

文章编号:2095-7386(2020)01-0013-05

DOI:10.3969/j.issn.2095-7386.2020.01.003

基于色彩空间的航拍油菜花图像分割

刘昌华¹,胡林龙¹,杨康¹,伍晓明²

(1. 武汉轻工大学 数学与计算机学院,湖北 武汉 430023;

2. 农业部油料作物生物学与遗传育种重点实验室 中国农业科学院油料作物研究所,湖北 武汉 430062)

摘 要:针对花期识别过程中油菜花航拍图像分割效果不理想的问题提出一种基于色彩空间分割油菜花图像的方法。首先在 Lab 空间利用 K-means 算法对油菜图像进行初次处理,然后在 HSV 空间使用颜色分割算法对聚类结果进行二次处理,最后把分割出的油菜花映射到 RGB 空间,以期完成油菜与油菜花的有效分割。选取了 60 幅不同的航拍油菜图像进行分割实验。结果表明采用此方法不仅能够较好地分割出油菜花,而且还能有效地避免光照的影响。此实验结果能为后期识别油菜花花期生长规律的研究提供参考。

关键词:航拍油菜图像;K-means 聚类算法;颜色分割算法;油菜花

中图分类号:TP 391

文献标识码:A

Color space-based aerial rapeseed flower image segmentation

LIU Chang-hua¹, HU Lin-long¹, YANG Kang¹, WU Xiao-ming²

(1. School of Mathematics and Computer Science, Wuhan Polytechnic University, Wuhan 430023, China;

2. Oil Crops Research Institute, Chinese Academy of Agricultural Sciences, Wuhan 430062, China)

Abstract: Aiming at the problem of unsatisfactory segmentation effect of aerial image of rapeseed flowers in the process of florescence recognition. This paper proposes a method combining K-means algorithm and color segmentation algorithm to segment rapeseed image. Firstly, the K-means algorithm was used to first process the rapeseed image in Lab space. Then, the clustering results were processed once again in HSV space using color segmentation algorithm. Finally, the segmented rapeseed was subjected to morphological treatment to complete the effective segmentation of rapeseed and rapeseed flowers. Sixty different aerial rapeseed images were selected for segmentation experiments. The results show that this method can not only segment rapeseed well, but also effectively avoid the influence of illumination. The results of this experiment can provide reference for the later study of the flowering period of rapeseed.

Key words: drone rapeseed image; K-means clustering algorithm; color segmentation algorithm; image segmentation

1 引言

油菜作为我国重要的油料作物,在我国油料作

物供给中一直占据着较大的比重,种植面积以及作物产量都位于油料作物首位,随着全球经济化的到来以及我国社会经济的迅速发展,国民经济水平的

收稿日期:2019-12-20.

作者简介:刘昌华(1963-),男,教授,E-mail:494456957@qq.com.

基金项目:国家重点研发计划(2016YFD0100202).

提高,我国对食用植物油的需求越来越大^[1]。适宜的开花期对增强油菜抗病性以及提高油菜产量有着非常重要的作用,因此就如何判别油菜花期是一个值得我们去探究的课题。但由于在获取大田油菜生长信息时需要消耗大量的人力物力资源,操作过程繁琐、时间较长,容易错过获取大田油菜生长信息的最佳时间,同时也容易受人们的主观因素影响出现误判等情况。

随着数字图像处理技术和现代农业的快速发展,大田农业作物生长信息的图像数量急剧增长,因此就如何高效处理和分析这些图像数据已成为研究的热点问题^[2]。如学者刘国成^[3]等用色度值作为特征向量进行 K 均值聚类,将对图像进行分割的任务转化为对数据集合进行分类的任务,从背景中能有效地识别出叶片上的叶螨害虫,平均识别率达 93.95%。但 K 均值采用的是迭代的方法,只能得到局部最优解,且对于噪音和异常点比较敏感。因此许多学者将 K 均值算法与其他算法结合进行改进。王雪光等^[4]在 Lab 颜色空间中采用 K 均值聚类算法分割草莓图像;徐黎明等^[5]在 Lab 空间下结合同态滤波和 K 均值算法对彩色杨梅图像进行分割;赵文昌等^[6]利用融合改进人工蜂群和 K 均值聚类的算法分割图像;以及刘智杭^[7]等提出一种融合局部异常因子(LOF)算法和 K 均值聚类算法的图像分割算法对葡萄果穗图像进行分割。以上学者的算法在某些特定领域应用其效果较为理想,但在不同的场景下应用不能保证其算法的有效性。

