

# 基于图像处理和聚类算法的待考种大豆主茎节数统计

王跃亭<sup>1</sup> 王敏娟<sup>1</sup> 孙石<sup>2</sup> 杨斯<sup>1</sup> 郑立华<sup>1</sup>

(1. 中国农业大学现代精细农业系统集成研究教育部重点实验室, 北京 100083;

2. 中国农业科学院作物科学研究所, 北京 100081)

**摘要:** 为了实现待考种大豆植株主茎节数的快速、高效测量,提出一种基于图像处理和聚类算法的待考种大豆主茎节数统计方法。首先,获取不同视角下的已脱叶待考种大豆植株图像,随机抽取训练集与验证集样本植株,并设定初始图像采集间隔与抽样步长;其次,通过植株分割、骨架提取、主茎节点去噪等操作,获取分布于植株主茎上的待检测大豆茎节点;通过基于空间距离的数据转换方法将分布离散的大豆茎节点转换至便于聚类的数据集内;利用HDBSCAN聚类算法对不同采集视角下的待检测大豆茎节点进行聚类,统计、记录主茎节数识别准确率,筛选最优采集间隔;最后,利用最优采集间隔对剩余样本植株主茎节数进行统计、分析。在63株“中黄30”待考种大豆植株中抽取21株植株作为训练集,并进行实验测试,发现在采集间隔为90°时,以最小聚类簇为2,融合处理4幅大豆图像,大豆主茎节数识别效果最优。据此对42株验证集样本植株进行主茎节数识别和分析,结果表明,大豆主茎节数识别准确率可达98.25%。该方法能够快速、准确获取大豆主茎节数,可满足大豆考种需求。

**关键词:** 大豆考种; 主茎节数; 图像处理; 空间转换; HDBSCAN 聚类算法

中图分类号: TP391.41; S565.1 文献标识码: A 文章编号: 1000-1298(2020)12-0229-09

OSID:



## Statistics of Seed-testing Soybean Main Stem Nodes Based on Image Processing and Clustering Algorithm

WANG Yueting<sup>1</sup> WANG Minjuan<sup>1</sup> SUN Shi<sup>2</sup> YANG Si<sup>1</sup> ZHENG Lihua<sup>1</sup>

(1. Key Laboratory of Modern Precision Agriculture System Integration Research, Ministry of Education, China Agricultural University, Beijing 100083, China

2. Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China)

**Abstract:** Aiming at measuring the number of main stem nodes of soybean plants quickly and efficiently, a statistical method of soybean main stem nodes was proposed based on image processing and clustering algorithm. Firstly, multiple perspectives of the soybean plants were obtained by using a camera, the initial image collection interval and sampling step were set, and then some of the plants were extracted as the train set and validation set. Secondly, through the operations such as plant segmentation, skeleton extraction, and denoising of the main stem nodes, the soybean main stem nodes to be detected were obtained. Meanwhile, the multiple dimensional scaling (MDS) was used to convert the data into a space which was easy to cluster. Then, the hierarchical density-based spatial clustering of applications with noise (HDBSCAN) clustering algorithm was used to cluster the soybean stem nodes to be detected from multiple perspectives, and the recognition accuracy of the number of main stem nodes were recorded. Finally, the optimal collection interval was used to determine the number of main stem nodes of the remaining sample plants and conduct statistical analysis. The experiments were carried out based on the above method by using 63 samples which variety was called Zhonghuang 30. 21 plants were selected as the training set, and it turned out that, under the condition of 90° interval, and four soybean images were captured and fused with the minimum cluster of 2, the node number recognition results were mostly distributed in the effective range. To identify and analyze the main stem node number of the remaining 42 sample plants, the corresponding soybean main stem node number recognition accuracy rate can reach

收稿日期: 2020-03-26 修回日期: 2020-07-01

基金项目: 国家重点研发计划项目(2016YFD0200600-2016YFD0200602)和国家现代农业产业技术体系建设专项(CARS-04)

作者简介: 王跃亭(1996—),男,博士生,主要从事农业信息化与自动化技术研究,E-mail: wyt\_study@163.com

通信作者: 郑立华(1969—),女,教授,博士生导师,主要从事农业信息化研究,E-mail: zhenglh@cau.edu.cn

98.25%。The experiment results showed that this method can meet the needs of soybean plant test requirements.

**Key words:** soybean variety test; number of main stem nodes; image processing; spatial transformation; HDBSCAN clustering algorithm

## 0 引言

为了熟悉新培育出来的大豆品种,研究人员需要对大豆植株性状进行详细的观察和记录(即考种)<sup>[1-2]</sup>。传统的人工考种存在操作误差大、效率低下、成本过高等问题。为了优化作物考种流程,研究者进行了大量研究:机器视觉、深度学习等技术应用于作物籽粒性状检测,提升了检测速率与准确性<sup>[3-5]</sup>;三维视觉技术的引入为获取作物茎秆信息,以及作物产量、抗倒伏等性状分析提供了支持<sup>[6-8]</sup>;高光谱技术与图像分析技术的结合为作物叶片性状的采集与监测提供了新的方式<sup>[9-11]</sup>。

豆荚多生长于主茎上的茎节点处,主茎节数关系到大豆的产量,故主茎节数是大豆考种过程的重要指标之一。文献检索发现,目前利用机器学习和图像处理方法进行大豆主茎节点和主茎节数统计的研究鲜有报道,但其他作物有关茎节点和茎节的研究均基于图像处理展开。文献[12-13]通过对甘蔗图像进行图像分割、局部像素点统计等对甘蔗茎节点进行识别;MOSHASHAI等<sup>[14]</sup>根据茎节处粗大的形态特点,通过检测甘蔗边缘信息,将局部最大值确定为茎节。上述方法虽然在甘蔗茎节点识别中获得不错的效果,但因大豆植株存在豆荚、分枝等因素的干扰,局部像素点较多处和边缘粗大处等不一定为大豆茎节点。葡萄、荔枝等果实采摘点与大豆豆荚生长习性类似,果实串采摘点的确定对大豆主茎节点的识别有一定借鉴意义。熊俊涛等<sup>[15-16]</sup>在识别、分割葡萄、荔枝等作物果实串的基础上,确定果梗感兴趣区域,并辅以角点检测等算法确定采摘点,为农业采摘机器人的实际应用提供了技术支持。上述方法应用于大豆茎节点识别时,需要解决大豆因色彩相近、遮挡严重等原因造成的豆荚识别困难,以及豆荚与主茎连接处较短、不易识别等问题。

