化考种数据实时分析系统, 最终实现胡麻种子自动化考种的研究。本研究在工业相机获取的胡麻种子图像上进行试验。[结果和讨论] 提出的自动化考种方法对不同品种胡麻种子的统计识别准确率达97.28%, 百粒种子平均处理时长69.58 ms, 相较于极限腐蚀算法、基于距离变换的分水岭算法, 平均计算准确率比极限腐蚀算法提升19.6%, 平均运算时间低于直接使用分水岭算法所需时间。[结论] 自动化考种方法具有更好的计算准确率和处理速度, 能够更准确地批量获取胡麻种子的形态学特征参数, 使测量误差能够保持在10%以内, 可为今后胡麻考种相关工作提供技术支撑, 助力相关产业发展。

关键词: 胡麻种子; 机器视觉; 自动化考种; 图像分割; 软件系统

中图分类号: S565; TP391.1; TP18

文献标志码: A

文章编号: SA202309011

引用格式: 毛永文, 韩俊英, 刘成忠. 基于机器视觉的胡麻种子自动化考种方法[J]. 智慧农业(中英文), 2024, 6(1): 135-146. DOI: 10.12133/j.smartag.SA202309011

MAO Yongwen, HAN Junying, LIU Chengzhong. Automated Flax Seeds Testing Methods Based on Machine Vision[J]. Smart Agriculture, 2024, 6(1): 135-146. DOI: 10.12133/j.smartag.SA202309011 (in Chinese with English abstract)

0 引言

胡麻具有生长周期短、适应性强的特点^[1], 是中国北方地区主要的经济作物之一, 因其用途广泛、品质特殊, 在中国油料作物和纤维作物中占有重要地位^[2]。种子是作物的芯片, 胡麻种子质量的好坏直接关系到胡麻植株产量的高低。目前, 胡麻植株抗倒伏差、籽粒产量低等问题制约着胡麻产业的发展^[3]。为保证农作物的优质生产, 通过了解胡麻籽粒表型特征、根据育种目标选用不同的资源做亲本等方法, 对于提高育种效率以及开展胡麻考种工作具有十分重要的指导意义^[4]。

考种是胡麻育种过程中的重要环节。关于胡麻考种时形态方面的相关参数主要包括胡麻籽粒的周

长、面积、长短轴、千粒重等^[5,6]。胡麻种子的大小和轮廓特征各不相同, 通过测量这些数据可以帮助人们区分不同品种、筛选饱满健康的胡麻种子, 并为后期考种分类工作提供重要依据。当前, 考种的数据统计方式主要分为人工统计与机器视觉统计两种。传统的人工统计方式是当前运用最为广泛的方式, 但大量且繁琐的统计工作不仅对统计人员专注力要求极高, 而且由于胡麻种子相对其他农作物籽粒而言颗粒更小、轮廓更奇异, 使胡麻考种工作不仅耗时而且难以利用现有物理手段得到精确形态数据。机器视觉技术的出现则从根本上解决了胡麻考种过程中数据统计不精准、测量主观性强等问题, 实现了农业产业的智能化、现代化与自动化。

收稿日期: 2023-09-11

基金项目: 中国甘肃省高校创新基金项目(2021A-056); 中国甘肃省高校产业扶持与引导项目(2021CYZC-57)

作者简介: 毛永文, 研究方向为农业信息化。E-mail: halomyw@163.com

*通信作者: 韩俊英, 教授, 研究方向为机器视觉、深度学习在农业中的应用。E-mail: hanjy@gsau.edu.cn

copyright©2024 by the authors

针对农作物考种, 吴迪^[7]提出基于机器视觉与深度学习算法的大豆籽粒考种与选种研究, 并设计出一套基于计算机视觉技术的玉米考种系统, 从多个方面优化了现有算法, 实现了自动化、高通量地获取玉米表型性状参数。该系统具有精度高、效率高、全自动等特点。周洪垒^[8]完成了基于图像处理的水稻考种系统的设计与实现, 结合 Qt Designer 软件设计了水稻考种系统, 实现了对水稻种子的自动计数, 较适合对类圆锥形的农作物种子进行计数及形态学分析。周成全等^[9]利用机器视觉与深度学习技术, 对西兰花表型快速提取方法进行了研究, 可以为作物田间长势监测研究提供重要参考。李锦明^[10]针对玉米籽粒提出基于机器视觉的玉米考种技术研究, 通过机器视觉、数字图像处理, 以及深度学习等技术, 对玉米籽粒的多种外形参数和品种的分类进行了研究。Zhang 等^[11]针对小麦、玉米等作物提出了一种基于短轴的腐蚀算法和改进的角点法, 开发了基于 Android 设备的应用程序, 可以用于快速计算粮食数量。Liu 等^[12]根据颗粒接触区域的角点数和封闭区域数与颗粒接触区域内颗粒数的关系, 提出了一种颗粒计数方法, 并开发了基于 Android 的谷物计数移动应用程序。

针对图像粘连分割问题, 马正华等^[13]提出基于极限腐蚀的重叠苹果果实分割方法, 对果实的正确分割率达到 96.5%。吴叔珍等^[14]提出基于凹点寻找标记的分水岭算法来分割粘连谷粒, 对 5 种谷物的粘连谷粒图像进行了分割测试, 结果获得了 96.4% 的谷粒计数准确率。柳冠伊等^[15]综合分水岭算法的形态学概念与区域生长算法迭代生长的思想, 对玉米果穗进行了分割, 粘连籽粒的分割准确率为 94.9%。Liang 等^[16]提出了一种针对不同种类、形状和大小的复杂接触颗粒的分割新方法, 在分层分水岭算法的基础上, 设计了一种分割线检测算法, 实现了复杂接触颗粒的精确分割, 平均测量精度达到 99.65%。从以上研究方法中可以看出, 这些基于形态学进行图像分割计数的方法, 大多都是对图像本身进行处理, 通过对图像的不断变换, 使得图像种子原始像素比例发生了明显的改变, 最终得到了籽粒与籽粒之间明显的区分界线。这些分割方法虽然统计准确率较高, 但只能适用于籽粒计数工

作, 而不能更好地在已有图像分割基础上对籽粒其他形态数据进行精确的计算与统计。

综上所述, 不同品种的农作物需要根据实验对象找寻适合的方法。本研究针对胡麻考种实际需求, 设计出胡麻种子自动化考种数据实时分析系统, 并且针对籽粒可能出现的重叠现象提出基于融合角点特征的轮廓拟合图像分割方法, 在尽可能分割胡麻种子轮廓的基础上, 最大限度保留胡麻籽粒原始图像特征, 获取胡麻种子的形态数据以及考种信息。

