

数字图像中绿色植物的快速提取方法

柯秋红 张俊梅* 田野
(北京林业大学工学院 北京 100083)

摘要 可见光图像可以反应植物的表型特征,在高通量表型测量的要求下,提出一种从数字图像中提取绿色植物区域的方法。首先对原图像运用高斯函数进行滤波预处理,然后利用图像的绿、红、蓝分量进行色彩分割,并通过设置阈值对图像进行二值化处理,针对二值化处理后的杂散点采用形态学开运算算法去除,最后对未提取出的植物区域的孔洞进行填充。从实验结果可以看出,原图像中的绿色植物区域被完整地提取出来。该方法能快速有效地提取图像中的绿色植物区域,实现对植物的无损测量。

关键词 植物提取 色彩分割 形态学算法

中图分类号 TP391 文献标识码 A DOI:10.3969/j.issn.1000-386x.2013.10.073

RAPID EXTRACTION METHOD FOR GREEN PLANTS IN DIGITAL IMAGES

Ke Qiuhong Zhang Junmei* Tian Ye
(Institute of Technology, Beijing Forestry University, Beijing 100083, China)

Abstract Visible light images can reflect the phenotypic characteristics of plants. A method to extract the green plant from digital image on the requirement of high throughput phenotyping measurement is presented. First, the original image is filtered by using Gaussian function. Then the green, red and blue components of the image are used to make colour segmentation. Binarisation of the image is processed by setting the threshold. For desultory points incurred from the binarisation processing, the morphological opening operation algorithm is applied to remove them. Finally the cavities in the plant areas which have not been extracted are filled. From the experimental result it can be seen that the green plant area in original image is completely extracted. The method can quickly and effectively extract the green plant area from image, and achieves the non-destructive measurement of the plant.

Keywords Plant extraction Colour segmentation Morphological algorithm

0 引言

植物表型组学用于提供植物表现型信息数据,植物的表现型是生长环境、基因组、遗传变异的必然结果^[1-3]。对植物表型的研究在基因学、遗传学等微观领域与宏观表象之间构建起连接的桥梁,在植物育种、改善温室效应和培养新作物等方面提供更为高效的手段与研究方法^[4-8]。植物的表型主要是通过光谱图像获得,其中很重要的部分是可见光图像。受到植物本身的限制如何高效准确地从可见光图像中提取植物部分,是进行表型分析的前提。本文在高通量表型测量要求的前提下,提出了从数字图像中提取植物的方法。对植物的可见光图像处理的传统方法是将叶片采摘下来,通过扫描得到图像进而进行特征提取^[9]。本研究的切入点是直接从数字图像中快速查找绿色植物并准确提取。拟解决的关键问题包括如何克服图像中土壤、容器、背景的干扰及不同光照条件对图像的影响。

1 材料与方法

利用数码相机获取植物图像如图 1 所示。该图像中包括绿色植株区域、白色的花盆容器、泥土、红色的背景。由于叶片上

覆盖有蜡质层,白天在室内光照条件下,图像中的叶片部分存在反光的情况,这些都为植物的提取带来了困难。

1.1 图像预处理

获取图像后,首先需要对图像进行滤波预处理,降低噪声。由于噪声属于高频信号,本文选用高斯滤波器去除噪声,其滤波传递函数的形式如公式(1)所示,其中 $D(u, v)$ 是距频率矩形中心的距离, D_0 是截止频率,即方差^[10]。

$$H(u, v) = e^{-D^2(u, v)/2D_0^2} \tag{1}$$

经离散化后,取 7×7 高斯模板进行滤波。

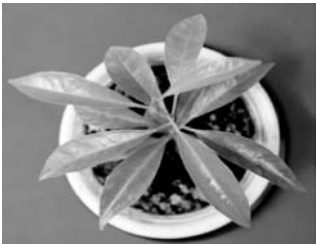


图 1 原始图像:箭头所指区域即为明显的反光区域,叶片在图像中表现出白色

1.2 基于色彩的图像分割

预处理完成后,紧接着进行植物提取。也就意味着将植物从图像中分割出来。常用的分割方法有聚类法、基于区域的方法等。聚类法是对图像中的每一个像素点的色彩特征进行聚类分析;基于区域的方法包括区域生长、区域分裂和聚合等^[10]。这些方法运算时间较长,不能满足高通量环境下的高速运算需要,且由于图像中包含多种特征区域,分割算法不稳定。