综合分析相关研究可知,对于大豆植株存在的豆荚、分枝干扰问题,可利用骨架提取等技术进行图像细化,配合角点检测可进一步实现对主茎节点的确定。针对因拍摄角度造成大豆茎节点无法有效识别的问题,可利用聚类算法<sup>[17]</sup>对多角度下的图像信息进行统计和分析,以实现大豆植株茎节点的识别、统计和分析。其中,HDBSCAN (Hierarchical density-based spatial clustering of applications with

noise)<sup>[18-21]</sup>结合了密度聚类、层次聚类算法的优点,无需设定初始聚类簇大小、领域阈值等参数,能对噪声点进行有效识别和剔除,适用于未知节数情形下的大豆主茎节数统计研究。

本文在上述研究基础上,提出一种基于图像处理和聚类算法的待考种大豆主茎节数统计方法。以不同拍摄角度下的大豆图像作为输入,引入图像处理与聚类算法对大豆主茎节数进行逐级筛选、提纯,最终实现大豆主茎节数的统计。

## 1 材料与方法

### 1.1 图像采集和处理

实验对象为已脱叶、处于考种阶段的高产大豆品种“中黄30”,由中国农业科学院作物科学研究所提供。采集图像时,将种植于花盆内的待测植株放置于转盘上,背景采用黑色幕布,使用三脚架将相机架设到适宜高度,数据采集场景设计如图1a所示。拍摄时,转动转盘实现不同角度的大豆植株图像获取,图像采集实验场景如图1b所示。

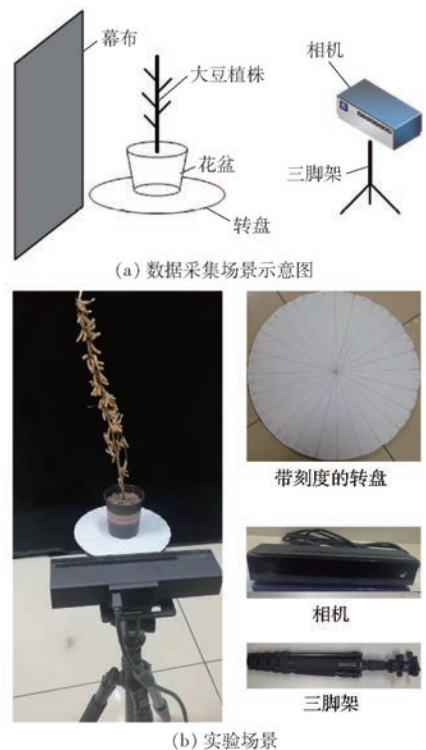


图1 数据采集实验场景

Fig. 1 Scene of data acquisition

相机分辨率为1 920像素×1 080像素,可获得彩色图像。为了避免单一拍摄角度下处于视觉盲区

的茎节点丢失,实验时设定旋转基础步长为  $15^\circ$ , 转盘旋转一周可获得 24 幅不同角度的单株大豆植株图像。本文在 Windows 10 操作系统下, 利用 PyCharm 集成开发平台, 采用 Python 语言, 实现大豆植株图像的植株分割、骨架提取、角点识别、主茎节点筛选、数据转换以及密度聚类统计等操作, 进而完成大豆植株主茎的茎节统计。

### 1.2 大豆植株主茎节数统计算法

为了有效统计大豆植株主茎节数, 本文设计了如图 2 所示的基于图像处理和聚类算法的待考种大豆植株主茎节数统计算法。算法将采集到的大豆植株图像随机分为训练集和验证集, 并进行对应的训练、验证操作。其中, 训练的目的在于获得主茎节数识别准确率最大值所对应的图像采集间隔以及聚类参数, 据此表明对同一品种大豆, 只需训练少量样本即可获得稳定、收敛的结果。验证集则可进一步验证上述算法的效果。

在训练阶段(图 2a), 首先通过人工输入初始图像采集间隔和抽样步长, 初始化算法参数。其次, 根

据上述参数抽取训练集样本, 并经过图像分割、骨架提取、角点识别、主茎节点筛选等操作, 获取待检测茎节点数据集。再次, 对分布离散、距离较难计算的二维数据点进行数据降维, 转化为更易聚类的一维空间数据。然后, 利用 HDBSCAN 算法对不同采集间隔下的待检测大豆主茎节点进行聚类分析, 完成当前大豆植株的主茎节数的识别、统计, 并计算对应的大豆植株识别准确率。最后, 依据抽样步长减小图像采集间隔, 重复上述步骤直至训练集数据运算完毕, 根据主茎节数识别准确率筛选出最优采集间隔与聚类参数。

在验证阶段(图 2b), 根据训练阶段所获得的最优采集间隔与聚类参数, 对验证集图像进行对应图像的处理、聚类操作, 并对获得的大豆识别准确率进行统计, 进一步验证算法的准确性、有效性。