1 图像采集与处理

1.1 图像采集

试验胡麻种子分别为 XB1、NM-21-10、ZC-62、ZZ-173、1009-1、08006-375、张亚 2 号、坝亚 21 号和同白亚 3 号, 共计 9 个品种, 由甘肃省农业科学院提供。

图像采集系统由工业相机 (维视智造 MV-HP505GM (500 万像素))、环形光源、LED 灯板、镜头固定槽、支架、微型计算机 (Intel (R) Core (TM) i7-7700HQ CPU @ 2.80 GHz、NVIDIA GeForce GTX 1060 with Max-Q Design、16 G 内存) 等组成, 试验采用的系统环境为 Windows10、编译环境为 Python3.10。该系统可以快速、批量获取胡麻种子图像并进行数据处理分析。胡麻种子图像采集设备如图 1 所示。

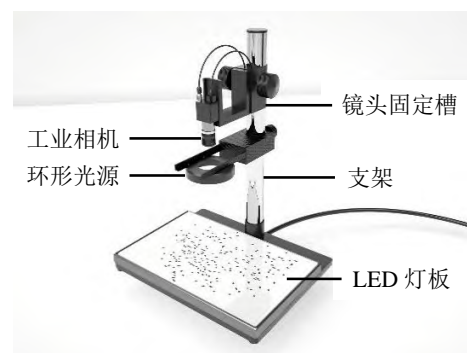


图 1 胡麻种子图像采集设备

Fig. 1 Image acquisition equipment for flax seeds

1.2 图像处理流程

首先对胡麻种子图像进行预处理, 将采集的彩

色图像转换为灰度图像，并对图像进行滤波平滑处理，得到二值化图像^[17]。针对粘连种子图像，基于融合角点特征的轮廓拟合图像分割方法，通过自适应阈值筛选对种子粘连区域进行轮廓逼近、角点

检测、轮廓拟合等处理，实现胡麻种子图像的分割计数，最终通过像素点转化，实现胡麻种子数据的汇总。图像处理流程图如图2所示。

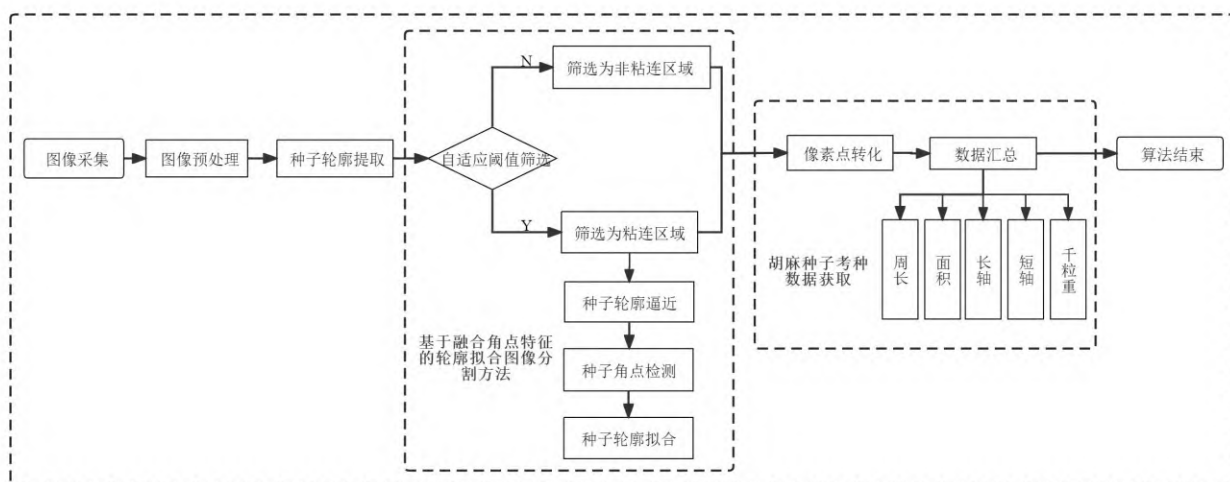


图2 胡麻籽粒自动化考种图像处理流程

Fig. 2 Image processing flow for automated flax seeds testing

2 基于融合角点特征的轮廓拟合图像分割方法

2.1 胡麻种子轮廓特点及空间排布规律

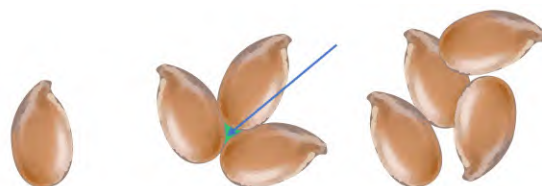
胡麻种子大致轮廓呈扁卵形^[18]，表面光滑且有光泽，通常呈褐色或淡黄色，面积较小且长度大于宽度，整体长约4~7 mm，宽约2~4 mm，厚约1.5 mm，种子一端较为钝圆，另一端则略尖，尖端会向一侧倾斜，存在轻微变形或者局部凹陷等情况。胡麻种子相较于其他农作物籽粒颗粒较小，该特点可能会造成人工统计籽粒误差大、测量形态数据不准确的问题，而且由于胡麻种子尖端倾斜、局部凹陷的特点，可能会影响现有凹点匹配算法对于胡麻种子分割的精确程度^[19]。胡麻种子示意图如图3所示。



图3 胡麻种子示意图

Fig. 3 Schematic image of flax seeds

胡麻种子的空间排布结构根据有无角点及孔洞的情况大致可以分为3种：①单籽粒结构；②多籽粒粘连有孔洞结构；③多籽粒粘连无孔洞结构。单籽粒胡麻种子没有角点和孔洞，易于区分，而多籽粒粘连胡麻种子会对轮廓边缘区域造成干扰，影响数据的统计。胡麻种子空间排布结构示意图如图4所示。



(a)单籽粒 (b)多籽粒粘连有孔洞 (c)多籽粒粘连无孔洞

图4 胡麻种子空间排布结构示意图

Fig. 4 Schematic image of spatial arrangement structure of flax seeds

2.2 自适应阈值筛选

采集胡麻种子图像时会出现图4所示的3种空间排布结构。图像中既会存在粘连种子也会存在非粘连种子，若对图像中每一粒胡麻种子进行复杂的

轮廓分割操作,必然会徒增算法计算量,甚至会降低识别的准确率,也不能满足实时自动化考种的需求。因此,当进行图像轮廓边缘检测时,引入自适应阈值筛选,需要首先根据标准差因子对图像中的胡麻种子进行整体筛选,只有满足条件的多籽粒目标区域才会被执行图像分割处理操作,而单籽粒区域则会跳过图像分割步骤直接进行考种数据的汇总。

