

基于自适应三分图的玉米穗丝精细化提取^{*}

熊俊逸^{1,2}, 朱德利^{1,2}, 文 瑞^{1,2}

(1. 重庆师范大学 计算机与信息科学学院; 2. 重庆市数字农业服务工程技术研究中心, 重庆 401331)

摘要:玉米穗丝是玉米的受粉器官,它的生长发育状况将直接影响玉米产量。为了实现穗丝性状的自动监测与评价,提出一种基于自适应三分图的玉米穗丝精细化提取方法。首先基于直方图对比度的显著性检测算法预估目标显著性,形成显著图,再利用自适应阈值分割对显著图初步预分割;然后寻找最大连通域获得单一目标二值图像,根据穗丝特点确定形态学处理的最佳卷积核并生成包含目标、背景和不确定区域的自适应三分图,以 Laplace 矩阵优化 closed-form matting 算法中目标函数的分块矩阵;最后根据玉米穗丝特点,结合三分图结果和优化过的算法对玉米穗丝进行软分割实现精细化提取。实验结果表明:所提的研究方法像素精度值为 97.96%,综合评价指标值为 94.16%;其中像素精度值相较于基于点运算的 OTSU 算法、Grabcut 算法和基于神经网络的深度学习算法 SeFormer 和 DeeplabV3,分别提高了 10.43、2.35、1.67 和 1.91 百分点,综合评价指标值分别提高了 23.96、5.59、4.28 和 5.13 百分点。此方法能有效弥补硬分割算法对玉米穗丝这种特殊对象的分割缺陷,提取的目标区域接近真实区域,准确度高,从而为智能化监测玉米穗丝生长状况提供技术支持。

关键词:玉米穗丝;显著性检测;图像分割;目标提取

中图分类号:TP391.41

文献标志码:A

文章编号:1672-6693(2024)01-0100-09

玉米是中国主要农作物之一,保证玉米稳产高产是实现粮食安全的关键任务。玉米穗丝是玉米的受粉器官,它的生长发育状况与产量密切相关。现阶段对玉米穗须生长状况的监控,主要通过人工巡检的方式,此方法成本高,效率低,远远无法满足实际生产的需要。目前,机械化农业已向智慧农业转变,人工检测的方式则逐渐被信息化方式取代。在智能化监测玉米穗丝生长状况的工作中,准确提取穗须是首要条件。

提取图像中的农作物,需要将目标与叶片、土壤等背景进行区别,目前对目标进行提取的主流方法有目标检测和分割。张万红等人^[1]用基于 H 和 S 颜色分量的 K 均值聚类结合色差运算的分割算法能够在弱光条件对玉米覆膜影像进行分割;王春雷等人^[2]提出一种基于遗传算法和阈值滤波的玉米根茬行图像分割方法,能在复杂背景下和强光照下,准确分割出玉米根茬行;许高建等人^[3]用基于 k-means 算法在 RGB、HSV 和 Lab 颜色空间中对小麦扬花到黄熟期感染赤霉病的麦穗图像进行分割,提取出麦穗的病害部分,能够较为精确地识别患病麦穗;车金庆等人^[4]采用基于频率调谐的显著性检测算法分别提取光线以正常区域为主和以高亮区域为主的显著图,再阈值分割,将 2 个二值图像合并,由此获得黄色和绿色苹果的果实区域。但是,上述基于像素点运算的分割算法受光照和背景影响较大,算法鲁棒性较差。因此,Boykov 等人^[5]提出了一种用于 n 维图像通用的交互式分割技术 GraphCut;Rother 等人^[6]在 Boykov 的基础上提出 GrabCut 将基于灰度分布的模型替换为高斯混合模型以支持彩色图片分割。但 GraphCut 和 GrabCut 需要人工标记前景和背景,耗时费力。而深度学习为图像检测和分割带来了新的研究方向,许景辉等人^[7]在 VGG-16 模型的基础上,提出了一种基于迁移学习的卷积神经网络玉米病害图像识别模型,对玉米健康叶、大斑病叶和锈病叶图像的平均识别准确率为 95.33%;He 等人^[8]扩展了 Faster R-CNN^[9],通过添加一个用于预测对象掩码的分支与用于边界框识别的现有分支并行,有效地检测图像中的对象,同时为每个实例生成高质量的分割掩码;王璨等人^[10]提出改进的双注意力语义分割方法,能够对复杂田间场景中幼苗期玉米和杂草进行识别与区域划分。

^{*} 收稿日期:2022-10-18 修回日期:2023-11-10 网络出版时间:2023-07-04T12:30

资助项目:重庆市教育委员会科学技术研究计划项目(No. KJQN201800536);重庆市高校创新研究群体项目(No. CXQT20015);重庆师范大学研究生创新项目(No. YKC21038)

第一作者简介:熊俊逸,女,研究方向为数字图像处理,E-mail:672144937@qq.com;通信作者:朱德利,男,教授,博士,E-mail:delizhu@qq.com

网络出版地址:https://link.cnki.net/urlid/50.1165.N.20230704.1010.002

上述例如果实、叶片等需要分割或检测的目标在影像中都呈块状且都有清晰光滑的边界,然而玉米穗丝却不同于这些目标。玉米穗丝须具有特殊性:一方面穗丝错综复杂,边界杂乱不清晰;另一方面穗丝不同生长阶段有不同形态和颜色。这些情况都加大了精细化提取穗丝的难度,使得上述方法对穗丝提取不够精细。若采用硬分割算法提取,穗丝目标和背景交界处提取的效果较差;若用目标检测的方式,得到的穗丝则会包含过多背景信息。但引入 matting^[11-17] 技术实现对穗丝的软分割,可以解决上述问题,使提取到的穗丝边缘清晰准确,更接近真实区域。现有 matting 算法多数需要用户手动标注提供三分图(trimap)或是草图(scribble),以此提高算法的准确度,但手动标注又存在花费高、效率低等问题。为解决上述问题,本文提出一种基于自适应三分图方法对玉米穗丝进行精细化提取。