#### 1.2.1 植株分割

在本文设计实验场景下采集的大豆植株图像包含花盆、转盘、幕布等其他背景信息以及干扰噪声。为了获得准确、有效的大豆植株图像, 需要对图像进

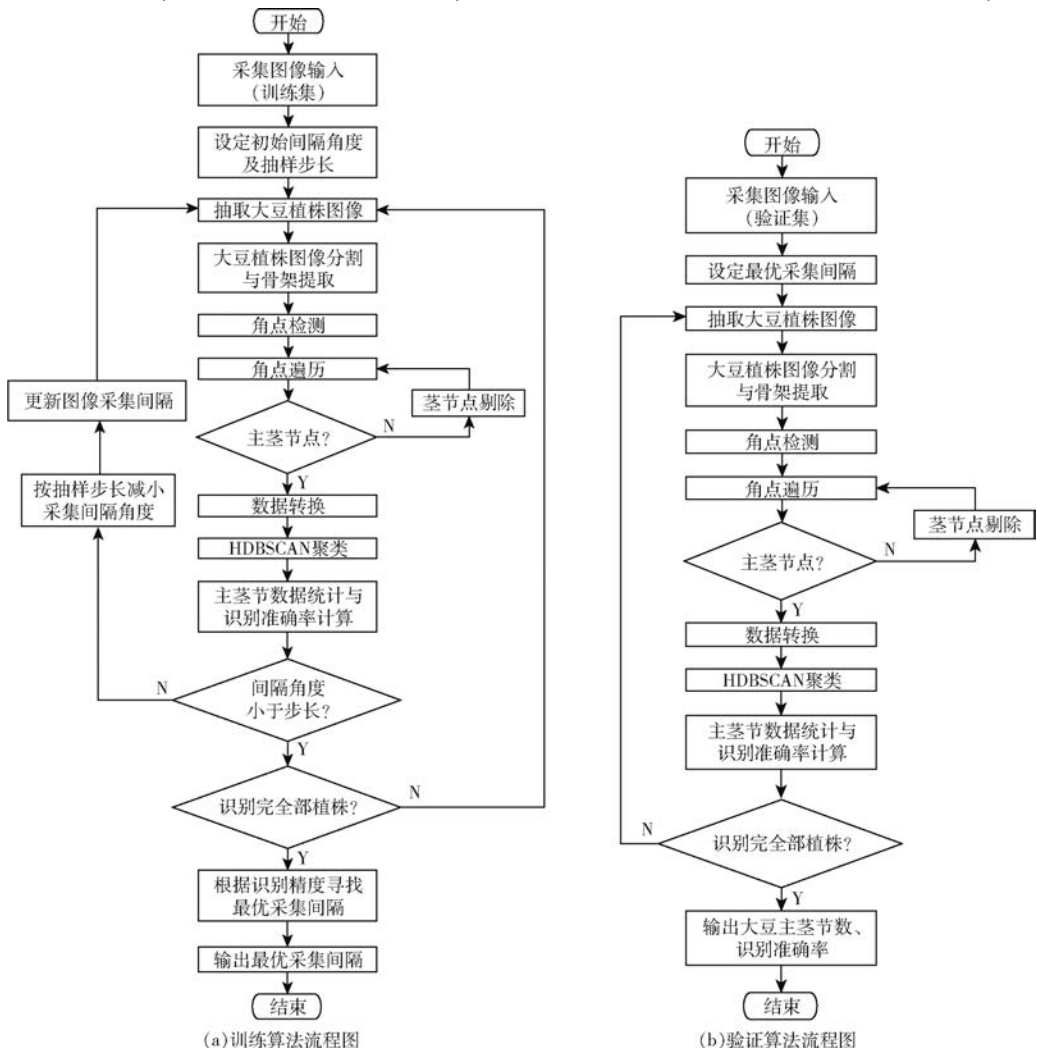


图 2 大豆植株主茎节数统计算法流程图

Fig. 2 Flow chart of statistical algorithm for number of main stem nodes of soybean plants

行分割以获得仅含有目标的图像。对此,本文利用HSV色彩模型制作掩模,并以此实现对大豆植株的图像分割,获取只含大豆植株的图像。其中,HSV色彩模型是一种面向视觉感知的颜色模型,相较于RGB色彩空间更加符合人的视觉特征。HSV色彩模型由3个要素组成,即色调( $H$ )、饱和度( $S$ )以及亮度( $V$ )。

### 1.2.2 骨架提取

植株茎节点是大豆主茎节数计数的标志点,通过识别、标识、统计植株茎节点即可确定主茎节数,进而实现大豆产量的评估。若直接进行茎节点识别,因存在豆荚遮挡、主茎分枝等干扰,将引起茎节点遗漏、重复识别、错误识别等问题(图3)。故本文利用骨架提取算法进行大豆植株图像细化,以保证主茎节点的准确、有效提取。

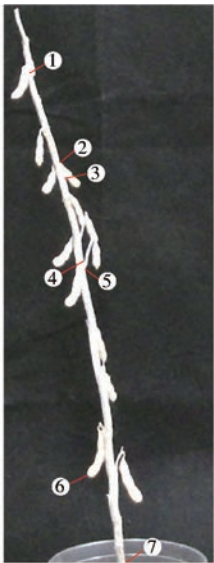


图3 大豆植株图像噪声分布示意图

Fig.3 Noise distribution schematic of soybean plant

在图3中,编号1处的茎节点存在豆荚遮挡问题;编号2、3处因豆荚与主茎的交叉而引入2个噪声点;编号4、5处则因分枝与主茎相交而引入2个噪声点;编号6处的噪声点由豆荚末端引入;编号7处则为大豆主茎与成像边界相交而引入的噪声点。

本文首先利用大津法(Otsu)<sup>[22]</sup>对图像分割后的大豆植株图像进行二值化,然后选择由LEE<sup>[23]</sup>提出的骨架提取算法进行大豆植株的骨架提取。

### 1.2.3 大豆主茎节点检测

大豆主茎节点是植株主茎上两节茎秆的连接部分,而相邻两节茎秆的直径、方向不同等使得其在该点存在属性差异(即角点)。为了获取大豆植株主茎上待检测的茎节点,本文选择Shi-Tomasi特征点检测算法<sup>[24]</sup>对大豆植株骨架进行角点检测。

经过角点检测后,骨架端点处仍存在因豆荚、分

枝以及成像限制等产生的噪声点。为了剔除噪声点,本文首先设定端点识别阈值,选定卷积模板;然后,根据设定的卷积模板尺寸构建卷积模板,对骨架提取后的大豆植株图像各像素点进行卷积操作;最后,将各卷积结果与端点识别阈值进行比较,小于阈值的点即为所寻找的端点。通过对20组不同植株在任意角度下拍摄的图像进行主茎节点提取试验,观察发现当卷积模板尺寸设置为 $3 \times 3$ 、端点阈值设置为4像素时,端点识别效果较好,能够识别出所有非遮挡的骨架节点,平均准确率达90.39%。