图像中的像素灰度值服从正态分布,位于该分布平均值减去 η 倍标准差之间的像素就会在整体数据中占据相应大小的比例范围。因此,只需得到合适的标准差因子,就可以根据阈值将该范围内的目标像素认作粘连种子区域,而该范围以外的目标像素就是非粘连种子区域,从而可以进行排除。同时,由于每幅图像胡麻种子周长的平均值以及标准差是不一致的,所以采用该方法就可以达到自适应阈值筛选的效果,从而得到一个相对可靠的阈值。自适应阈值 F 反映物体形态特征在整体数据中的包容度,其计算见公式(1)。

$$F = A - \eta S \quad (1)$$

式中: A 为胡麻种子轮廓像素周长的平均值; η 为标准差因子; S 为胡麻种子轮廓像素周长的标准差。

每个品种胡麻种子周长大小、粘连情况各不相同。从9种胡麻种子样本中随机挑选5个品种(1009-1、XB1、ZC-26、ZZ-173和张亚2号),每个品种选择50、100、150和200粒时图像,将提取轮廓后得到的胡麻种子区域数与实际颗粒数进行对比,二者对比结果之差的绝对值小于10的图像判定为正确划分的图像,统计各个标准差因子下正确划分的图像数,从而进行最佳标准差因子的计算。胡麻种子标准差因子分布如图5所示。

为获得更加直观的图像数据,将上文随机挑选的品种数据扩大10倍,并将正确划分图像数为0的数据设置为1。由图5可知,当标准差因子数值取4时,图中横坐标数值最高,意味着该标准差因子可以在最大程度上对粘连种子区域进行划分,因此,本研究将4作为标准差因子,以其自适应阈值作为判断轮廓是否粘连的标准。

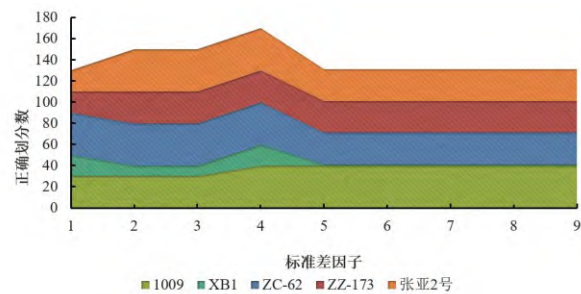


图5 胡麻种子轮廓像素标准差因子分布

Fig. 5 Distribution of pixel standard deviation factors for the contour of flax seeds

2.3 轮廓逼近

图像的轮廓由若干个像素点组成。这些像素点中可能存在多余、重复或不必要的信息。轮廓逼近就是通过一定的近似算法在给定图像轮廓信息的基础上,将图像轮廓形状近似为另一种含有少量顶点的轮廓形状,并将原始的轮廓曲线简化为更加规则、紧凑的线段或多边形的算法。一个理想的轮廓逼近算法应当尽可能小地降低计算的复杂度,将图像中所有线条形状特点尽可能地标记出来^[20]。在实际应用中,出于计算速度与计算精度的考虑,该算法常用于简化轮廓、减少噪声干扰和提高程序运行效率,具体算法过程如下:

假设采集的胡麻种子图像轮廓信息中有 n 个像素点,每个像素点用 $P = \{P_0, P_1, P_2, \dots, P_n\}$ 表示。在图像轮廓曲线上选择起点 P_0 和终点 P_n ,将这两点之间的连线定义为一条线段,计算任意一条 P_0 到 P_n 之间的轮廓曲线子线段与曲线的距离,并且找到线段距离的最远点 P_k 与其所对应的距离,将距离最远点 P_k 与预设阈值进行比较,如果该距离小于预设阈值,则线段上所有的点都可以用直线连接 P_0 和 P_n 来代替;如果该距离大于或等于预设阈值,则需要将曲线分为 P_0 到 P_k 和 P_k 到 P_n 两段,随后分别对这两段执行上述递归过程。这样就可以得到一个尽可能接近原始曲线的轮廓逼近结果,其中预设阈值则代表轮廓的逼近程度,而越小的阈值会使得逼近结果更加精确,但可能会导致出现过多的顶点;相反,越大的阈值会使得结果中的顶点数量减少,但可能会丢失一些曲线的细节信息。经过这样多次反复递归迭代,最终可以得到一个由少量顶点组成的

轮廓曲线，从而实现轮廓的逼近处理。胡麻种子轮廓逼近示意图如图6所示。

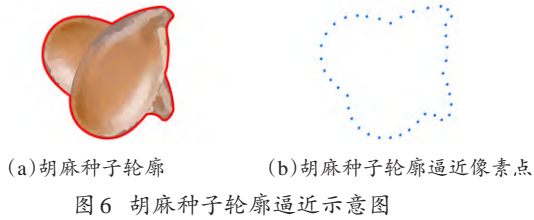


图6 胡麻种子轮廓逼近示意图

2.4 角点检测

在图像处理领域中，角点通常指图像中出现两个或多个轮廓且边缘相交的像素点。角点检测则是机器视觉领域中常用于检测图像角点信息的算法，其基本思想在于图像角点在不同方向上的像素强度变化差异较大，所以基于像素灰度值的二阶导数矩阵可以实现角点的检测。在上述图像轮廓逼近算法的基础上，进行轮廓角点检测，可以得到每个角点的坐标，将得到的轮廓点及角点作为输出，则能够进一步提高后续轮廓拟合方法的精度和准确率，具体算法如下。

首先，需要对灰度图像中的每个像素点进行操作，计算出该像素点所在位置图像局部区域的二阶导数矩阵 \mathbf{M} （结构矩阵）；其次，通过结构矩阵 \mathbf{M} 计算出响应函数 R 的值；最后，根据其响应函数 R 值来判断其是否为角点。具体地，对于坐标 (x, y) 处的像素，其二阶导数矩阵 \mathbf{M} 的定义如公式（2）。

$$\mathbf{M} = \begin{bmatrix} \sum [w(x, y) \times I_x^2] & \sum [w(x, y) I_x I_y] \\ \sum [w(x, y) I_x I_y] & \sum [w(x, y) \times I_y^2] \end{bmatrix} \quad (2)$$

式中： I_x 和 I_y 是指像素点周围 x 、 y 方向上的梯度； $w(x, y)$ 是不同位置像素值加权的高斯窗口函数； $\sum [w(x, y) \times I_x^2]$ 表示 x 方向上梯度加权和平方的和； $\sum [w(x, y) I_x I_y]$ 表示 x 、 y 方向上梯度乘积的加权和平方的和； $\sum [w(x, y) \times I_y^2]$ 表示 y 方向上梯度加权和平方的和。通过以上对每个像素点的计算，就可以得到该像素点周围的结构矩阵，计算出相应的响应函数 R 。响应函数 R 的定义如公式（3）。

$$R = \det(\mathbf{M}) - k \times \text{trace}(\mathbf{M})^2 \quad (3)$$

式中： $\det(\mathbf{M})$ 为 \mathbf{M} 的行列式； $\text{trace}(\mathbf{M})$ 为 \mathbf{M} 的迹。图像中的每个像素点响应函数越大，则越有