1 材料与方法

1.1 研究数据与预处理

玉米穗丝图像于 2021 年 6 月中旬到 7 月初采集于重庆市云阳县农田,分别采集了玉米穗丝不同生长时期图像,为避免采集的玉米穗丝图像尺寸大小不一、重复率较高等问题,去除冗余图像样本以及穗丝特征不清等杂糅图,分辨率统一调整为 480×640 像素,得到 400 张有效样本。

对比实验采用了 4 种分割算法,包括基于点运算的经典方法 OTSU^[18]、Grabcut^[19] 和基于神经网络的语义分割算法 SegFormer^[20] 和 DeeplabV3^[21]。其中 SegFormer 和 DeeplabV3 的数据集是从田间采集的玉米穗丝图像中随机选取 220 张,使用图像标记工具 Labelme 手工标记玉米穗丝边界,尽可能准确清楚,得到手工标记图,由于数据量较少,又采用了包括旋转、随机加噪声、裁剪和翻转等方式进行数据增强。共得到 1 100 张图片,其中 990 张作为训练集,110 张作为验证集。

1.2 整体研究思路

为精细化提取玉米穗丝,首先基于图像直方图对比度计算每个像素点的显著性值,得到显著图,再使用自适应阈值分割算法对图像进行初步预分割,然后通过寻找最大连通域去除图像中的噪声和孤岛现象,根据穗丝特点确定形态学处理的最佳卷积核并生成包含目标、背景和不确定区域的自适应三分图,再优化 closed-form matting 算法的目标函数,简化计算,最后结合三分图结果和优化过的算法实现穗丝精细化提取。实验整体流程如图 1 所示。

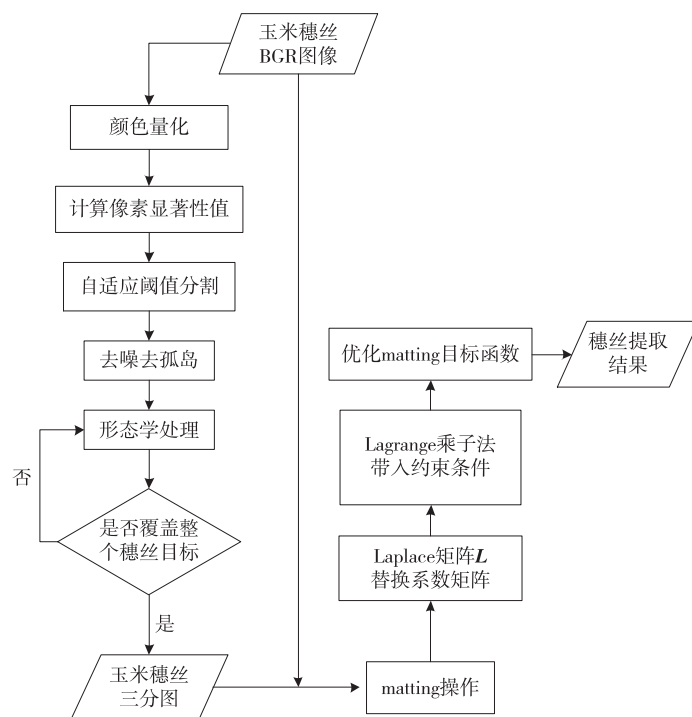


图 1 实验流程

Fig. 1 Experiment process

1.3 自适应生成三分图

1.3.1 显著性目标检测

通过对采集的穗丝图像进行观察分析,可知穗丝与背景玉米枝叶、玉米果实的颜色和性状有较为明显的区别,人的视觉注意力会更多的停留在穗丝所在的显著区域。Itti 等人^[22]将视觉显著性从认识心理学和神经学迁移到计算机视觉领域中,使显著性成为一种可以被计算的内容^[23]。显著性检测模拟人类视觉注意机制去对图像信息进行处理,目前,典型的显著性检测算法有 AC、LC、FT 和 HC。

由于玉米穗丝和其他背景在颜色上有比较明显的差别,因此显著性检测算法可以从像素颜色的对比入手。LC 和 HC^[24]算法都是基于 1 个像素和全局所有像素颜色的对比度来确定像素的显著性值,LC 算法只考虑到了图像的灰度信息,不适用于本文中 BGR 穗丝图像,而 HC 算法是包含彩色信息的基于直方图对比度的显著性检测算法,通过计算像素与图像中其他像素在 Lab 空间中的颜色距离为度量,来生成显著图。一张图片中每个像素点的显著性值是由它和剩下全部图像像素点的对比度来定义的,即 1 个像素的显著性值为 $S(I_k) = S(c_l) = \sum_{j=1}^n f_j D(c_l, c_j)$, 其中: c_l 是像素 I_k 的颜色值, n 是不同像素颜色的数量, f_j 是图像 I 中像素颜色 c_j 的概率。

通过实验验证 HC 算法对穗丝图像信息的感知灵敏度较高,能得到较为完整的显著区域。使后续操作会优先关注显著性区域,大大提高对图像信息进行处理效率。4 种显著性检测算法的实验结果如图 2 所示。