### 1.2.4 基于空间距离的数据转换

大豆植株生长过程中的弯曲、地面不平整等,使得不同拍摄角度的大豆主茎节点分布无序,从而使得后续的聚类算法缺乏合理的聚类依据。故需将空间分布无序的大豆茎节点进行数据转换,得出便于数据空间聚类的数据集。

本文选择MDS降维算法<sup>[25]</sup>将不同分布的待检测二维空间点映射至一维空间内,简化后续的聚类操作。MDS算法是一种基于空间距离的数据降维算法,其要求在低维空间内样本距离能够保持原始空间中的各样本之间的距离<sup>[26]</sup>。

算法假定 $m$ 个样本在原始空间(其维度为 $d$ )的距离矩阵为 $D \in \mathbf{R}^{m \times m}$ ,其第 $i$ 行、第 $j$ 列的元素 $d_{ij}$ 为样本 $x_i$ 到 $x_j$ 的距离。样本在 $d'$ 维空间的标识形式为 $Z \in \mathbf{R}^{d' \times m}$ , $d' \ll d$ ,且任意两个样本 $z_i, z_j$ 在 $d'$ 维空间中的欧氏距离等于原始空间中的距离,即 $\|z_i - z_j\| = d_{ij}$ 。

令 $B = Z^T Z \in \mathbf{R}^{m \times m}$ ,其中 $B$ 为降维后样本的内积矩阵。对矩阵 $B$ 做特征值分解, $B = VAV^T$ ,其中 $V$ 为特征向量矩阵; $A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_d)$ 为特征值构成的对角矩阵, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$ 。假定其中有 $d^*$ 个非零特征值,它们构成对角矩阵 $A^* = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_{d^*})$ ,令 $V^*$ 表示相应的特征向量矩阵,则 $Z$ 可表达为

$$Z = \sqrt{A^*} V^{*T} \quad (Z \in \mathbf{R}^{d^* \times m}) \quad (1)$$

### 1.2.5 HDBSCAN 聚类算法

HDBSCAN算法是一种密度聚类与层次聚类相结合的聚类算法。其只需确定每个簇中样本数的最小值(即最小簇大小)即可实现样本集的可视化聚类。其在未知大豆主茎节点的情况下,只需考虑待检测茎节点在总的抽样样本集中出现的频次即可,更加适合大豆主茎茎节点的聚类统计。算法具体步骤如下:

(1)根据数据点分布密度进行空间变换。为确定不同簇的聚类依据,算法引入了核心距离与互达

距离的概念。核心距离  $c_k(x) = d(x, N^k(x))$  指样本与第  $k$  个最邻近样本点的距离。 $a, b$  两点的互达距离  $d_{mreach-k}(a, b)$  计算公式为

$$d_{mreach-k}(a, b) = \max \{c_k(a), c_k(b), d(a, b)\} \quad (2)$$

式中  $d(a, b)$ —— $a, b$  两点的直线距离

对于空间中存在的点  $A, B, C$ , 当  $k=5$  时,  $A, B$  两点和  $A, C$  两点的互达距离分别为  $c_5(x), d(A, C)$ 。互达距离示意图如图 4 所示。

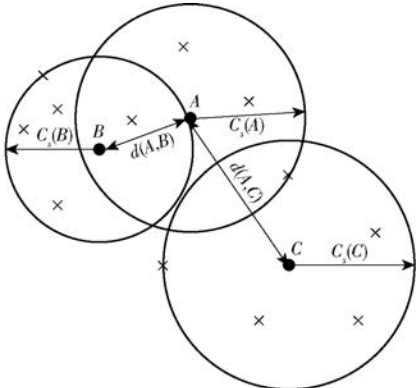


图 4 互达距离示意图

Fig. 4 Diagram of reaching distance

(2) 依据互达距离, 建立最小生成树。为了能够获取数据集内所有数据点的信息, HDBSCAN 算法利用 Prim 算法, 以数据点为顶点、任意两点间互达距离为权重, 构建最小生成树。

(3) 最小生成树转换为簇层次结构。簇状结构能够对不同层次的数据集进行划分, 故需将最小生成树转化为簇层次结构。算法以互达距离为权重, 对树的边按升序进行排序、遍历, 为每条边创建一个新的合并簇。

(4) 压缩簇层次结构。为了剔除数据集中的噪声点、压缩簇层次结构, 需要人工设定最小簇的值。在每次簇层次结构分割时, 比较新分割出的簇点数与最小簇大小: 当新簇不少于最小簇时, 新分割的簇保留, 否则该簇被剔除且保留较大的父簇; 最终获得一个拥有较少分枝的层级树结构。

(5) 提取稳定聚类簇。在完成簇层次结构压缩后, 为保证选择的聚类簇的强壮性, 需要进行稳定簇的提取。算法引入距离的倒数  $\lambda = \frac{1}{d}$  作为指标来度量聚类结果的稳定性, 并对于树中的某个节点引入 3 个量: 聚类簇形成时的  $\lambda$  值 ( $\lambda_{birth}$ ), 聚类簇分裂为 2 个子簇时的  $\lambda$  值 ( $\lambda_{death}$ ), 以及从父簇分离出去时的值 ( $\lambda_p$ )。则每个聚类簇的稳定度  $s_{cluster}$  可以表示为

$$s_{cluster} = \sum_{p \in cluster} (\lambda_p - \lambda_{death}) \quad (3)$$

式中  $c_{cluster}$ ——聚类簇集合

将所有子簇均声明为选定簇, 通过反向拓扑排序遍历比较子簇稳定度与父簇稳定度: 当子簇稳定度之和大于父簇稳定度时, 则将簇稳定度设置为子簇稳定度之和; 否则, 将父簇声明为选定簇, 并取消选定其子簇; 当到达根节点时, 当前选定簇集合即为最终提取的稳定簇集合。

### 1.2.6 识别准确率计算及寻优

为了衡量训练集大豆主茎节数的识别效果, 需要设计、制定合理的评价标准。因单个植株主茎节数较小(“中黄 30”大豆多为 12~18 节), 即使较小的误差(1~3 节)也将影响每一样本的精度统计结果。故针对不同采集间隔下的识别准确率  $P$ , 其计算公式为

$$P = \begin{cases} \left(1 - \frac{|N_I - N_T|}{N_T}\right) \times 100\% & (N_I < 2N_T) \\ 0 & (N_I \geq 2N_T) \end{cases} \quad (4)$$