可能是角点，同时根据初始预设的阈值，只有当像素的响应函数 R 大于该阈值时，才会被视为角点。胡麻种子角点检测示意图如图7所示。

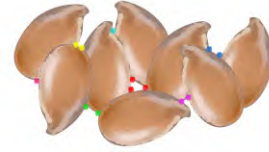


图7 胡麻种子角点检测示意图

Fig. 7 Schematic image of corner point detection of flax seeds

2.5 轮廓拟合

轮廓拟合是通过对检测到的图像曲线轮廓进行拟合，得出其形状及相关参数的过程，可以更加精细地描述物体的形状，同时去除轮廓离散点带来的噪声。在机器视觉领域中，常用的轮廓拟合方法包括最小二乘法拟合、多项式拟合等。对比多项式拟合，最小二乘法拟合稳定性高、计算过程简单、对异常值有一定的鲁棒性，可以减少干扰点对拟合结果的影响。根据上文对胡麻种子轮廓特点的分析，本研究采用最小二乘法对胡麻种子轮廓进行椭圆拟合，具体算法过程如下。

首先，为提高轮廓拟合的精度，本研究采用自适应阈值筛选、轮廓逼近、角点检测等方法获取了图像轮廓二维离散点数据集，假设该数据集有 n 个二维离散数据点 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ ，利用这些数据点就可以计算出相应的协方差矩阵，通过该矩阵的特征向量和特征值则能够得到椭圆的长轴、短轴、旋转角度等几何参数。具体地，椭圆可以用公式（4）表示。

$$(x - cx)^2/a^2 + (y - cy)^2/b^2 = 1 \quad (4)$$

式中： (cx, cy) 代表椭圆的中心坐标； a 和 b 分别代表椭圆的长轴和短轴。

其次，在拟合椭圆轮廓时，将椭圆与数据点之间的距离平方和最小化，利用最小二乘法找到一组最优的椭圆参数（中心坐标、长轴、短轴、旋转角度），使得该椭圆到所有点的距离平方和最小，以尽可能减小实际观测值与理论值之间的距离差。

最后，将轮廓点数据集作为输入，计算得到图像轮廓的几何矩，基于最小二乘法及椭圆形状的几何特征，通过数值优化方法求出椭圆参数的最优

解,拟合出给定图像的最佳轮廓。胡麻种子轮廓拟合图像如图8所示。



图8 胡麻种子轮廓拟合图像

Fig. 8 Fitted image of flax seeds contour

3 试验结果与分析

为验证本研究方法针对胡麻种子图像分割的优势,随机选取一幅胡麻种子粘连图像,分别利用极限腐蚀算法^[21]、基于距离变换的分水岭算法^[22],以及本研究方法进行图像分割测试,各方法对比如图9所示。

试验中,二值化图像共有186颗胡麻籽粒,极限腐蚀算法分割识别到180颗胡麻籽粒;基于距离变换的分水岭算法分割识别到176颗胡麻籽粒;本研究方法分割识别到了所有胡麻籽粒。可见,本研究方法在胡麻籽粒的分割中较为准确。

由图9(b)可知,极限腐蚀算法有着较强的图像分割能力,但也可以看出该算法并没有如图中橙色箭头所示,将粘连严重的胡麻种子区域全部分开,同时,该算法很大程度上取决于用户预设的初始腐蚀程度。当腐蚀程度过小时,会出现粘连种子区域没有被分开的情况;而当腐蚀程度过大时,则会出现某些较小种子区域消失的情况。并且由于该算法主要依赖于对图像轮廓的腐蚀,导致使用该算法处理后的胡麻种子图像区域会比实际二值化种子图像区域小很多,造成原轮廓变形,不能满足本研究后期自动化考种的需求。

相较于极限腐蚀算法,基于距离变换的分水岭算法最大不同在于该算法保留了胡麻种子图像区域的完整轮廓,该算法对于轻微粘连的胡麻种子图像区域分割效果很好,但是如图9(c)中橙色箭头所示,该算法对于严重粘连的胡麻种子图像区域分割能力较差,同时,如图9(c)中黄色箭头所示,该

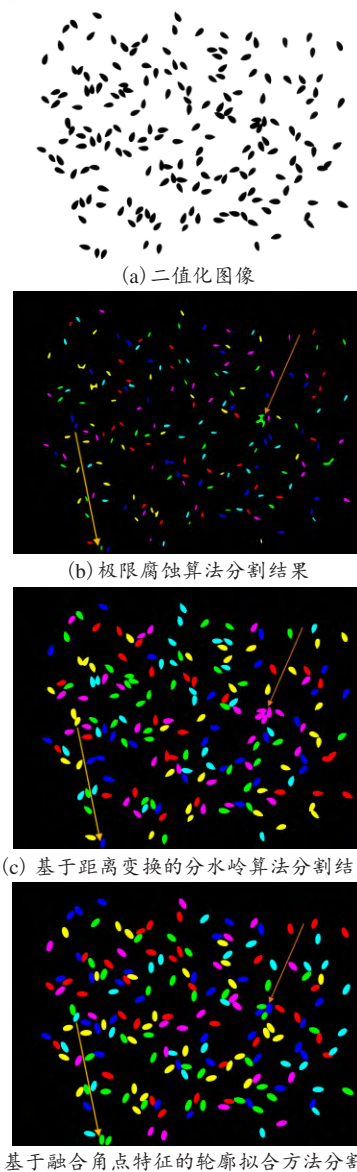


图9 胡麻种子图像分割方法对比图

Fig. 9 Comparison of methods for image segmentation of flax seeds

算法也出现了丢失种子轮廓信息的情况。除此之外,该算法的计算量庞大,在处理大规模图像时,需要消耗大量时间和计算资源,同样,也不能满足本研究后期自动化考种的需求。

同时,本研究所采用的基于融合角点特征的轮廓拟合图像分割方法,相较于以上两种算法而言,不仅对于胡麻种子粘连图像区域分割效果较好,并且能够在保证分割准确率的同时,最大程度上绘制出胡麻种子的完整轮廓特征,如图9(d)中橙色、黄色箭头所示,本研究基于融合角点特征的轮廓拟