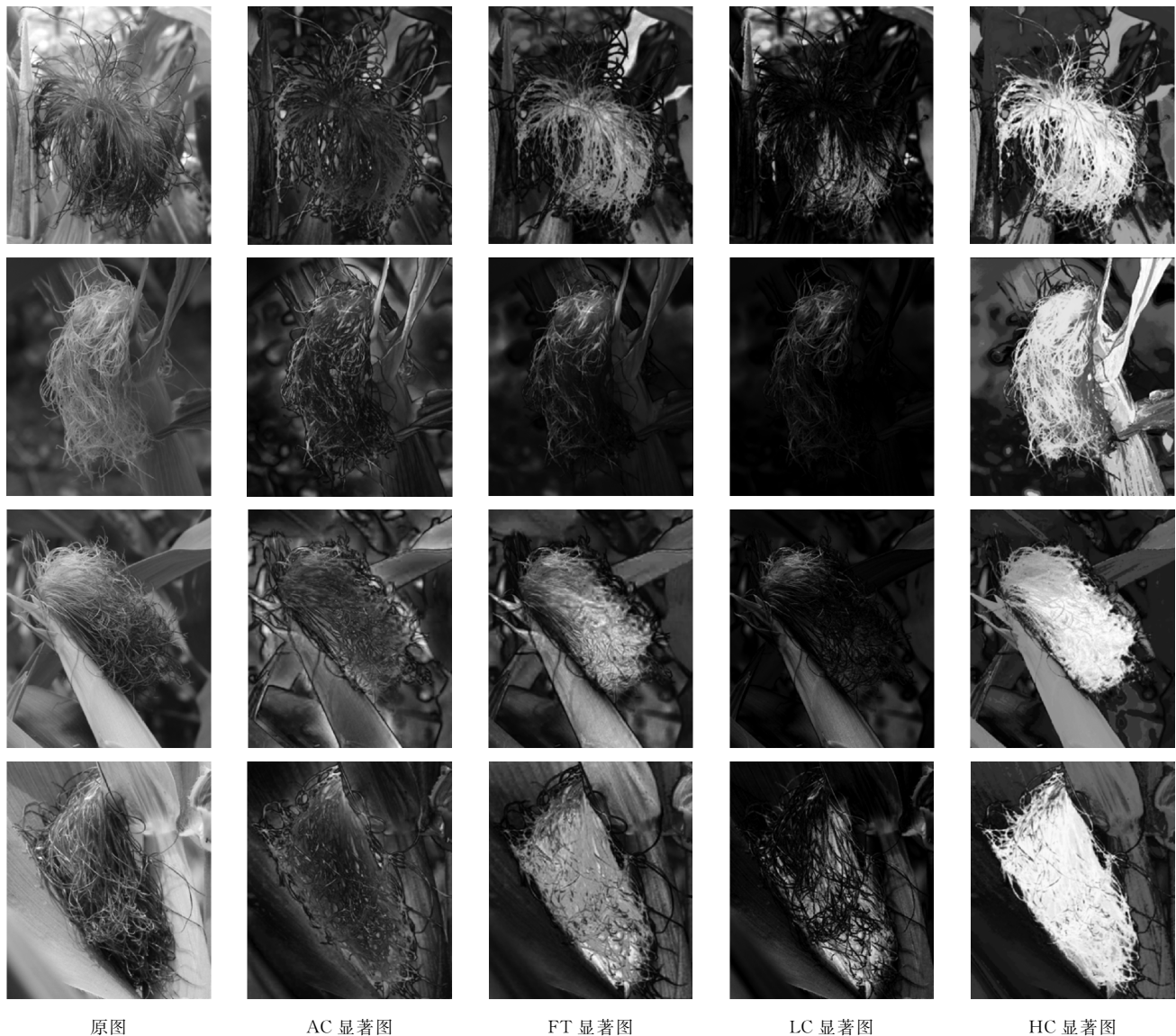


图 2 4 种显著性检测算法结果对比

Fig. 2 The results of the saliency detection algorithm

显著性检测算法给予每个像素 1 个显著性值,将穗丝与背景进行区别,得到显著图。显著图中玉米穗丝区域显著性较强,周围叶片和土壤显著性相对较弱,并且得到的显著区域较为完整,边缘也较为清晰。

1.3.2 预分割

得到图像的显著图后,首先采用最大类间方差法^[18]对显著图进行预分割,将它变为二值图,如图 2FT 显著图所示。完成分割后,发现此时二值图目标区域不只是穗丝,还残存着部分背景,造成这种现象通常是因为玉米叶片由于光照有较亮的反光部分,天空背景和土壤背景对图像显著性也造成了一定影响,导致 OTSU 分割存在大量非目标区域。

观察分割图像发现,其中穗丝区域占目标区域的很大一部分,是前景中最大的区域,且穗丝部分是一片连续区域,穗丝周围分布着零散的孤岛部分。针对这样的情况,本文通过寻找分割图像中最大连通域,将一些小面积区域作为背景像素去除,通过保留分割后目标中面积最大的区域来得到所需要的唯一穗丝目标,算法流程如下: 1) 将图像进行腐蚀,目的为了使一些与穗丝相连但并不是穗丝的区域断开,不再连通。2) 将所有连通区域标记为不同颜色。3) 统计每种颜色的数量,找到最大值。4) 遍历所有像素点,像素点属于最大连通区域则赋值 255 为目标穗丝,其余的像素全部赋值为 0 使之成为背景。最终得到唯一目标二值图如图 2LC 显著图所示。

1.3.3 形态学处理

图像中穗丝区域由于光照、重叠和遮挡,会产生孔洞,这些区域会被误认为背景,在后续腐蚀操作时,会将这些孔洞放大,从而影响三分图的准确度。因此需要先将穗丝区域的孔洞填充,在此使用膨胀操作,采用椭圆形核,核大小为 5,将这些孔洞进行填充。

腐蚀计算公式为: $X \ominus S = \{x | S + x \subseteq X\}$; 膨胀计算公式为: $X \oplus S = \{x | S + x \cup x \neq \emptyset\}$ 。其中: X 是目标图像, S 是结构元素。

准确的三分图白色区域和黑色区域一定要是确定的前景和背景,灰色部分是不确定区域,这是作为先验条件参与 matting 算法。由存在唯一目标二值图得到包含前景(白色区域)、背景(黑色区域)和不确定区域(灰色区域)的三分图,需要对唯一目标二值图进行先腐蚀再膨胀的操作。具体操作如下: 1) 先将目标腐蚀。为缩小穗丝前景(白色区域)所覆盖的面积,以防预分割后得到的唯一目标区域还是包含部分属于背景的像素所带来的误差。此时的核设置应当偏小,实验选取 45 作为核的大小,因为经过上述预分割和孔洞填充操作,此时的目标穗丝区域已经较为接近准确区域。2) 再将腐蚀后的目标膨胀。为让膨胀后的区域足够覆盖整个穗丝所在的区域,此时核设置应该偏大,实验选取 200 作为膨胀的核大小,以确保能够覆盖整个穗须目标。3) 最后再将上述 2 个步骤后得到的图像做加法运算。因为都是二值图,计算结果只有 0、255、510 这 3 种情况。当相加值为 0 时使结果为 0,呈背景黑色;相加值为 255 时令结果为 128,呈不确定区域的灰色;当相加值为 510 时令结果为 255,呈前景白色。从而得到准确无误的三分图。自适应生成三分图的过程如图 3 所示。