考虑到植物的色彩表现为绿色,可以在 RGB 色彩空间中,将绿色作为特征,通过使用阈值的方法将图像分割为绿色和非绿色部分。随后便可以实现植物的提取。在 RGB 空间中,任一像素点的颜色可以看作红、绿、蓝三个分量不同比例的组合,一幅彩色图像即可以看成是红、绿、蓝三个灰度平面组合而成,如果某点处的三个分量为灰度的最大值则该点表现为白色。但是由于 RGB 色彩空间受光照影响很大,为了减少光照的影响,首先在 RGB 空间中进行归一化处理。即从图像中提取出绿分量 g 、红分量 r 和蓝分量 b ,归一化后为 R 、 G 、 B (见式(2))。为进行绿色特征的判断,考虑使用 $\begin{bmatrix} -1 & 2 & -1 \end{bmatrix}$ 模板对图像各点处 (R, G, B) 分量进行处理,如式(3)所示。该模板强调绿色分量,压制红色和蓝色分量,且模板和为 0。模板具有对称性,且长度为奇数,因此具有线性相位,保证时不变要求。理想情况下,对于蓝色和红色显著的区域,使用模板运算后的结果必为负值,而白色区域使用模板运算后的结果为 0,只有绿色区域在模板运算后表现为大于 0 的结果。其他颜色区域模板运算结果必为负值,这个特点非常重要,只要使用大于 0 作为阈值判断的依据便可以很方便地将绿色分割出来,避免了其他算法中阈值的选择,体现了算法对图像的无关性。由于从归一化到模板运算是每个像素点使用加法和乘法,因此该算法可以在数字信号处理器 DSP 上高效实现,在保证 DSP 性能发挥的同时,也可以实现高速处理,满足了高通量环境下对处理速度的要求。

$$\begin{cases} G = \frac{g}{g+r+b} \\ R = \frac{r}{g+r+b} \\ B = \frac{b}{g+r+b} \end{cases} \quad (2)$$

$$ROI = \begin{bmatrix} -1 & 2 & -1 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3)$$

1.3 去除杂散点

对图像采用式(3)的方法处理后,用 0 进行阈值化,即产生二值图像,只要大于 0 便可以判断为绿色区域。此时图像会产生一些杂散点,可采用形态学算法中的开运算去除^[10]。

结构元 B 对图像 A 的开运算记作 $A \circ B$,如式(4)所示,对图像先腐蚀后膨胀。腐蚀可以去除杂散点,但同时也使叶片区域变小,故需通过膨胀操作来恢复。

$$A \circ B = (A \ominus B) \oplus B \quad (4)$$

腐蚀记作 $A \ominus B$,如式(5)所示,其中, A^c 是 A 的补集, \emptyset 是空集。

$$A \ominus B = \{z \mid (B)_z \cap A^c \neq \emptyset\} \quad (5)$$

膨胀记作 $A \oplus B$,如式(6)所示,其中, \hat{B} 是 B 关于原点对称的映像, \emptyset 是空集。

$$A \oplus B = \{z \mid (\hat{B})_z \cap A \neq \emptyset\} \quad (6)$$

1.4 图像重构

图像开运算后,虽然杂散点被完全消除,但是边缘却会发生变化,需要利用形态学重构算法进一步恢复图像^[10],即将上述经过开运算后的图像作为标记图像 F ,阈值化后没有经过处理的图像作为掩膜图像 G , F 对 G 的形态学重构 $R_G^D(F)$ 为:

$$R_G^D(F) = D_G^{(k)}(F) \quad (7)$$

通过反复迭代,直至 $D_G^{(k)}(F) = D_G^{(k+1)}(F)$,其中:

$$D_G^{(1)}(F) = (F \oplus B) \cap G \quad (8)$$

$$D_G^{(n)}(F) = D_G^{(1)}[D_G^{(n-1)}(F)] \quad (9)$$

$D_G^{(n)}(F)$ 表示标记图像 F 关于图像模板 G 的 n 次侧地膨胀, B 为结构元素,即将标记图像 F 的膨胀限定在掩膜图像 G 中。

1.5 填充孔洞

原始图像中,叶片反光使图像中的叶片区域表现为白色,而白色区域中绿分量 G 、红分量 R 、蓝分量 B 相等,用式(3)方法算出的结果为 0,该结果在阈值化后变成背景,从而导致提取的叶片区域出现孔洞;而叶片之间存在间隙,用式(3)处理并阈值化后也产生孔洞,两类孔洞混在一起。为消除该影响,需要对孔洞进行判断,若是叶片上面的孔洞,才进行填充。判断方法是读取孔洞所在区域的原始图像的红蓝分量,若是其值大于定义的阈值,则该孔洞由原始图像中叶片上的白色区域产生,需要对其进行填充,否则是背景,不进行填充。

孔洞填充算法是从孔洞中的一个给定点 X_0 开始,按照式(10)进行迭代,直到 $X_k = X_{k-1}$ ($k = 1, 2, 3, \dots$),迭代结束^[10]。

$$X_k = (X_{k-1} \oplus B) \cap A^c \quad (10)$$

其中 A 表示一个具有 8 连通边界的集合,边界包围孔洞, A^c 是 A 的补集, B 为结构元素, X_k 包含所有被填充的孔洞。

1.6 特征提取

提取出叶片区域后即可对其利用函数进行特征提取,绘制质心和最小凸多边形^[10,11]。

2 结果与分析

图 2 所示为原始图像和经过滤波预处理后的图像。原始图像的大小为 1024×768 ,图中植物的叶片与的花盆、泥土和背景形成鲜明对比,从图中还可以看出叶子的反光情况。

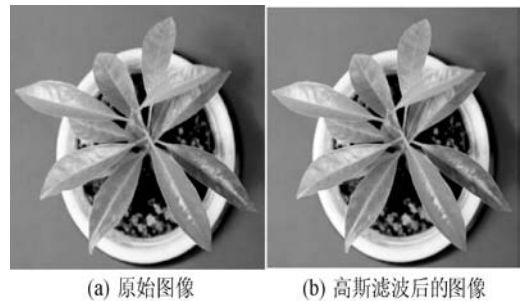
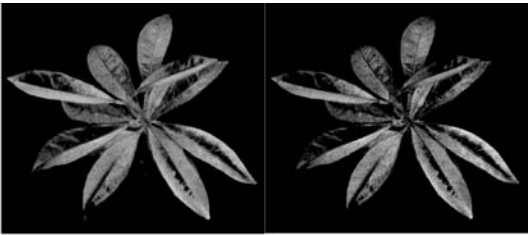


图 2 图像预处理

滤波后提取出图像的绿 G 、红 R 、蓝 B 三个分量,进行归一化,并用式(3)进行处理。处理结果如图 3(a)所示,可以发现,在图 3(a)中土壤和容器部分也被提取出来,原因在于光谱中红色与绿色谱线有部分重叠和 CCD 成像过程中的噪声,该公式需增加修正项“ $-1.4 \times R + G$ ”。



(a) 用公式(3)处理后的图像, 图中有一部分背景被提取出来 (b) 用增加修正项的公式处理后的图像, 背景被消除

图3 色彩分割

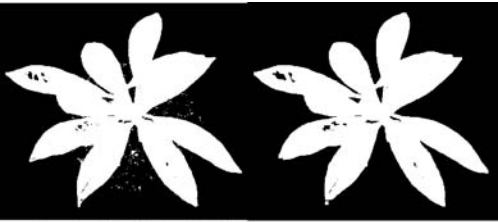
用带有修正项的公式处理图像后, 对其进行阈值化, 形成二值图像。当阈值选择过大时, 将导致部分提取出的叶片边缘的点缺失严重, 故选用 0 作为阈值, 如图 4 所示。