针对油菜花图像分割问题,在本文中,笔者提出一种 Lab 空间 K -means 聚类与 HSV 空间颜色分割相结合的算法。采用颜色分割算法来弥补 K -means 聚类算法在光照条件下的过分割问题,首先将油菜花图像转换到 Lab 空间,取出 Lab 空间的 a、b 分量进行重组,设置 K -means 聚类个数,对重组后的图像进行聚类,用颜色分割算法对聚类后的油菜花图像再次进行分割,剔除过分割的区域,可以有效地分割出油菜花的花形、颜色等特性。为了突显此方法分割油菜图像的效果较好,将会对比 K -means 聚类算法以及 EXG 因子颜色算法^[8]分割油菜图像的效果图。

2 K -means 与颜色阈值相结合的分割方法

2.1 K -means 算法原理^[9-10]

K -means 算法是一种无监督的聚类方法,也是一种准确、高效的目标分割算法。之所以被称为

K -means 是因为它可以自发地将很多样本聚成 k 个不同的类别,每一个类别即为一个簇,并且簇的中心是由簇中所有的点的均值计算得出的。

假设有样本数据集 $D = \{x_1, x_2, \dots, x_m\}$, 为 n 维向量,表示样本数据集中的每一个样本数据, n 为样本数据的属性个数。聚类是将样本数据集 D 中相近的样本数据划分到一个集合内,称其集合为簇,用 C 表示, C 的个数用 k 表示。每个簇都有一个中心点,即簇中所有点的中心(即为质心),用 μ_k 表示。因此,使用 K -means 算法能把 $D = \{x_1, x_2, \dots, x_m\}$ 划分为 $C = \{C_1, C_2, \dots, C_k\}$, 每个划分好的簇中的各点到 μ_k 的距离平方之和为误差平方和,其标准估值差 SSE 为式(1)所示。

$$SSE = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2. \quad (1)$$

若要 K -means 算法到达 $C_1, C_2, \dots, C_k$ 内部的样本数据相近性大,簇与簇直接的样本数据相似性小的效果,即尽可能地减少 SSE 的值。输入为:样本数据集 D , 簇的数量 k , 即 k 个划分好的簇。算法具体流程如下:

(1) 在样本数据集 D 中随机选取 k 个聚类中心,即 $\{\mu_1, \mu_2, \dots, \mu_k\}$ 。

$$C = \{C_1, C_2, \dots, C_k\}.$$

(2) 计算样本数据集 D 中每个样本数据 x_i 到每个聚类中心 μ_i 的距离,即计算公式如(2)所示。

$$d_i = (x_i - \mu_i)^2. \quad (2)$$

(3) 如果 d_i 的距离最小,则将 x_i 标记为簇 C_j 中的样本数据,即有 $C_j = \{x_i\}$ 。

(4) 把所有的样本数据划分到不同的集合之后,计算新的聚类中心,即 C_j 中所有点的平均值 μ'_j , 通过(3)式计算所有点的平均值以及误差平方和。

$$\mu'_j = \frac{1}{|C_j|} \sum_{x \in C_j} X. \quad (3)$$

(5) 比较前后两次误差平方和的差值和设定的阈值,如果大于阈值,则重复步骤(2)~(4)。

(6) 如果误差平方和的变化小于设定的阈值,说明聚类已经完成。

2.2 颜色算法原理

目前常用的颜色空间有两种^[11-12], 一种是面向硬件设备的 RGB 颜色空间,另一种是面向视觉感知的 HSV 颜色空间。除此之外还有 Lab、CMYK、YUV、YCbCr 等颜色空间。在本文中,通过无人机拍摄采集的油菜图像属于 RGB 颜色空间,它由三种基本的原色(红、绿、蓝)组成,我们肉眼所看到的颜色

都可以通过这 3 种颜色混合而成, 在实际应用中可以通过分析其图像颜色特征把目标区域与背景分割出来。Lab 颜色空间弥补 RGB 的不足, 它是一种与设备无关的颜色空间, 也是一种基于生理特征的颜色空间。L 表示亮度、a 和 b 表示两个颜色通道。同 RGB 颜色空间相比, Lab 是一种不常用的色彩空间。RGB 颜色空间不能直接转换到 Lab 颜色空间, 需借助 XYZ 颜色空间把 RGB 颜色空间转换到 XYZ 颜色空间, 之后再吧 XYZ 颜色空间转换到 Lab 颜色空间。