1.4 优化 closed-form matting 算法实现玉米穗丝目标提取

Levin 等人^[25]提出了闭合型抠图算法 closed-form matting。假设图像局部平滑,则 F 和 B 为常数,生成 1 个关于 α 的二次代价函数。通过解稀疏线性方程组得到 α 是损失函数的全局最优。对于灰度图,下标 i 表示像素索引。假设在图片的任意一个足够小的像素窗口 w_j 上, F_i, B_i 近似为常值,则得:

$$I_i \approx \alpha_i F_i + (1 - \alpha_i) B_i, i \in w_j。$$

F_i, B_i 近似为常值,由上式变换可得 $\alpha_i \approx a_j I_i + b_j, i \in w_j$, 其中: $a = \frac{1}{F_j - B_j}, b = \frac{B_j}{F_j - B_j}$, 由于图片是许多相互重叠的窗口组成的,那么定义损失函数为:

$$J(\alpha, a, b) = \sum_{j \in W} \left(\sum_{i \in w_j} (\alpha_i - a_j I_i - b_j)^2 + \epsilon a_j^2 \right),$$

其中: W 表示窗口索引集合, ϵ 表示正则项系数。假设总窗口数为 n , 窗口尺寸为 3×3 。用矩阵的方法重写上式

$$\text{可得: } J(\alpha, a, b) = \sum_{k=1}^n |G_k \cdot c_k - \alpha_k|^2, \text{ 其中: } G_k = \begin{bmatrix} I_{k1} & 1 \\ \vdots & \vdots \\ I_{k9} & 1 \\ \sqrt{\epsilon} & 0 \end{bmatrix}, \alpha_k = \begin{bmatrix} \alpha_{k1} \\ \vdots \\ \alpha_{k9} \end{bmatrix}, c_k = \begin{bmatrix} a_k \\ b_k \end{bmatrix}。$$

定义 $J(\alpha) = \min J(\alpha, a, b)$, 通过最小二乘拟合可以将 a 和 b 用 α 表示,再经过求偏导和一系列矩阵运算可

得算法目标函数为:

$$J(\alpha) = \sum_{k=1}^n \|\mathbf{G}_k \cdot \mathbf{c}_k - \alpha_k\|^2 = \sum_{k=1}^n (\mathbf{G}_k \cdot \mathbf{c}_k - \alpha_k)^T (\mathbf{G}_k \cdot \mathbf{c}_k - \alpha_k) = \sum_{k=1}^n \alpha_k^T \mathbf{G}_k^T \mathbf{G}_k \alpha_k = \sum_{i=1}^n \alpha_k^T \mathbf{L}_k \alpha_k$$

其中: $\mathbf{L}_k = \mathbf{G}_k^T \mathbf{G}_k = \delta_{ij} - \frac{1}{|W_k|} \left(1 + \frac{1}{\frac{\epsilon}{|W_k|} + \sigma_K^2} (I_i - \mu_k)(I_j - \mu_k) \right)$, δ_{ij} 是 Kronecker 函数, 当 $i=j$ 时, $\delta_{ij}=1$, 其