式中  $N_I$ ——全部训练集识别的主茎总节数

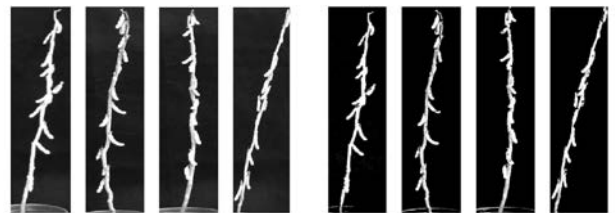
$N_T$ ——实际主茎总节数

在不同采集条件下对不同植株的识别准确率进行记录, 将最高识别准确率的采集条件作为最优采集间隔, 以供剩余植株的主茎节数识别、统计。

## 2 结果与分析

### 2.1 大豆植株图像分割

为了方便大豆植株的分割, 在植株后方布置灰色幕布, 因此 HSV 色彩模型选定灰色色彩空间 ( $H_{\min} = 0, H_{\max} = 180; S_{\min} = 0, S_{\max} = 30; V_{\min} = 46, V_{\max} = 220$ ) 进行背景色彩描述; 然后, 通过取反操作制作掩模与拍摄图像进行与操作完成大豆植株的图像分割。图 5 为 4 株已脱叶的待考种“中黄 30”大豆在转角  $0^\circ$  时的原始图像及目标分割后的图像。



(a) 原始图像

(b) 目标分割后图像

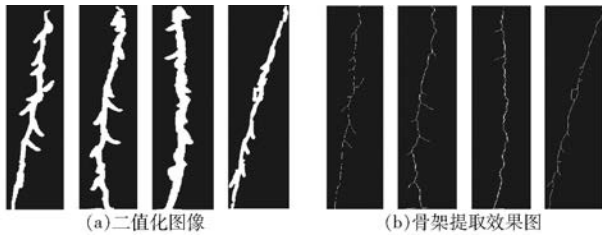
图 5 大豆植株原始图像及分割后图像

Fig. 5 Images of original and image segmentation

### 2.2 大豆植株骨架提取与角点检测

#### 2.2.1 大豆植株骨架提取

在获取到只含有大豆植株的图像后, 首先利用天津法对图像进行二值化处理, 将大豆植株所在区域灰度设置为 255 (即白色), 背景区域灰度设置为 0 (即黑色); 然后, 利用骨架提取算法进行二值图像的骨架提取, 对图 5 进行处理的结果如图 6 所示。



(a)二值化图像 (b)骨架提取效果图

图6 大豆植株二值化图像与骨架提取效果

Fig. 6 Images of binarization and skeleton extraction

通过骨架提取将大豆植株细化为宽度为1像素的二值化图像,减少同一豆荚、分枝因其空间不同而引入的噪声点数量。此外,在宽度为1像素的植株图像中更容易确定茎节点的位置。

### 2.2.2 大豆植株角点检测

为保证获得所有可能茎节,针对考种阶段的大豆植株,通过前期试验设定最大识别角点数为100、识别等级为0.01、最小间隔为30像素进行大豆植株骨架的角点检测,图7为图6的处理结果,并将对应的图像坐标信息存入Excel表格内备用。

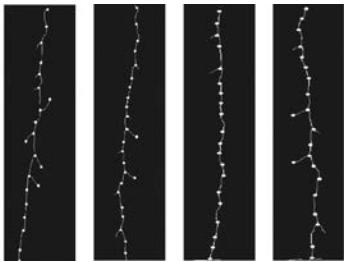


图7 不同视角图像的待检测茎节点识别结果

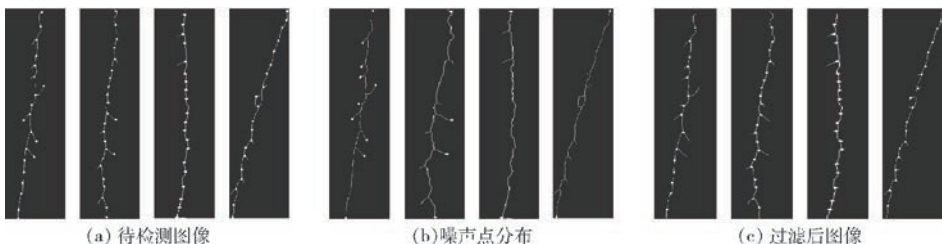
Fig. 7 Recognition results of nodes from different perspectives

由图7可知,通过角点检测获得的待检测茎节点中存在由豆荚骨架末端以及茎秆与图像边界相交所引入的噪声点,需要进行大豆植株主茎节点的筛选和提纯。

### 2.3 大豆植株主茎节点筛选

为了将骨架端点处的噪声点剔除,本文对骨架图像端点进行了识别、记录,并将其从待检测茎节点数据集中剔除,图7中各目标的处理结果如图8所示。

观察发现,通过图像的卷积操作能够保留植株主茎上的茎节点,并滤除植株、分枝、豆荚等部位的



(a)待检测图像

(b)噪声点分布

(c)过滤后图像

图8 大豆植株骨架端点识别及筛选效果

Fig. 8 Recognition and filtering of skeleton endpoint and screening

末端噪声点干扰。因已经滤除主茎末端噪声干扰,现存的主茎节点个数即可应用于后续聚类分析。

### 2.4 基于空间距离的数据转换

完成单幅图像的主茎节点筛选后,需要将不同视角下的主茎节点进行汇总,并将数据转换为更适宜聚类的一维空间数据。因不同角度下采集的图像尺寸(1920像素×1080像素)相同,且每幅图像具有相同的像素坐标系,故不同视角下的主茎节点可放于同一文件内直接进行比较。

图9为从同一株大豆植株的3组图像中提取的主茎节点组合以及经过MDS数据转换的茎节点空间分布效果图。这3组图像分别为拍摄角度相隔180°的2幅图像(图9a)、相隔120°的3幅图像(图9b)、相隔90°的4幅图像(图9c)。其中,左图为茎节点图像组合,右图为降维转换后的数据(为了便于数据点描述,X坐标统一置零,Y坐标为转换后的数值)。