HSV 颜色空间是一种主观空间, H 表示色调、S 表示饱和度、V 表示亮度。相比于 RGB 颜色空间, 在 HSV 空间中分割彩色图像更能直观地表达出图像色彩的明暗、色调以及鲜艳程度, 便于肉眼观察颜色之间的不同。在分割油菜 RGB 图像中, 本文中采用颜色直方图作为油菜花图像的颜色特征, 它反映了图像颜色分布的统计特性, 适用于描述难以自动分割的图像。由于油菜花图像为 RGB 颜色空间, 若要在 HSV 空间准确识别出油菜花图像中的油菜花需把油菜花图像按公式(4)转至 HSV 空间^[13-14]。

$$\begin{aligned} V &= \frac{1}{3}(R + G + B). \\ S &= 1 - \frac{3}{(R + G + B)}[\min(R, G, B)]. \\ H &= \cos^{-1} \left\{ \frac{[(R - G) + (R - B)]/2}{(R - G)^2 + (R - B)(R - G)^{\frac{1}{2}}} \right\} / 360. \end{aligned} \quad (4)$$

待图像转至 HSV 空间之后, 提取图像的 H、S、V 颜色直方图并分析其颜色特征。通过确定 H、S、V 三个分量的范围来准确判断某一像素点的颜色。假设 J 为图像数据的某一点像素, J_H 为此像素点的 H 分量、 J_S 为此像素点的 S 分量、 J_V 为此像素点的 V 分量。

则当:

$$\begin{cases} H_{\min} < J_H < H_{\max} \\ S_{\min} < J_S < S_{\max} \\ V_{\min} < J_V < V_{\max} \end{cases} \quad (5)$$

令:

$$\begin{cases} J = J_H \\ J = J_S \\ J = J_V \end{cases} \quad (6)$$

否则:

$$J = 0.$$

上述式(5)中, J_H 的取值范围为 0 到 360° , J_S 以及 J_V 的取值范围为 0 到 255。固定式(5)中 J_S 、

J_V 的区间, 采用二分法调动 J_H 的区间范围, 选取最适宜的 J_H 取值范围使得图像分割的效果最佳, 类推可以确定另外两个分量的取值范围。通过大量实验验证得像素点 J 的约束条件如式(7)。

$$\begin{cases} 26 < J_H < 34. \\ 43 < J_S < 255. \\ 46 < J_V < 255. \end{cases} \quad (7)$$

确定图像在 HSV 空间中每个分量的阈值范围遍历图像中所有像素点, 筛选符合条件的像素点即可把图像中感兴趣的目标区域分割出来并将背景设置为黑色($L=0$)。

2.3 算法描述

无人机拍摄的油菜花图像中主要由油菜绿叶、油菜花以及其它背景等组成, 我们最终的目的是要把油菜花从油菜图像中完整地分割出来。由于 K-means 聚类算法不能适应不同光照下的油菜图像分割, 还需将聚类后的油菜花图像再次经过颜色算法处理准确分割出油菜花。具体算法步骤如下:

(1) 读取一幅 RGB 颜色空间的油菜图像, 记为 I 。

(2) 设置格式创建一个颜色转换结构, 把 I 转换至 Lab 空间, Lab 空间的油菜图像记为 J 。提取 J 的 a 和 b 分量并设置其行数与列数, 使用 reshape 函数对 J 的 a、b 通道进行重组处理。

(3) 对重组后的 J 进行聚类, 设置聚类个数为 3, 迭代次数为 50 次, 初始化聚类中心, 并将不同类别进行 labels 标记, 不断更新聚类中心, 当聚类中心不再发生改变或是达到最大迭代次数时完成聚类。定义一个数组存储聚类结果并显示其聚类效果图。