余情况 $\delta_{ij}=0$, W_k 是窗口中像素个数, μ_k 和 σ_k^2 是当前窗口 k 的均值和方差。

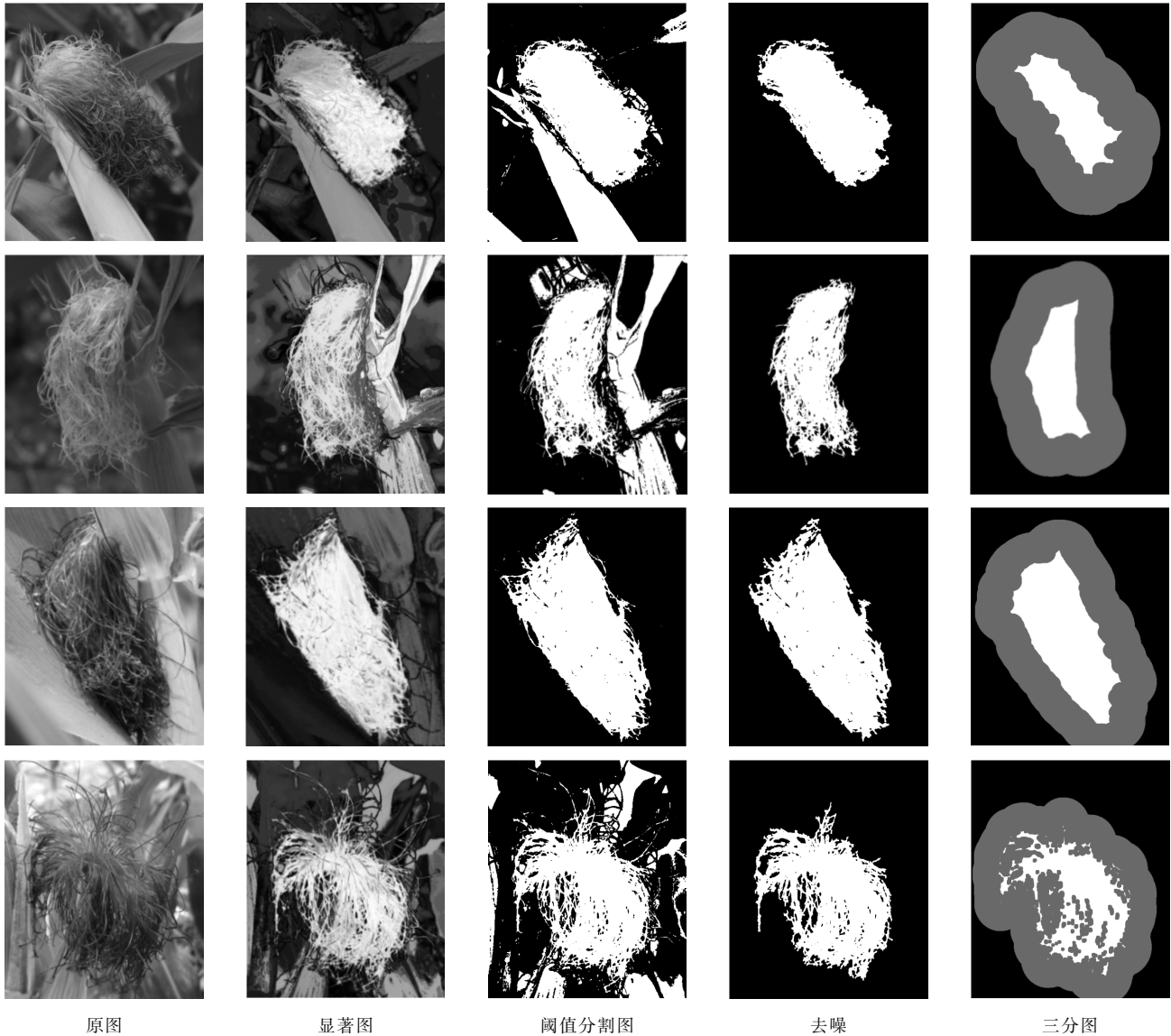


图 3 自适应生成三分图的过程

Fig. 3 Process of adaptively generating trimap

该算法将图像分解为前景像素和背景像素的线性组合, 对于单通道图片, 假设局部区域像素近似为常值; 对于彩色图片, 则假设局部区域的像素值在色彩空间中保持线性。这样的假设将图像方程求解问题变成了优化问题。

依据 Levin 等人^[25]提出的算法思想, 本文通过构建 Laplace 矩阵 $\mathbf{L}$ 替换目标函数中的分块的系数矩阵优化 closed-form matting 的目标函数, 从而简化目标函数。

证明 假设矩阵 $\mathbf{A}$ 和向量 $\mathbf{v}$ 分块: $\mathbf{A} = \begin{bmatrix} \mathbf{A}_{00} & \mathbf{A}_{01} \\ \mathbf{A}_{10} & \mathbf{A}_{11} \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} \mathbf{v}_0 \\ \mathbf{v}_1 \end{bmatrix}$, 则有:

$$\mathbf{v}^T \mathbf{A} \mathbf{v} = [\mathbf{v}_0^T \quad \mathbf{v}_1^T] \begin{bmatrix} \mathbf{A}_{00} & \mathbf{A}_{01} \\ \mathbf{A}_{10} & \mathbf{A}_{11} \end{bmatrix} \begin{bmatrix} \mathbf{v}_0 \\ \mathbf{v}_1 \end{bmatrix} = \mathbf{v}_0^T \mathbf{A}_{00} \mathbf{v}_0 + \mathbf{v}_0^T \mathbf{A}_{10} \mathbf{v}_1 + \mathbf{v}_1^T \mathbf{A}_{01} \mathbf{v}_0 + \mathbf{v}_1^T \mathbf{A}_{11} \mathbf{v}_1。$$

若 $\mathbf{A}_{10} = 0, \mathbf{A}_{01} = 0$, 即 $\mathbf{A}$ 是一个带状矩阵, 则有: $\mathbf{v}^T \mathbf{A} \mathbf{v} = \sum_{i=0}^1 \mathbf{v}_i^T \mathbf{A}_{ii} \mathbf{v}_i$ 。这与 closed-form matting 的目标函数相似。当向量 $\mathbf{v}$ 的某些位为 0 时, 去掉为 0 的位和 $\mathbf{A}$ 中相应的行和列, 结果是等价的。证毕

由上述例子, 可以推断存在一个带状矩阵 $\mathbf{L}$, 使得 $J(\alpha) = \sum_{i=1}^n \alpha_i^T \mathbf{L}_k \alpha_k = \alpha^T \mathbf{L} \alpha$ 。矩阵 $\mathbf{L}$ 通过 $\mathbf{L}_k$ 组装得来。 $\alpha = [\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_N]^T$ 是由每个像素的透明度组成的向量, $\alpha_k = [\alpha_{k_1} \quad \alpha_{k_2} \quad \cdots \quad \alpha_{k_g}]^T$ 是第 k 个窗口的所有像素透明度组成的向量, 窗口尺寸为 3×3 。因为图片是二维的, 按从左到右、从上到下的顺序依次扫描, 可知 α_k 是向量 α 的一部分。通过 $\mathbf{L}_k$ 组装得出 $\mathbf{L}$, 需要明确 2 个矩阵之间的对应关系。

$$\text{假设 } \alpha = [\alpha_1 \quad \alpha_2 \quad \alpha_3 \quad \alpha_4]^T, \alpha_k = [\alpha_1 \quad \alpha_2 \quad \alpha_4]^T, \text{ 且 } \mathbf{L}_k = \begin{bmatrix} L_{k,11} & L_{k,12} & L_{k,13} \\ L_{k,21} & L_{k,22} & L_{k,23} \\ L_{k,31} & L_{k,32} & L_{k,33} \end{bmatrix}。 \text{按照上述向量乘法的要求, 可知 } \mathbf{L}_k \text{ 在 } \mathbf{L} \text{ 中的位置分布如下:}$$

$$\mathbf{L}_k = \begin{bmatrix} L_{k,11} & L_{k,12} & 0 & L_{k,13} \\ L_{k,21} & L_{k,22} & 0 & L_{k,23} \\ 0 & 0 & 0 & 0 \\ L_{k,31} & L_{k,32} & 0 & L_{k,33} \end{bmatrix}。$$