通过对数据转换后的图像进行分析可知,当采集间隔较大、输入图像较少时,转换后的各待检测茎节点分布较为疏散,但数据量较少;随着采集间隔减小、输入图像增多,数据点分布逐渐稠密,在多幅图像中出现的茎节点重叠使得图像中该点颜色加深。

### 2.5 茎节点聚类

在HDBSCAN算法聚类过程中将最小簇设置为2(在2幅图像中均稳定出现的茎节点即认定为主茎节点),部分聚类效果如图10所示,其中被圈出的是稳定的聚类簇,即被认定为大豆植株主茎上的茎节点。

由图10可知,当图像采集间隔较大时,输入的待检测茎节点较少,样本量不足以及部分样本点相距较远,导致识别的茎节点数少于真实值。随着图像采集间隔的不断减小,输入的待检测茎节点总数不断增加,更多的待检测茎节点被判定为茎节点,但也更容易受到噪声点的影响。

### 2.6 最优采集间隔选取与识别效果分析

为了确定最优的图像采集间隔,本文选择21株待考种的“中黄30”大豆植株作为训练集,以180°采集间隔为初始采集间隔,设定抽样步长为30°,分别

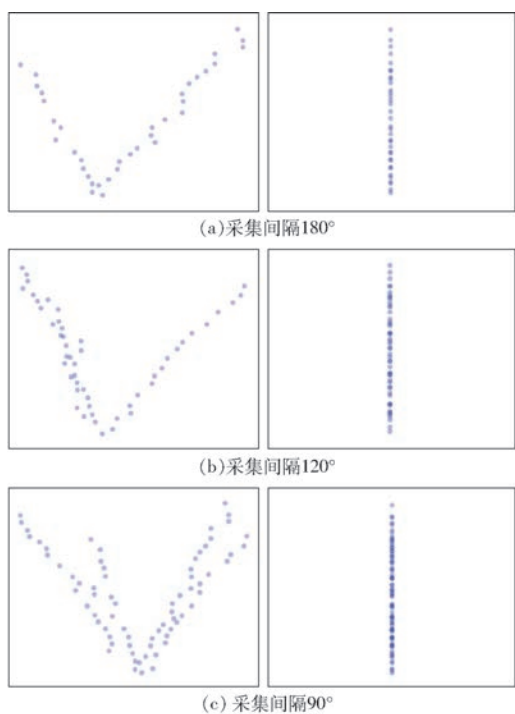


图 9 大豆植株主茎节点数据降维效果

Fig.9 Effect diagrams of dimensionality reduction of node data

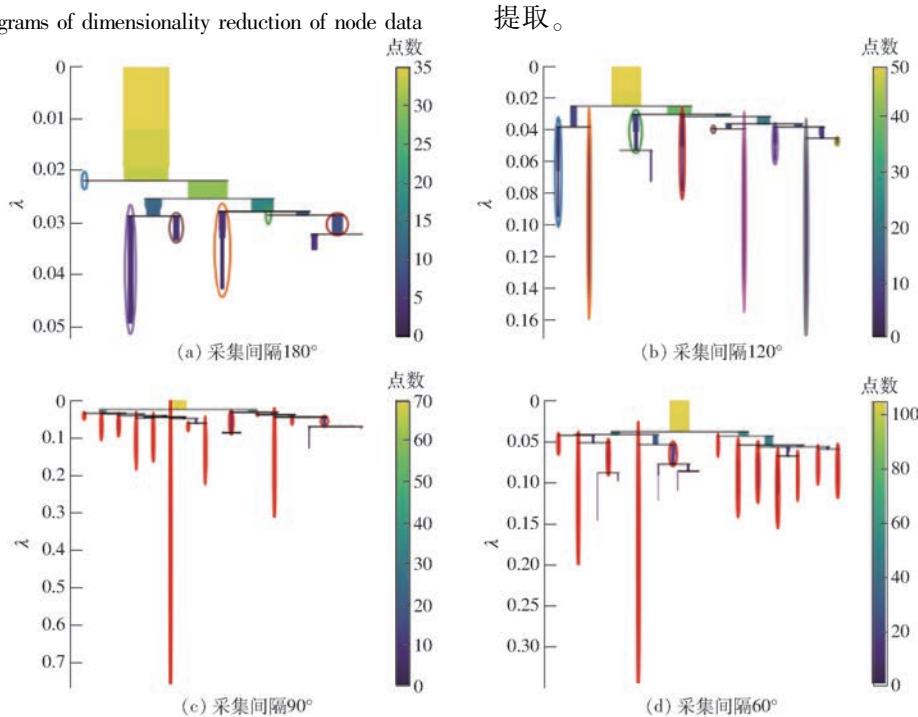


图 10 大豆植株主茎节点聚类效果

Fig. 10 Clustering of main stem nodes of soybean plants

对采集间隔为 180°、150°、120°、90°、60°、30°时获得的主茎节数识别准确率进行记录、统计和对比分析,结果如表 1 所示。

由表 1 可知,当图像采集间隔为 90°时,其识别准确率为 91.11%,识别效果最优,因此将其作为剩余样本的最优图像采集间隔。本文对 7 种不同主茎节数的 42 株验证集大豆植株(即每种主茎节数包括 6 株样本)进行主茎节数识别、统计,并计算其识别准确率,结果如表 2 所示。

由表 2 可知,在图像采集间隔为 90°的条件下,不同主茎节数的大豆植株的平均识别准确率可达 98.25%,单株考种大豆主茎节数识别准确率大于 88%,能够满足大豆实际考种需求。

### 3 结论

(1)根据大豆植株的生长期及其骨架角点分布情况,通过实验设定卷积模板尺寸,并进行卷积计算,实现大豆植株主茎骨架上的待检测数据点提取。

表 1 训练集不同采集间隔下的主茎节数识别结果

Tab.1 Results of soybean main stem nodes from different collection intervals

采集间隔/ (°)	实际主茎节 总数/节	识别主茎节 总数/节	识别准确率/ %
180	315	99	31.43
150	315	251	79.68
120	315	237	75.23
90	315	343	91.11
60	315	496	42.54
30	315	870	0