合图像分割方法并没有出现上述两种方法中，粘连严重区域未分割或丢失种子轮廓信息等情况，同时该方法数据计算量小、运算速度快，能够满足系统自动化考种的需求。

为进一步验证本研究方法针对胡麻种子图像检测具有的优势及识别准确率，选取剩余4个不同品种的胡麻种子（同白亚3号、坝亚21号、NM-21-10和08006-375）并采集图像样本，每种样本采集3组，种子数依次为100、150和200粒，分别利用极限腐蚀算法、分水岭算法，以及本研究方法进行准确率和运算时间的测试。准确率计算如公式（5）。

$$P(\%) = (1 - |T - R| / T) \times 100 \tag{5}$$

式中：*P*为计算准确率；*T*为胡麻种子实际粒数；*R*为胡麻种子统计粒数。

从表1可以看出，针对不同品种的胡麻种子，使用3种方法对相同图像进行处理时，本研究方法在胡麻种子计算准确率方面，结果优于极限腐蚀算法，略高于基于距离变换的分水岭算法，同时在运算时间方面，基本与极限腐蚀算法相当，但远快于基于距离变换的分水岭算法。通过表1中坝亚21号

样本的数据可以看到，极限腐蚀算法准确率下降到41.94%，由此可以看出极限腐蚀算法十分依赖预设腐蚀程度，说明该方法鲁棒性差，易产生严重误差；同时，观察4组样本整体试验数据可知，基于距离变换的分水岭算法平均运算时间都在3 400 ms，该方法图像处理所需时间远远超出其他两种方法，由此也可以看出，该算法庞大的数据计算量。同时，本研究方法在同白亚3号和08006-375样本中也出现了统计籽粒个数超出实际个数、统计时出现正偏差的问题。该问题是由于算法在严重粘连种子区域中出现轮廓过拟合，使两个或多个重叠种子相交区域出现多余轮廓信息，导致分割轮廓过多，亟待改进。总体而言，相较于极限腐蚀算法和基于距离变换的分水岭算法，本研究算法在保证较快运算时间的情况下，依然能够获得最高的计算准确率。经过4组试验测试，本研究算法平均准确率达到97.28%，平均运算时间为69.58 ms，平均准确率比极限腐蚀算法高19.6%，平均运算时间比分水岭算法快3 361 ms（约49倍），能够满足当前胡麻种子自动化考种时图像处理的需求。

表1 胡麻种子图像分割方法对比表

Table 1 Comparison of image segmentation methods for flax seeds

品种	编号	个数	各方法统计种子数			各方法计算准确率/%			各方法运算时间/ms		
			极限腐蚀 (正负偏差)	分水岭 (正负偏差)	本文算法 (正负偏差)	极限腐蚀	分水岭	本文算法	极限腐蚀	分水岭	本文算法
同白亚3号	1	100	93(-7)	98(-2)	107(+7)	93.00	97.96	93.00	52	2 353	64
	2	150	137(-13)	146(-4)	160(+10)	91.33	97.26	93.33	56	3 778	66
	3	200	180(-20)	193(-7)	201(+1)	90.00	96.37	99.00	62	4 578	72
	平均值	150	136.67	145.67	156.00	91.44	97.20	95.11	56.67	3 569.67	67.33
坝亚21号	1	100	36(-64)	95(-5)	96(-4)	36.00	95.00	96.00	46	2 390	62
	2	150	65(-85)	141(-9)	150-	43.33	94.00	100.00	47	3 437	69
	3	200	93(-107)	179(-21)	200-	46.50	89.50	100.00	49	4 350	71
	平均值	150	64.67	138.33	148.67	41.94	92.83	98.67	47.33	3 392.33	67.33
NM-21-10	1	100	85(-15)	97(-3)	99(-1)	85.00	97.00	99.00	53	2 221	68
	2	150	125(-25)	141(-9)	149(-1)	83.33	94.00	99.33	55	3 280	75
	3	200	169(-31)	188(-12)	200-	84.50	94.00	100.00	64	4 157	79
	平均值	150	126.33	142.00	149.33	84.28	95.00	99.44	57.33	3 219.33	74.00
08006-375	1	100	92(-8)	96(-4)	105(+5)	92.00	95.83	95.00	52	2 411	64
	2	150	139(-11)	145(-5)	158(+8)	92.67	96.55	94.67	56	3 453	71
	3	200	189(-11)	194(-6)	204(+4)	94.50	96.91	98.00	61	4 764	74
	平均值	150	140.00	145.00	155.67	93.06	96.43	95.89	56.33	3 542.67	69.67
平均值		150	116.92	142.75	152.42	77.68	95.37	97.28	54.42	3 431.00	69.58

注：(+)表示出现正偏差；(-)表示出现负偏差；-表示无正负偏差。

4 自动化考种与系统运行

4.1 胡麻种子形态学特征参数测定

上文提出的基于融合角点特征的轮廓拟合图像分割方法在最大程度上保留了原始图像的特征信息，可以很好地针对胡麻种子轮廓特征信息进行图像分割，也为实现胡麻自动化考种的数据获取提供了强力的技术支撑。

数字图像中的每个像素点都是大小、形状均匀的小方块，所以胡麻种子的周长、面积、长短轴可以利用籽粒在图像中所占的像素点个数总和来表示^[23]，而胡麻种子的千粒重则可以通过籽粒总重量与轮廓检测得到的籽粒总个数作比值运算，得出单个籽粒的重量，从而换算获得^[24]。针对胡麻自动化考种过程中需要将籽粒的图像像素点转换为实际形态大小的问题，本研究使用中国人民银行发行的2019版第五套人民币中的1元和1角硬币，作为像素点转换的参考对象。硬币是日常生活中常见的物品，十分容易获取，并且大小、厚度适中，便于采集图像信息，同时，硬币的材质质地坚硬，不易发生变形及磨损的情况。此外，由于硬币是国家统一发行的货币，对于其尺寸都有着严格的标准，所以使用硬币进行像素点信息的转化更加精确。

根据中国人民银行发布的相关数据可知^[25]，2019版第五套人民币1元硬币的直径为22.25 mm，周长约为69.90 mm，面积约为388.82 mm²，1角硬币的直径为19 mm，周长约为59.69 mm，面积约为283.53 mm²。

假设采集的硬币实际周长为 C_1 ，实际面积为 S_1 ，实际直径为 D_1 ，硬币周长像素点个数为 C_2 ，面积像素点个数为 S_2 ，直径像素点个数为 D_2 ，那么在一定高度下， $C=C_1/C_2$ 代表单个周长像素点实际大小，同理 $S=S_1/S_2$ 、 $D=D_1/D_2$ 代表单个像素点面积、直径的实际大小，最终只需要将像素点个数乘以对应像素点大小，即可得到实际胡麻种子的周长、面积、长短轴数据。为验证该方法的可行性，选用中国人民银行发行的全新没有磨损及使用痕迹的2019版第五套人民币中的1元和1角硬币，并同时1元硬币和1角硬币垂直置于距离相机物距29.5 cm（白板完全置于相机视野内）正对上方位置进行图

像采集。

通过轮廓检测计算得到1元硬币周长、面积、直径像素点实际大小分别为0.069 962、0.005 464和0.074 167；1角硬币周长、面积、直径像素点实际大小分别为0.074 293、0.006 192和0.078 603。对硬币像素点数据进行对比分析，可得图像像素点转化正确率为95%左右。