上述例子说明说明 $\mathbf{L}_k$ 在 $\mathbf{L}$ 中的位置与 α_k 元素下标有关, 对应关系为: $\mathbf{L}_k(i, j) \rightarrow \mathbf{L}(k_i, k_j)$, 其中: k_i 是 α_k 中第 i 个元素的下标。当不同的 $\mathbf{L}_k$ 在 $\mathbf{L}$ 上的位置重叠时, 由于乘法矩阵是简单的线性操作, 将它们相加即可。

由此得证, 存在一个矩阵 $\mathbf{L}$, 使得 $\sum_{i=1}^n \alpha_i^T \mathbf{L}_k \alpha_k = \alpha^T \mathbf{L} \alpha$, 玉米穗丝三分图作为约束条件用 Lagrange 乘子法加入损失函数^[25], 可得最终目标函数: $\alpha = \arg \min \alpha^T \mathbf{L} \alpha + \lambda (\alpha^T - \mathbf{b}_s^T) \mathbf{D}_s (\alpha^T - \mathbf{b}_s^T)$, 其中: $\mathbf{b}_s$ 为用户指定像素的 α 值, 如果该像素是确定的前景或者背景, 则 $\alpha - \mathbf{b}_s = 0$, 这表示该像素的 α 值与用户指定的一致; $\mathbf{D}_s$ 是一个 $n \times n$ 的对角矩阵, 用来标记已知的前景和背景像素位置。

2 实验结果与分析

2.1 实验评价指标

实验以田间拍摄的不同形态和不同颜色的玉米穗丝图像为原始数据, 采用像素精度(pixel accuracy, PA)、准确率(precision)、召回率(recall)和综合评价指标(F-measure)作为本文结果的评价指标, 分别定义 x_{PA} 、 $P_{\text{precision}}$ 、 R_{recall} 和 F 。计算公式分别为 $x_{PA} = \frac{I_{TP} + I_{TN}}{I_{TP} + I_{TN} + I_{FP} + I_{FN}}$, $P_{\text{precision}} = \frac{I_{TP}}{I_{TP} + I_{FP}} \times 100\%$, $R_{\text{recall}} = \frac{I_{TP}}{I_{TP} + I_{FN}} \times 100\%$, $F = 2 \times \frac{P_{\text{precision}} \times R_{\text{recall}}}{P_{\text{precision}} + R_{\text{recall}}} \times 100\%$, 其中: I_{TP} 是正确分割为穗丝的像素数, I_{FP} 是错误分割为穗丝的像素数, I_{FN} 为错误分割为背景的像素数。

2.2 实验结果分析

为验证本文方法的优越性, 分别用基于点运算经典方法 OTSU、Grabcut 和基于神经网络的深度学习方法 SegFormer 和 DeeplabV3 这 4 种分割方法, 与本文方法对玉米穗丝目标提取进行对比, 实验结果平均值如表 1 所示。

由实验结果可知, 本文方法提取玉米穗丝 F-measure 值为 94.16%, 高于 Grabcut 的 88.57% 和 SegFormer 的 89.88%, 同时高于 DeeplabV3 语义分割 5 个百分点。PA 值为 97.96%, 分别比 OTSU 和 Grabcut 高出 10.43% 和 3.35%, 同时比基于神经网络的深度学习方法 SegFormer 和 DeeplabV3 高出 1.67% 和 1.91%, 数据表明本文方法相较于硬分割方法表现优异, 玉米穗丝能被正确提取。

本文方法与对比实验方法提取穗丝效果如图 4 所示。

分析图 4 可知, OTSU 阈值分割受背景土壤、叶片影响很大, 且光线下穗丝明暗程度也极大地影响了分割结

果,无法将穗丝完整提取,相比之下 Grabcut 算法能够极大地提取到穗丝目标,但是 Grabcut 算法需要进行多次人机交互,才能得到较好的分割效果。基于神经网络的深度学习算法 SegFormer 和 DeeplabV3 分割效果较好,能够将形态各异的穗丝提取,且提取较为完整,但在前景和背景交界处,无法精细展示穗丝细节。而本文方法提取的穗丝,与真实区域吻合,较好地解决了硬分割方法对于穗丝这种特殊对象的分割缺陷,既能将穗丝完整提取,又能对穗丝细节有更精细的表现。

表 1 评价结果对比

Tab. 1 The results of the evaluation

%

方法	x_{PA}	$P_{precision}$	R_{recall}	F
OTSU	87.53	62.05	80.81	70.20
Grabcut	95.61	84.80	93.10	88.57
SegFormer	96.29	83.95	96.71	89.88
DeeplabV3	96.05	84.89	93.59	89.03
本文方法	97.96	94.42	93.90	94.16