(a) 采用的阈值为 0 (b) 采用的阈值为 0.03 箭头标注位置的叶片阈值化处理后缺失严重 (c) 采用的阈值为 0.05 箭头标注位置的叶片阈值化处理后缺失严重

图4 阈值处理

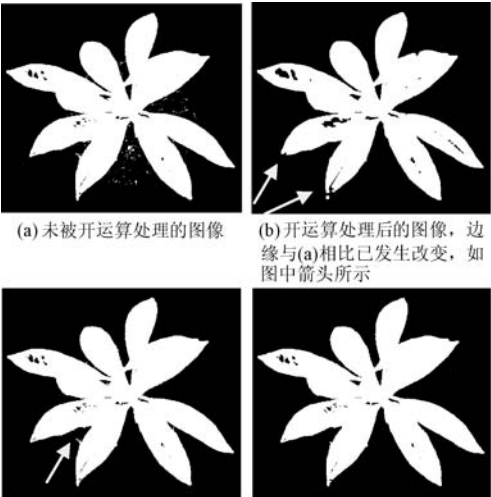
得到二值图像后, 采用形态学的开运算去除叶片周围的杂散点, 如图 5 所示。



(a) 为阈值后的原始二值图像 (b) 开运算处理后的图像, 杂点已完全消除

图5 开运算处理

叶片周围的杂散点已经被完全消除, 但是叶片的边缘却发生了变化, 采用形态学重构恢复边缘。重构后如图 6 所示, 处理后发现边缘已被较好地恢复。

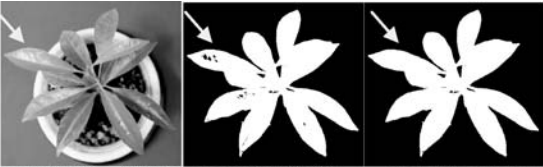


(a) 未被开运算处理的图像 (b) 开运算处理后的图像, 边缘与(a)相比已发生改变, 如图中箭头所示 (c) 重构算法处理后的图像, 其边缘已经很好地恢复, 但同时出现了杂点 (d) 再次采用开运算处理后的结果, 杂点已被消除

图6 图像重构

重构后需要填补图像中叶片区域的孔洞, 首先找出所有的

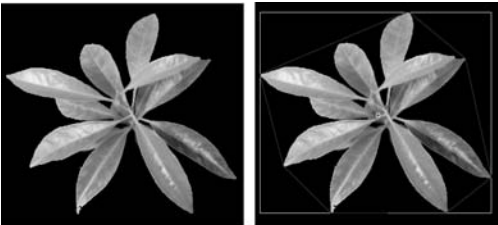
孔洞, 进而判断该孔洞属于背景, 即叶子之间的间隙, 还是叶片本身的反光区。判断方法是读取孔洞对应区域的原始图像的红分量和蓝分量, 若其超过阈值, 则其为白色, 即叶片上面的反光区, 否则就是背景, 即叶片之间的间隙。通过多次实验最终阈值选定为 100。结果如图 7 所示。



(a) 原始图像, 箭头标注处的叶片为明显的反光区域, 即其在图像中表现为白色 (b) 对应的带有孔洞的二值图像 (c) 填充孔洞后的结果, 如箭头所示, 孔洞已被填充

图7 孔洞填充

获取到完整的植物区域的二值图像后, 将其作为标记, 从原始图像提取出植物, 然后利用区域特征提取方法, 计算出植物的区域描绘子。图 8 给出了描述植物区域的最小凸多边形及其质心。