如表2所示，通过像素点转化率测得不同品种的胡麻种子平均短轴长度为3.19 mm、平均长轴长度为6.17 mm、平均周长长度为14.24 mm、平均面积大小为16.07 mm²。

表2 胡麻种子形态数据
Table 2 Morphological data of flax seeds

品种	编号	个数	短轴/ mm	长轴/ mm	周长/ mm	面积/ mm ²
同白亚3号	1	100	3.13	6.20	14.23	15.59
	2	150	3.06	6.13	14.01	15.52
	3	200	2.92	5.85	13.37	14.37
	平均值	150	3.04	6.06	13.87	15.16
坝亚21号	1	100	3.43	6.60	15.24	18.14
	2	150	3.25	6.31	14.52	17.20
	3	200	3.42	6.59	15.21	17.98
	平均值	150	3.37	6.50	14.99	17.77
NM-21-10	1	100	3.22	5.91	13.83	15.94
	2	150	3.18	5.79	13.60	15.34
	3	200	3.29	6.04	14.15	16.22
	平均值	150	3.23	5.91	13.86	15.83
08006-375	1	100	3.05	6.03	13.83	15.01
	2	150	3.17	6.24	14.33	15.64
	3	200	3.21	6.34	14.56	15.90
	平均值	150	3.14	6.20	14.24	15.52
平均值		150	3.19	6.17	14.24	16.07

同时，为验证本研究方法获取到的胡麻种子形态数据的准确性，试验对上述同白亚3号、坝亚21号、NM-21-10、08006-375这4个品种，分别取编号为1图像识别后的第1、2、3粒胡麻种子，使用螺旋测微仪进行测量，结果如表3所示。

从表3可以看出，由考种系统处理得到的胡麻籽粒平均短轴长度为3.19 mm，而通过实际测量获取到的胡麻籽粒平均短轴长度为3.14 mm，同时，由考种系统处理得到的胡麻籽粒平均长轴长度为

表3 胡麻种子测量数据对比

Table 3 Comparison of measured data for flax seeds

品种	编号	考种短轴/mm	考种长轴/mm	测量短轴/mm	测量长轴/mm	平均误差/%
同白亚3号	1	3.13	6.20	2.91	5.78	7.41
	2	3.06	6.13	2.98	5.83	3.92
	3	2.92	5.85	2.72	5.68	5.71
	平均值	3.04	6.06	2.87	5.76	5.50
坝亚21号	1	3.43	6.60	3.28	6.19	5.60
	2	3.25	6.31	3.03	6.10	5.35
	3	3.42	6.59	3.17	6.15	7.52
	平均值	3.37	6.50	3.16	6.15	6.16
NM-21-10	1	3.22	5.91	3.48	6.12	5.45
	2	3.18	5.79	3.23	6.05	2.92
	3	3.29	6.04	3.52	6.19	4.48
	平均值	3.23	5.91	3.41	6.12	4.28
08006-375	1	3.05	6.03	3.27	6.24	5.05
	2	3.17	6.24	2.98	6.12	4.17
	3	3.21	6.34	3.13	6.01	4.02
	平均值	3.14	6.20	3.13	6.12	4.41
平均值		3.19	6.17	3.14	6.04	5.09

注：*由于胡麻籽粒很小，手工测量存在误差，最终取3个人测量的平均值。

6.17 mm，而通过实际测量获取到的胡麻籽粒平均长轴长度为6.04 mm。由此可以得出，本研究所提出的针对胡麻种子形态学特征参数的测定方法，短轴和长轴的平均误差为5%，整体测量误差在10%的范围以内，能够满足胡麻种子自动化考种的实际需求。

4.2 胡麻种子自动化考种数据实时分析系统

目前，研究人员已经根据不同需求为各类农作物设计开发了广泛多样的系统软件^[26-27]。本研究采用图像处理技术及基于融合角点特征的轮廓拟合图像分割方法，基于OpenCV（Open Source Computer Vision Library）结合Qt Designer设计了胡麻种子自动化考种数据实时分析系统。系统运行界面展示图如图10所示。

当用户通过选择目录功能，指定图像路径后，执行图像处理功能，即可在相同目录下生成对应的图像处理文件夹，自动存放处理后的图像数据。图像处理功能展示图如图11所示，左图为原始图像，右图为处理后图像。

当用户通过图片处理功能获取到处理的图像数据以后，即可通过执行数据汇总功能，自动生成对应数据表格文件夹，存放相应的图像数据汇总表格（包括个数、面积、周长、长轴、短轴、千粒重等）。数据汇总展示图如图12所示。



图10 胡麻种子考种系统运行界面展示图

Fig. 10 Display diagram of the operating interface of flax seeds testing system

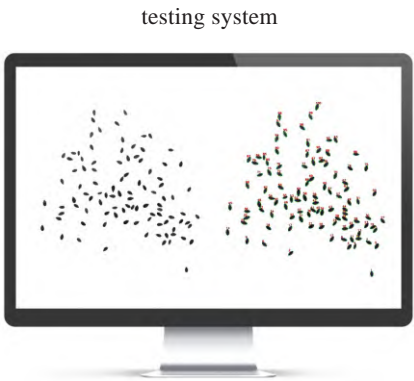


图11 胡麻种子考种系统图像处理展示图

Fig. 11 Image processing display diagram of flax seeds testing system

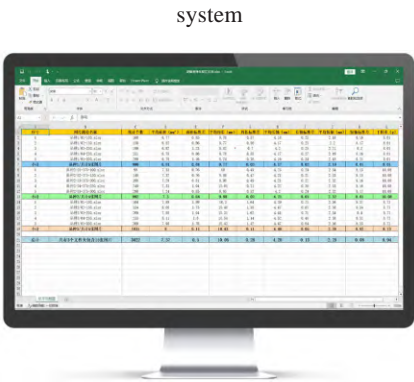


图12 胡麻种子考种系统数据汇总展示图

Fig. 12 Summary and display of data for the flax seeds testing system

5 讨论与结论

本研究针对胡麻种子重叠现象提出基于融合角

点特征的轮廓拟合图像分割方法,并根据该方法设计出胡麻种子自动化考种数据实时分析系统,能够实时处理、获取胡麻种子图像的周长、面积、长短轴、千粒重等数据信息。该方法通过自适应阈值筛选得到粘连种子图像区域信息,利用轮廓逼近、角点检测获得图像实际标注点,并对标注点图像区域进行轮廓拟合,完成图像的分割。该方法相较于极限腐蚀算法及基于距离变换的分水岭算法,不仅具有较高的准确率,而且也具有更快的处理速度,平均计算准确率比极限腐蚀算法提高19.6%,平均运算时间低于直接使用分水岭算法所需时间,能够在保证分割准确率的同时,最大程度上绘制出胡麻种子的完整轮廓特征,从而获取到精确的籽粒形态数据。考虑到自动化考种的实际需求,该方法没有使用膨胀、腐蚀等方法对图像进行形态学处理,所以在最大程度上保留了图像的原始特征,该方法平均计算准确率达97.28%,运算时间短,具有较强的鲁棒性、适用性,并且测量误差能够保持在10%的范围以内,能够满足胡麻种子自动化考种的实际需求。