注:加黑的字体表示实验中取得最好效果的指标。

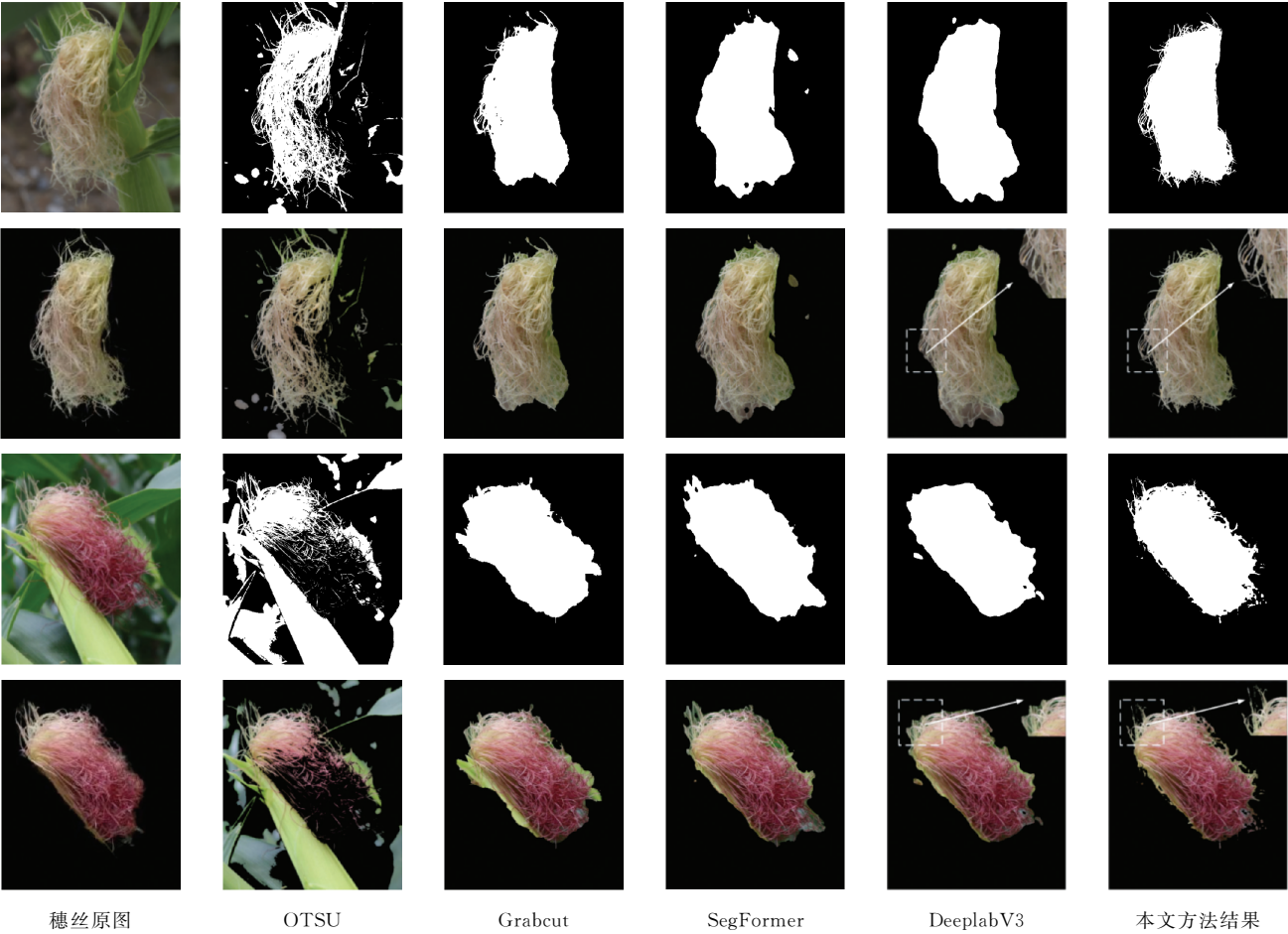


图 4 实验结果图

Fig. 4 Pictures of experimental results

3 结论

玉米穗丝不同于其他有较为清晰边界的农作物,因此对穗丝进行提取存在一定难度。如用现有的硬分割算法对穗丝进行提取,提取的结果呈现块状,在穗丝与背景交界处的表现较差。针对这一缺陷,本文从新的角度实

现对穗丝的提取:

1) 引入 matting 软分割算法对穗丝进行提取。通过对应关系,将原分块的系数矩阵整合为一个 Laplace 矩阵来优化 closed-form matting 算法。

2) 提出基于视觉注意机制的显著性检测从而自适应生成三分图的方法。三分图准确度决定 matting 算法的精确度。首先基于直方图对比度的显著性检测算法预估目标显著性,形成显著图,再利用自适应阈值分割对显著图初步预分割,然后寻找最大连通域获得单一目标二值图像,最后根据穗丝特点确定形态学处理的最佳卷积核并生成包含目标、背景和不确定区域的自适应三分图。

实验结合三分图结果和优化过的算法实现穗丝精细化提取。结果表明该方法提取不同形态的穗丝都有较好表现,能有效弥补硬分割算法对玉米穗丝这种特殊对象的分割缺陷,提取的目标区域接近真实区域,从而为智能化监测玉米穗丝生长状况提供依据。

参考文献:

- [1] 张万红,刘文兆. 覆膜玉米冠层图像分割方法[J]. 浙江大学学报(农业与生命科学版),2017,43(5):649-656.
ZHANG W H,LIU W Z. Image segmentation method of plastic-film corn canopy[J]. Journal of Zhejiang University (Agriculture and Life Sciences),2017,43(5):649-656.
- [2] 王春雷,卢彩云,陈婉芝,等. 基于遗传算法和阈值滤波的玉米根茬行图像分割[J]. 农业工程学报,2019,35(16):198-205.
WANG C L,LU C Y,CHEN W Z,et al. Image segmentation of maize stubble row based on genetic algorithm and Image segmentation of maize stubble row based on genetic algorithm and threshold filterina noise[J]. Transactions of the Chinese Society of Agricultural Engineering,2019,35(16):198-205.
- [3] 许高建,沈杰,徐浩宇. 基于 Lab 颜色空间下的小麦赤霉病图像分割[J]. 中国农业大学学报,2021,26(10):149-156.
XU G J,SHEN J,XU H Y. Image segmentation of wheat scab based on Lab color space[J]. Journal of China Agricultural University,2021,26(10):149-156.
- [4] 车金庆,王帆,王艺洁,等. 基于视觉注意机制的黄绿色苹果图像分割[J]. 江苏农业学报,2018,34(6):1347-1353.
CHE J Q,WANG F,WANG Y J,et al. A segmentation method of yellow and green apple images based on visual attention mechanism[J]. Jiangsu Journal of Agricultural Sciences,2018,34(6):1347-1353.
- [5] BOYKOV Y Y,JOLLY M P. Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images[C]// WERNER B. Proceedings Eighth IEEE International Conference on Computer Vision. Piscataway:IEEE,2001:105-112.
- [6] ROTHER C,KOLMOGOROV V,BLAKE A. "GrabCut" interactive foreground extraction using iterated graph cuts[J]. ACM Transactions on Graphics,2004,23(3):309-314.
- [7] 许景辉,邵明烨,王一琛,等. 基于迁移学习的卷积神经网络玉米病害图像识别[J]. 农业机械学报,2020,51(2):230-236.
XU J H,SHAO M Y,WANG Y S,et al. Recognition of corn leaf spot and rust based on transfer learning with convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery,2020,51(2):230-236.
- [8] HE K M,GKIOXARI G,DOLLÁR P,et al. Mask r-cnn[C]//O' CONNER L. Proceedings of the 2011 IEEE international conference on computer vision. Piscataway:IEEE,2017:2961-2969.
- [9] REN S Q,HE K M,GIRSHICK R,et al. Faster r-cnn: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2017,39(6):1137-1149.
- [10] 王璨,武新慧,张燕青,等. 基于双注意力语义分割网络的田间苗期玉米识别与分割[J]. 农业工程学报,2021,37(9):211-221.
WANG C,WU X H,ZHANG Y Q,et al. Recognition and segmentation of maize seedlings in field based on dual attentionsemantic segmentation network[J]. Transactions of the Chinese Society of Agricultural Engineering,2021,37(9):211-221.
- [11] CHUANG Y Y,CURLESS B,SALESIN D H,et al. A Bayesian approach to digital matting[C]//TACDBS A,BALDWIN T. Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision & Pattern Recognition. Piscataway:IEEE,2003.
- [12] LEVIN A,RAV-ACHA A,LISCHINSKI D. Spectral matting [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence,2008,30(10):1699-1712.
- [13] SUN J Y,JIA J Y,TANG C-K,et al. Poisson matting[C]//MARKS J. ACM SIGGRAPH 2004 Papers. New York:Association for Computing Machinery,2004:315-321.
- [14] CHENG Z Z,YANG Q X,SHENG B. Deep colorization [C]// O' CONNER L. Proceedings of 2015 IEEE International