表 2 采集间隔为 90°下的主茎节数识别结果

Tab.2 Results of soybean main stem nodes at 90°

实际 节数/节	平均识别 茎节数/节	实际主茎节 总数/节	识别主茎节 总数/节	识别准确 率/%
12	13.17	72	79	90.25
13	14.50	78	87	88.46
14	13.83	84	83	98.79
15	14.33	90	86	95.53
16	16.50	96	99	96.86
17	17.17	102	103	99.00
18	17.33	108	104	96.28

(2) 基于空间距离的数据转换算法解决了因大豆植株自然生长、地势起伏等原因而引起数据点分布离散的问题,为聚类算法提供较好的样本数据集。

(3) 在未知茎节点个数的情况下,本文方法能

够统计出大豆植株主茎节数,并可剔除分枝中茎节点的干扰。实验发现,对于42株待考种“中黄30”大豆植株,利用90°采集间隔获得的图像的主茎节数平均识别准确率达98.25%。

### 参 考 文 献

- [1] 赵朝森,赵现伟,杨中路,等. 国家区试长江流域春大豆品种农艺、产量及品质性状的演变[J]. 大豆科学,2019,38(1): 41-48.  
ZHAO Chaosen, ZHAO Xianwei, YANG Zhonglu, et al. Evolution of agronomic, yield and quality traits of spring soybean varieties attending national regional test of Yangtze reiver basin [J]. Soybean Science, 2019, 38(1):41-48. (in Chinese)
- [2] 宋豫红,白艳凤,包荣军,等. 东北大豆种质群体生态性状在北安地区的表现及其潜在的育种意义[J]. 大豆科学,2018, 37(6):829-838.  
SONG Yuhong, BAI Yanfeng, BAO Rongjun, et al. Performance and breeding potential of northeast soybean germplasm population in Beian [J]. Soybean Science, 2018, 37(6):829-838. (in Chinese)
- [3] 赵志衡,宋欢,朱江波,等. 基于卷积神经网络的花生籽粒完整性识别算法及应用[J]. 农业工程学报,2018,34(21): 195-201.  
ZHAO Zhiheng, SONG Huan, ZHU Jiangbo, et al. Identification algorithm and application of peanut kernel integrity based on convolution neural network[J]. Transactions of the CSAE, 2018, 34(21): 195-201. (in Chinese)
- [4] 陈进,顾琰,练毅,等. 基于机器视觉的水稻杂质及破碎籽粒在线识别方法[J]. 农业工程学报,2018,34(13):187-194.  
CHEN Jin, GU Yan, LIAN Yi, et al. Online recognition method of impurities and broken paddy grains based on machine vision [J]. Transactions of the CSAE, 2018, 34(13): 187-194. (in Chinese)
- [5] 杜建军,郭新宇,王传宇,等. 基于全景图像的玉米果穗流水线考种方法及系统[J]. 农业工程学报,2018,34(13):195-202.  
DU Jianjun, GUO Xinyu, WANG Chuanyu, et al. Assembly line variety test method and system for corn ears based on panoramic surface image[J]. Transactions of the CSAE, 2018, 34(13): 195-202. (in Chinese)
- [6] 冯佳睿,马晓丹,关海鸥,等. 基于深度信息的大豆株高计算方法[J]. 光学学报,2019,39(5):258-268.  
FENG Jiarui, MA Xiaodan, GUAN Haiou, et al. Calculation method of soybean plant height based on depth information[J]. Acta Optica Sinica, 2019, 39(5): 258-268. (in Chinese)
- [7] BUSEMEYER L, MENTRUP D, MOLLER K, et al. Breed vision a multi sensor platform for non-destructive field-based phenotyping in plant breeding[J]. Sensors (Basel), 2013, 13(3): 2830-2847.
- [8] PAULUS S, SCHUMANN H, KUHLMANN H, et al. High-precision laser scanning system for capturing 3D plant architecture and analysing growth of cereal plants[J]. Biosystems Engineering, 2014(121): 1-11.
- [9] 张宏鸣,谭紫薇,韩文霆,等. 基于无人机遥感的玉米株高提取方法[J/OL]. 农业机械学报,2019,50(5):241-250.  
ZHANG Hongming, TAN Ziwei, HAN Wenting, et al. Extraction method of maize height based on UAV remote sensing[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(5): 241-250. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20190528&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20190528&journal_id=jcsam). DOI:10.6041/j.issn.1000-1298.2019.05.028. (in Chinese)
- [10] 杨琦,叶豪,黄凯,等. 利用无人机影像构建作物表面模型估测甘蔗 LAI[J]. 农业工程学报,2017,33(8):104-111.  
YANG Qi, YE Hao, HUANG Kai, et al. Estimation of leaf area index of sugarcane using crop surface model based on UAV image[J]. Transactions of the CSAE, 2017, 33(8): 104-111. (in Chinese)
- [11] 冯青春,陈建,李翠玲,等. 基于光度立体视觉的蔬菜秧苗叶片形态测量方法[J/OL]. 农业机械学报,2018,49(5):43-50.  
FENG Qingchun, CHEN Jian, LI Cuiling, et al. Measurement method of vegetable seedling leaf morphology based on photometric stereo[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(5): 43-50. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20180505&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20180505&journal_id=jcsam). DOI:10.6041/j.issn.1000-1298.2018.05.005. (in Chinese)
- [12] 黄亦其,黄体森,黄媚章,等. 基于局部均值的甘蔗茎节识别[J]. 中国农机化学报,2017,38(2):76-80.  
HUANG Yiqi, HUANG Tisen, HUANG Meizhang, et al. Recognition of sugarcane nodes based on local mean [J]. Journal of Chinese Agricultural Mechanization, 2017,38(2):76-80. (in Chinese)
- [13] 张卫正,董寿银,齐晓祥,等. 基于图像处理甘蔗茎节识别与定位[J]. 农机化研究,2016,38(4):217-221.  
ZHANG Weizheng, DONG Shouyin, QI Xiaoxiang, et al. The identification and location of sugarcane internode based on image processing [J]. Journal of Agricultural Mechanization Research, 2016, 38(4):217-221. (in Chinese)
- [14] MOSHASHAI K, ALMASI M, MINAEI S, et al. Identification of sugarcane nodes using image processing and machine vision technology[J]. International Journal of Agricultural Research, 2008, 3(5):357-364.
- [15] 熊俊涛,何志良,汤林越,等. 非结构环境中扰动葡萄采摘点的视觉定位技术[J/OL]. 农业机械学报,2017,48(4):29-33,81.  
XIONG Juntao, HE Zhiliang, TANG Linyue, et al. Visual localization of disturbed grape picking point in non-structural environment[J/OL]. Transactions of the Chinese Society for Agricultural Machinery,2017,48(4):29-33,81. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20170403&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20170403&journal_id=jcsam). DOI:10.6041/j.issn.1000-1298.2017.04.003. (in Chinese)
- [16] 郭艾侠,熊俊涛,肖德琴,等. 融合 Harris 与 SIFT 算法的荔枝采摘点计算与立体匹配[J/OL]. 农业机械学报,2015,