(a) 提取出的绿色叶片区域 (b) 绘制的包围叶片的最小多边形, 并用“.”表示出区域的质心

图8 特征提取

3 结 语

获取具有复杂背景的绿色植物图像, 对图像进行处理, 包括滤波、色彩分割、去除杂散点、图像重构、填充孔洞等。其中色彩分割运用整幅图像的红分量、绿分量、蓝分量进行处理, 有效地提高了算法的效率。针对叶片反光导致产生孔洞的情况, 通过判断原始图像的红分量和蓝分量是否超过阈值来填充孔洞, 在实验中通过调整参数最终达到最佳效果, 结果表明该算法确实能有效地提取出植物的叶片区域, 以便进行特征提取, 实现了对植物的无损测量。具体程序可参看: http://blog.sina.com.cn/s/blog_6beeced7010133hx.html。

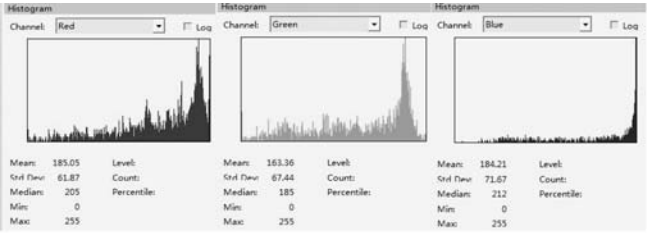
参 考 文 献

[1] Varki A, Altheide T K. Comparing the human and chimpanzee genomes; searching for needles in a haystack [J]. Genome Research, 2005, 15: 1746 - 1758.

[2] Oti M, Huynen M A, et al. Phenome connections [J]. Trends Genet, 2008, 24 (3) : 103 - 106.

[3] Wu X B, Liu Q F, et al. Align human interactome with phenome to identify causative genes and networks underlying disease families [J]. Bioinformatics, 2009, 25 (1) : 98 - 104.

[4] Kelsoe J R, Niculescu A B. Finding genes for bipolar disorder in the functional genomics era: from convergent functional genomics to phenomics and back [J]. CNS Spectr, 2002, 7 (3) : 215 - 216, 223 - 226.



(a) 最强光下的强度变化 (b) 次强光下的光照强度变化 (c) 弱光下光强度的变化

图 2 光强度的变化趋势

为了验证上述基于颜色梯度的背景差分法在不同强度光照强度变化下的目标分割的有效性,以 Visual Studio 2008 为平台,利用 C 语言进行编程,对同一个光照条件下,拍摄的人物图像进行分割,并传统的高斯混合模型检测进行比较。图 3 为两个视频序列的起始视频帧。



图 3 两个视频序列的起始帧图像

运用传统高斯混合模型在进行前景检测实验时,赋值 K 为 5,所有的高斯分布的原始均值是第一帧的像素均值,原始方差是 25,原始的权重是 0.2。通过本文算法计算,改进高斯模型参数同传统高斯模型无差异,公差取值是 0.5,选用前 25 帧对高斯混合模型进行原始的设计, $\alpha = 0.005$,并且,将检测的相应结果用彩色的方式表达,从而完成接下来更完善的操作。为更加明确呈现出相应的结果,设置相应的背景像素是 1(实际过程时,背景像素是 0),部分检测结果如图 4 所示。

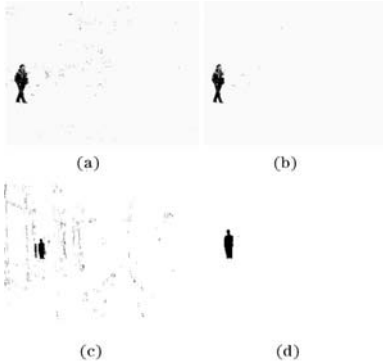


图 4 实验效果图

上述实验结果中,第一行为本文方法下得到的实验结果,检测得较为准确。第二行为传统的高斯模型检测的结果。通过对比可以看到,第二列为采用混合高斯模型的检测结果明显比第一行差,这是因为在强光照的变化环境下,传统算法有着较为明显的缺陷。对比传统混合高斯模型,很容易发现本文基于颜色梯度融合高斯模型的光照突变干扰消除算法明显优于使用传统混合高斯模型的算法。检测结果中,噪点明显减少。