- Conference on Computer Vision, Piscataway; IEEE, 2015; 415-423.
- [15] LARSSON G, MAIRE M, SHAKHNAROVICH G. Learning representations for automatic colorization [C]//LEIBE B, MATAS J, SEBE N, et al. European Conference on Computer Vision. Cham; Springer, 2016; 577-593.
- [16] CHO D, KIM S, TAI Y W, et al. Automatic trimap generation and consistent matting for light-field images [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 39(99): 1504-1517.
- [17] SINGH S, JALAL A S. Automatic generation of trimap for image matting [J]. International journal of machine intelligence and sensory signal processing, 2014.
- [18] OTSU N. A threshold selection method from gray-level histograms [J]. IEEE Transactions on Systems, Man, and Cybernetics, 1979, 9(1): 62-66.
- [19] ROTHER C, KOLMOGOROV V, BLAKE A. "GrabCut" -interactive foreground extraction using iterated graph cuts [J]. ACM Transactions on Graphics, 2004(3): 23.
- [20] XIE E Z, WANG W H, YU Z D, et al. SegFormer: simple and efficient design for semantic segmentation with transformers [EB/OL]. (2021-05-31)[2022-10-18]. <https://arxiv.org/abs/2105.15203>.
- [21] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation [EB/OL]. (2017-06-17)[2022-10-18]. <https://arxiv.org/abs/1706.05587v3>.
- [22] ITTI L, KOCH C, NIEBUR E. A model of saliency-based visual attention for rapid scene analysis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(11): 1254-1259.
- [23] 李婉蓉, 徐丹, 史金龙, 等. 显著性物体检测研究综述: 方法、应用和趋势 [J]. 计算机应用研究, 2022, 39(7): 1941-1950.
- LI W R, XU D, SHI J L, et al. Review of salient object detection: models, applications and prospects [J]. Application Research of Computers, 2022, 39(7): 1941-1950.
- [24] CHENG M M, ZHANG G X, MITRA N J, et al. Global contrast based salient region detection [C]//Computer Vision and Pattern Recognition, June 20-25, 2011, Colorado Springs, CO, USA. Piscataway; IEEE, 2011.
- [25] LEVIN A, LISCHINSKI D, WEISS Y. A closed-form solution to natural image matting [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2007, 30(2): 228-242.

Refined Segmentation of Corn Silk Based on Adaptive Trimap

XIONG Junyi, ZHU Deli, WEN Rui

(1. College of Computer and Information Science, Chongqing Normal University;

2. Research Center of Chongqing Digital Agricultural Service Engineering Technology, Chongqing 401331, China)

Abstract: The corn silk is the pollinating organ of maize, and its growth and development are closely related to the yield. In order to realize the automatic monitoring and evaluation of corn silk characters, proposing a refined extraction method of corn silk based on adaptive trimap. First, the significance detection algorithm based on histogram contrast is used to estimate the significance of the target and form a saliency image. Then, the saliency image is preliminarily pre segmented by using adaptive threshold segmentation, and finding the largest general domain to obtain a single target binary image. According to the characteristics of the silk, determining the best convolution kernel for morphological processing, and generating a trimap containing the background and uncertain regions. Laplace matrix is used to optimize the block matrix of the objective function in the closed form matching algorithm. Finally, according to the characteristics of the corn silk, the soft segmentation of the corn silk is carried out by combining the results of the trimap and the optimized algorithm to achieve refined extraction. The pixel accuracy value of this research method is 97.96%, and the comprehensive evaluation index F-measure value is 94.16%. Compared with OTSU algorithm and Grabcut algorithm based on point operation, and deep learning algorithms SeFormer and DeeplabV3 based on neural network, the accuracy is increased by 10.43, 2.35, 1.67 and 1.91 percentage points respectively, and the F-measure value is increased by 23.96, 5.59, 4.28 and 5.13 percentage points respectively. This method can effectively remedy the defect of hard segmentation algorithm in the segmentation of special objects such as corn ear silk. The extracted target area is close to its real area with high accuracy, thus providing technical support for intelligent monitoring of corn silk growth.

Keywords: corn silk; saliency detection; image segmentation; object extraction

(责任编辑 黄 颖)