46(12):11-17.

GUO Aixia, XIONG Juntao, XIAO Deqin, et al. Computation of picking point of litchi and its binocular stereo matching based on combined algorithms of Harris and SIFT[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(12):11-17. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20151202&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20151202&journal_id=jcsam). DOI:10.6041/j.issn.1000-1298.2015.12.002. (in Chinese)

- [17] 吴焕丽, 崔可旺, 张馨, 等. 基于改进 K-means 图像分割算法的细叶作物覆盖度提取[J/OL]. 农业机械学报, 2019, 50(1):42-50.  
WU Huanli, CUI Kewang, ZHANG Xin, et al. Improving accuracy of fine leaf crop coverage by improved K-means algorithm [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(1):42-50. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20190104&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20190104&journal_id=jcsam). DOI:10.6041/j.issn.1000-1298.2019.01.004. (in Chinese)
- [18] CAMPELLO R J, MOULAVI D, SANDER J. Density-based clustering based on hierarchical density estimates [C]//Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer Berlin Heidelberg, 2013:160-172.
- [19] CHAUDHURI K, DASGUPTA S. Rates of convergence for the cluster tree [C]//Advances in Neural Information Processing Systems, 2010:839-847.
- [20] MOULAVI D, JASKOWIAK P A, CAMPELLO R J, et al. Density-based clustering validation[C]//In SDM, 2014:839-847.
- [21] MALZER C, BAUM M. HDBSCAN( $\epsilon^*$ ): an alternative cluster extraction method for HDBSCAN [J]. arXiv preprint: 1911.02282.
- [22] OSTU N. A threshold selection method from gray-histogram[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1979, 9(1):62-66.
- [23] LEE T C. Building skeleton models via 3-D medial surface/axis thinning algorithms[J]. CVGIP: Graphical Models and Image Processing, 1994, 56(6):462-478.
- [24] SHI J, TOMASI C. Good features to track[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1994:593-600.
- [25] COX T F, COX M A A. Multidimensional scaling[J]. Journal of the Royal Statistical Society, 2001, 46(2):1050-1057.
- [26] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2017.

(上接第 183 页)

- [30] 关小克, 王秀丽, 张凤荣, 等. 生态刚性约束下的山区农村居民点整治与调控——以北京市门头沟区为例[J]. 资源科学, 2017, 39(2):220-230.  
GUAN Xiaoke, WANG Xiuli, ZHANG Fengrong, et al. Rehabilitation and adjustment of rural residences in mountainous areas under rigid ecological constraints[J]. Resources Science, 2017, 39(2):220-230. (in Chinese)
- [31] 关小克, 张凤荣, 王秀丽, 等. 北京市生态用地空间演变与布局优化研究[J]. 地域研究与开发, 2013, 32(3):119-124.  
GUAN Xiaoke, ZHANG Fengrong, WANG Xiuli, et al. Spatial evolution of urban ecological land and its distribution optimization in Beijing[J]. Areal Research and Development, 2013, 32(3):119-124. (in Chinese)
- [32] 傅伯杰. 土地资源系统认知与国土生态安全格局[J]. 中国土地, 2019(12):9-11.  
FU Bojie. Land resources system cognition and land ecological security pattern[J]. China Land, 2019(12):9-11. (in Chinese)
- [33] 关小克, 王秀丽, 陈伟强. 县域耕地边际化风险评价与验证[J/OL]. 农业机械学报, 2020, 51(2):153-160.  
GUAN Xiaoke, WANG Xiuli, CHEN Weiqiang. Risk evaluation and verification of marginalization of farmland in county area [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(2):153-160. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?file\\_no=2002017&flag=1](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=2002017&flag=1). DOI:10.6041/j.issn.1000-1298.2020.02.017. (in Chinese)
- [34] 任艳敏, 唐秀美, 刘玉, 等. 基于“生产-生态-区位”的都市区粮田布局方法[J/OL]. 农业机械学报, 2016, 47(5):321-327.  
REN Yanmin, TANG Xiumei, LIU Yu, et al. Layout method of cereal field based on production - ecology - location in metropolitan area [J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(5):321-327. [http://www.j-csam.org/jcsam/ch/reader/view\\_abstract.aspx?flag=1&file\\_no=20160544&journal\\_id=jcsam](http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20160544&journal_id=jcsam). DOI:10.6041/j.issn.1000-1298.2016.05.044. (in Chinese)
- [35] 赵爱栋, 许实, 曾薇, 等. 不稳定耕地利用困境: 基于粮食安全、农民收入和生态安全间的权衡——以甘肃省景泰县为例[J]. 资源科学, 2016, 38(10):1883-1892.  
ZHAO Aidong, XU Shi, ZENG Wei, et al. Utilization dilemma of unstable farmland: the tradeoffs between food safety, famers' income and ecological safety[J]. Resources Science, 2016, 38(10):1883-1892. (in Chinese)
- [36] 肖春艳. 小浪底大坝下游滨河湿地退化特征及修复途径研究[D]. 焦作: 河南理工大学, 2015.  
XIAO Chunyan. Study on degradation features and restoration approaches of downstream riparian wetland from the Xiaolangdi Dam of the Yellow River[D]. Jiaozuo: Henan Polytechnic University, 2015. (in Chinese)