课题组未来的研究方向致力于胡麻种子自动化考种数据实时分析系统的硬件与软件集成一体化研究,使软件系统的设计及使用更加人性化,使用户不再因系统操作晦涩难懂而望而却步。并且,针对本研究算法在严重粘连种子区域中出现轮廓过拟合,导致分割轮廓过多的问题,也需要进一步改进和优化相关算法,提高胡麻籽粒分割及识别的准确率,实现胡麻种子更加精确、快捷地自动化考种。

利益冲突声明: 本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献:

- [1] LI H B, TANG R R, DAI J L, et al. Recent progress in flax fiber-based functional composites[J]. *Advanced fiber materials*, 2022, 4(2): 171-184.
- [2] 马立晓,王树彦,王卓,等.国内外胡麻品种农艺性状及其品质评价[J]. *北方农业学报*, 2017, 45(3): 19-22.
MA L X, WANG S Y, WANG Z, et al. The agronomic traits and quality evaluation of foreign and domestic varieties for oil flax[J]. *Journal of northern agriculture*, 2017, 45(3): 19-22.
- [3] 郭芳,牛俊义,高珍妮.品种密度互作对胡麻抗倒伏特性的影响[J]. *甘肃农业大学学报*, 2016, 51(3): 72-76.
GUO F, NIU J Y, GAO Z N. Effects of variety and density interaction on flax lodging resistance[J]. *Journal of Gansu agricultural university*, 2016, 51(3): 72-76.
- [4] 王炜,陈军,叶春雷,等.甘肃胡麻地方及育成品种农艺性状分析及评价[J]. *中国种业*, 2022(12): 75-81.
WANG W, CHEN J, YE C L, et al. Analysis and evaluation of agronomic characters of local and bred linseed varieties in Gansu province[J]. *China seed industry*, 2022 (12): 75-81.
- [5] IVA S, OSCAR G, MARIE B, et al. Phenotypic evaluation of flax seeds by image analysis[J]. *Industrial crops and products*, 2013, 47: 232-238.
- [6] DANA W, IVO W. Computer image analysis of seed shape and seed color for flax cultivar description[J]. *Computers and electronics in agriculture*, 2008, 61(2): 126-135.
- [7] 吴迪.基于计算机视觉的玉米考种参数获取研究[D].杭州:杭州电子科技大学,2020.
WU D. Research on parameters acquisition of maize seeds' traits based on computer vision[D]. Hangzhou: Hangzhou Dianzi University, 2020.
- [8] 周洪奎.基于图像处理的水稻考种系统的设计与实现[D].成都:电子科技大学,2019.
ZHOU H L. The design and implementation of rice seed counting system based on image processing[D]. Chengdu: University of Electronic Science and Technology of China, 2019.
- [9] 周成全,叶宏宝,俞国红,等.基于机器视觉与深度学习的西兰花表型快速提取方法研究[J]. *智慧农业(中英文)*, 2020, 2(1): 121-132.
ZHOU C Q, YE H B, YU G H, et al. A fast extraction method of broccoli phenotype based on machine vision and deep learning[J]. *Smart agriculture*, 2020, 2(1): 121-132.
- [10] 李锦明.基于机器视觉的玉米考种技术研究[D].杭州:杭州电子科技大学,2019.
LI J M. Research on corn seed testing technology based on machine vision[D]. Hangzhou: Hangzhou Dianzi University, 2019.
- [11] ZHANG J, LIU S P, WU W, et al. Research on a rapid identification method for counting universal grain crops[J]. *PLoS one*, 2022, 17(9): ID e0273785.
- [12] LIU T, CHEN W, WANG Y F, et al. Rice and wheat grain counting method and software development based on Android system[J]. *Computers and electronics in agriculture*, 2017, 141: 302-309.
- [13] 马正华,申根荣,吕继东.基于极限腐蚀的重叠苹果果实分割方法[J]. *江苏农业学报*, 2017, 33(6): 1372-1378.
MA Z H, SHEN G R, LYU J D. Study on the method of separating apple fruits based on limiting corrosion[J]. *Jiangsu journal of agricultural sciences*, 2017, 33(6): 1372-1378.
- [14] 吴叔珍,姚青.基于凹点寻找标记的分水岭算法分割粘连谷粒[J]. *浙江农业科学*, 2017, 58(8): 1401-1405.
WU S Z, YAO Q. Segmentation of sticky grains by watershed algorithm based on pit finding markers[J]. *Journal of Zhejiang agricultural sciences*, 2017, 58(8): 1401-1405.