4 结 语

本文提出了一种基于颜色梯度融合高斯模型的光照突变干扰消除方法。首先通过在传统的高斯图像模型中加入颜色突变梯度融合算子,对传统的模型进行合理的优化,尝试性地消除光照突变对模型稳定性的影响;然后利用背景梯度分布获取的梯

度矢量概率对背景区域进行合理的区分。后期的计算机仿真实验结果表明:优化后的模型对于消除图像在光照突变环境下的干扰效果较好,最后通过仿真验证了该方法的有效性。

参 考 文 献

[1] Yao Y,Abidi B,Doggaz N. Evaluation of sharpness measures and search algorithms for the auto-focusing of high magnification images[C]//SPIE, Orlando,2006(6246).

[2] KuangChern N,Ang N P,Jr M H. Practical Issues in Pixel-Based Auto-focusing for Machine Vision [C]//Proceedings of the 2001 IEEE ICRA,2001(3):2791-2796.

[3] Rojas R. A short proof of the posterior probability property of classifier neural networks[J]. Neural Comutation,1996(8):41-43.

[4] Lu Y,Pyandeh S. Cooperative Hybrid Multi-camera Tracking for People Surveillance[J]. Canadian Journal of Electrical and Computer Engineering,2008,33(3):145-152.

[5] Geren W,Murphy T,Pankaskie T. Analysis of airborne GPS multipath effects using high-fidelity EM model[J]. IEEE Trans. on Aerospace Electronic and Systems,2008,44(2):711-723.

[6] Giremus A,Tourneret J Y,Calmettes V. A Particle Filtering Approach for Joint Detection/Estimation of Multipath Effects on GPS Measurements[J]. IEEE Trans. on Aerospace Electronic and Systems,2007,55(4):1275-1285.

[7] Lee S Y,Kumar Y,Cho J M,et al. Low-power passive focus measure operator based on the DCT for mobile phones[J]. The Imaging Science Journal,2008,56(6):301-306.

[8] Yang G,Nelson B J. Wavelet-Based autofocusing and unsupervised segmentation of microscope images[C]//Proceedings of the 2003 IEEE/RSJ Intl. Conference on Intelligence Roberts and Systems,2003(8):2143-2148.

[9] 陈晓彬. 基于小波的自动对焦算法研究[J]. 通信技术,2010,43(6):210-212.

[10] 王超,蒋远大,翟光洁,等. 基于数字图像处理的自控显微镜聚焦算法研究[J]. 仪器仪表学报,2009,30(6):1290-1295.

(上接第 268 页)

[5] Niculescu A B,Lulow L L,et al. PhenoChipping of psychotic disorders: a novel approach for deconstructing and quantitating psychiatric phenotypes[J]. Am. J. Med. Genet. B Neuropsychiatr. Genet,2006,141(6):653-662.

[6] Hegele R A,Joy T R,et al. Lipodystrophies:indows on adipose biology and metabolism[J]. Journal of Lipid Research,2007,48(7):1433-1444.

[7] Christophe R,Jan P,Frankard V,et al. Traitmill: a discovery engine for identifying yield-enhancement genes in Cereals [J]. Molecular Plant Breeding,2005,3(5):753-759.

[8] Rajendran K,Tester M,et al. Quantifying the three main components of salinity tolerance in cereals[J]. Plant, Cell & Environment,2009,32(3):237-249.

[9] 陈爱军,李东升,董光辉,等. 一种基于 MATLAB 的植物叶片参数测量系统[J]. 中国计量学院学报,2010,21(4):310-313.

[10] 冈萨雷斯,伍兹. 数字图像处理[M]. 北京:电子工业出版社,2011.

[11] 冈萨雷斯,伍兹. 数字图像处理(Matlab 版)[M]. 北京:电子工业出版社,2005.