- [15] 柳冠伊, 刘平义, 魏文军, 等. 玉米果穗粘连籽粒图像分割方法[J]. 农业机械学报, 2014, 45(9): 285-290.
LIU G Y, LIU P Y, WEI W J, et al. Method of image segmentation for touching maize kernels[J]. Transactions of the Chinese society for agricultural machinery, 2014, 45(9): 285-290.
- [16] LIANG N, SUN S S, YU J J, et al. Novel segmentation method and measurement system for various grains with complex touching[J]. Computers and electronics in agriculture, 2022, 202.
- [17] 祝保林. 基于机器视觉的小桐子种子识别分类研究[D]. 开封: 河南大学, 2020.
ZHU B L. Research on seed recognition and classification of *Jatropha curcas* based on machine vision[D]. Kaifeng: Henan University, 2020.
- [18] 李兴华, 方子森, 牛俊义. 胡麻种子形态结构与萌发特性的研究[J]. 种子, 2012, 31(11): 26-30.
LI X H, FANG Z S, NIU J Y. Study on configuration fabric and germination conditions of flax seed[J]. Seed, 2012, 31(11): 26-30.
- [19] 李毅念, 杜世伟, 姚敏, 等. 基于小麦群体图像的田间麦穗计数及产量预测方法[J]. 农业工程学报, 2018, 34(21): 185-194.
LI Y N, DU S W, YAO M, et al. Method for wheatear counting and yield predicting based on image of wheatear population in field[J]. Transactions of the Chinese society of agricultural engineering, 2018, 34(21): 185-194.
- [20] 段鹏, 程文播, 钱庆, 等. 基于瓶颈检测和分水岭算法的重叠宫颈细胞图像分割方法[J]. 中国医疗器械杂志, 2020, 44(1): 7-12.
DUAN P, CHENG W B, QIAN Q, et al. Overlapping cervical cell image segmentation based on bottleneck detection and watershed algorithm[J]. Chinese journal of medical instrumentation, 2020, 44(1): 7-12.
- [21] 姚业浩, 李毅念, 邹玮, 等. 油菜籽粒千粒重图像测定方法[J]. 中国油料作物学报, 2022, 44(1): 201-210.
YAO Y H, LI Y N, ZOU W, et al. Determination method on thousand-seed weight of rapeseed based on image processing[J]. Chinese journal of oil crop sciences, 2022, 44(1): 201-210.
- [22] 万路瑶. 基于图像识别的作物种子自动计数方法研究[D]. 成都: 成都大学, 2020.
WAN L Y. Research on automatic crop seed counting method based on image recognition[D]. Chengdu: Chengdu University, 2020.
- [23] 熊利荣, 郑宇. 基于形态学的稻谷种子品种识别[J]. 粮油加工, 2010(6): 45-48.
XIONG L R, ZHENG Y. Identification of rice seed varieties based on morphology[J]. Cereals and oils processing, 2010(6): 45-48.
- [24] 王刚. 基于机器视觉的玉米千粒重快速检测仪的研制[D]. 长春: 吉林大学, 2012.
WANG G. Development of 1000-kernel corn weight detector based on machine vision[D]. Changchun: Jilin University, 2012.
- [25] 人民币防伪特征[EB/OL]. [2023-09-09]. <http://www.pbc.gov.cn/huobijinyinju/147948/147954/148021/index.html>.
- [26] HU Y, ZHANG Z W. GridFree: A python package of image analysis for interactive grain counting and measuring[J]. Plant physiology, 2021, 186(4): 2239-2252.
- [27] 彭顺正, 岳延滨, 冯恩英, 等. 基于机器视觉的油菜籽计数系统开发与设计[J]. 计算机应用, 2020, 40(S2): 142-146.
PENG S Z, YUE Y B, FENG E Y, et al. Development and design of rapeseed counting system based on machine vision[J]. Journal of computer applications, 2020, 40(S2): 142-146.

Automated Flax Seeds Testing Methods Based on Machine Vision

MAO Yongwen, HAN Junying*, LIU Chengzhong

(College of Information Science and Technology, Gansu Agricultural University, Lanzhou 730000, China)

Abstract:

[Objective] Flax, characterized by its short growth cycle and strong adaptability, is one of the major cash crops in northern China. Due to its versatile uses and unique quality, it holds a significant position in China's oil and fiber crops. The quality of flax seeds directly affects the yield of the flax plant. Seed evaluation is a crucial step in the breeding process of flax. Common parameters used in the seed evaluation process of flax include circumference, area, length axis, and 1 000-seed weight. To ensure the high-quality production of flax crops, it is of great significance to understand the phenotypic characteristics of flax seeds, select different resources as parents based on breeding objectives, and adopt other methods for the breeding, cultivation, and evaluation of seed quality and traits of flax.

[Methods] In response to the high error rates and low efficiency issues observed during the automated seed testing of flax seeds, the

measurement methods were explored of flax seed contours based on machine vision research. The flax seed images were preprocessed, and the collected color images were converted to grayscale. A filtering and smoothing process was applied to obtain binary images. To address the issues of flax seed overlap and adhesion, a contour fitting image segmentation method based on fused corner features was proposed. This method incorporated adaptive threshold selection during edge detection of the image contour. Only multi-seed target areas that met certain criteria were subjected to image segmentation processing, while single-seed areas bypassed this step and were directly summarized for seed testing data. After obtaining the multi-seed adhesion target areas, the flax seeds underwent contour approximation, corner extraction, and contour fitting. Based on the provided image contour information, the image contour shape was approximated to another contour shape with fewer vertices, and the original contour curve was simplified to a more regular and compact line segment or polygon, minimizing computational complexity. All line shape characteristics in the image were marked as much as possible. Since the pixel intensity variations in different directions of image corners were significant, the second derivative matrix based on pixel grayscale values was used to detect image corners. Based on the contour approximation algorithm, contour corner detection was performed to obtain the coordinates of each corner. The resulting contour points and corners were used as outputs to further improve the accuracy and precision of subsequent contour fitting methods, resulting in a two-dimensional discrete point dataset of the image contour. Using the contour point dataset as an input, the geometric moments of the image contour were calculated, and the optimal solution for the ellipse parameters was obtained through numerical optimization based on the least squares method and the geometric features of the ellipse shape. Ultimately, the optimal contour was fitted to the given image, achieving the segmentation and counting of flax seed images. Meanwhile, each pixel in the digital image was a uniform small square in size and shape, so the circumference, area, and major and minor axes of the flax seeds could be represented by the total number of pixels occupied by the seeds in the image. The weight of a single seed could be calculated by dividing the total weight of the seeds by the total number of seeds detected by the contour, thereby obtaining the weight of the individual seed and converting it accordingly. Through the pixelization of the 1 yuan and 1 jiao coins from the fifth iteration of the 2019 Renminbi, a summary of the circumference, area, major axis, minor axis, and 1 000-seed weight of the flax seeds was achieved. Additionally, based on the aforementioned method, this study designed an automated real-time analysis system for flax seed testing data, realizing the automation of flax seed testing research. Experiments were conducted on images of flax seeds captured by an industrial camera.

[Results and Discussions] The proposed automated seed identification method achieved an accuracy rate of 97.28% for statistically distinguishing different varieties of flax seeds. The average processing time for 100 seeds was 69.58 ms. Compared to the extreme erosion algorithm and the watershed algorithm based on distance transformation, the proposed method improved the average calculation accuracy by 19.6% over the extreme erosion algorithm and required a shorter average computation time than the direct use of the watershed algorithm. Considering the practical needs of automated seed identification, this method did not employ methods such as dilation or erosion for image morphology processing, thereby preserving the original features of the image to the greatest extent possible. Additionally, the flax seed automated seed identification data real-time analysis system could process image information in batches. By executing data summarization functions, it automatically generated corresponding data table folders, storing the corresponding image data summary tables.

[Conclusions] The proposed method exhibits superior computational accuracy and processing speed, with shorter operation time and robustness. It is highly adaptable and able to accurately acquire the morphological feature parameters of flax seeds in bulk, ensuring measurement errors remain within 10%, which could provide technical support for future flax seed evaluation and related industrial development.

Key words: flax seeds; machine vision; automated seed testing; image segmentation; software systems

Foundation items: Gansu Provincial University Innovation Fund Project in China (2021A-056); Gansu Provincial University Industry Support and Guidance Project in China (2021CYZC-57)

Biography: MAO Yongwen, E-mail: halomyw@163.com

***Corresponding author:** HAN Junying, E-mail: hanjy@gsau.edu.cn

(登陆 www.smartag.net.cn 免费获取电子版全文)