(4) 把聚类之后得到的油菜花图像转至 HSV 空间, 另外创建一个空白图像 Y , 也转至 HSV 空间。

(5) 提取油菜花图像的颜色直方, 分析其颜色特性并确定 H、S、V 分量的阈值范围。再把满足(7)式要求的像素点数据选取出来并转到空白图像 Y 中。

(6) 将步骤(5)中所得图像 Y 映射到 RGB 彩色空间展示, 以此来保证油菜花在花色、花形上的完整性。

2.4 分割后图像评价指标

不失一般性, 本文中定义误分率 SE 作为对分割效果进行定量分析的标准, 设 N_1 为算法分割得到的图像像素值即目标像素值, N_2 为分割目标的实际像素值(通过 PS 获得)。则 SE 可定义为式(8)。

$$SE = \left| \frac{N_1 - N_2}{N_2} \right| \times 100\% . \quad (8)$$

式(8)中, SE 越小时表明其算法的分割效果越好, 反之则差。

3 实验结果与分析

3.1 图像采集及试验展示

本文中所有测试图像均为中国农业科学院油料作物研究所种质资源研究室采用无人机设备在规模化种植的油菜田中航拍所得。图像采集设备为大疆

公司的无人机, 型号为 WM331A, 图像格式为 jpg, 拍摄距离为离地面 5—10 m。数据采集时间为 2018 年 3 月 9 号下午 2:00 到 3:30 之间。采集油菜图像的当天是晴天, 采集了不同光照强度下的油菜图片。笔者对 60 幅不同的油菜图像进行实验, 选取其中的三幅油菜图像展示, 如图 1 所示。油菜图像在 Win10 系统下的 matlab R2016a 软件上实现算法仿真, 如图 2—图 4 所示(图 3 中标记圈为油菜绿叶或杂草非油菜花区域, 即分割误差区域)。



图 1 原始图像



图 2 EXG 算法效果图

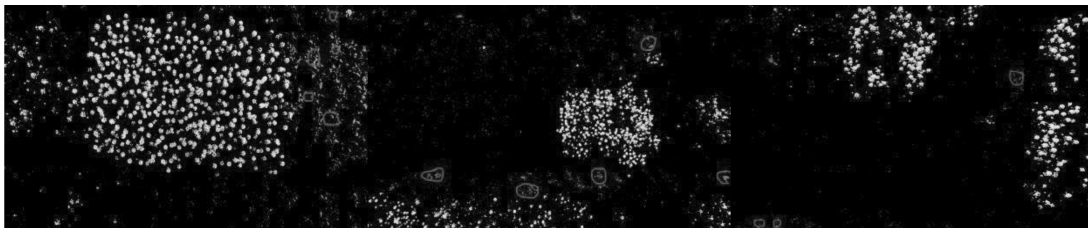


图 3 K-means 算法效果图



图 4 本文算法效果图

3.2 实验结果分析

对比图 1—图 4, 从图 2 中得出用 EXG 颜色算法分割油菜图像虽能大体上完成油菜图像分割, 但油菜花的有些部分没有显示出来且花色较暗从而产

生欠分割的现象。从图 3 中得出采用 K-means 算法分割油菜图像能完整地分割出油菜花且还能保证花色的亮度, 但是由于光照映射到油菜绿叶以及杂草上, 将其误认为是油菜花分割出来, 从而产生过分割

的现象。从图 4 中得出采用本文算法能够很好地分割出油菜花,还能保证油菜花在花形、花色上的完整性,其效果优于前两种算法。

3.3 实验结果指标评价

根据公式(7)计算出三种算法的分割误差率如表 1 所示。

表 1 三种算法的分割误差率

序号	EXG 算法分割 SE	K-means 算法分割 SE	本文算法分割 SE
1	0.241	0.277	0.043
2	0.337	0.349	0.039
3	0.400	0.138	0.046
4	0.337	0.230	0.060
5	0.267	0.263	0.062
6	0.080	0.332	0.043
7	0.097	0.353	0.027
8	0.321	0.305	0.074
均值	0.260	0.281	0.049

从表 1 中可看出本文分割算法的分割误差率 SE 的值最小,其均值也是最小。即采用本文算法分割油菜图像其效果比其它两种算法好。

4 结论

在本文中,笔者对田间规模化种植的油菜实现了精准的油菜花分割,对于实现高通量精准、鉴定油菜群体的开花期具有广阔的应用前景。在自然光照条件下采集的大田油菜航拍图像,受光照影响颇大,由于田间作物开花程度差异造成采光程度不同,在进行油菜花像素提取时,传统颜色索引算法对背景噪声过于敏感,而特定阈值分割方法处理范围单一,不适合处理色彩丰富,光照强度大的图像。基于此笔者得到以下结论:

(1)EXG 因子颜色分割算法虽方法较为简单、计算量小、效率较高,但对噪声较敏感,可能会导致区域内有空洞。另外当目标较大时分割速度较慢,因此不适于航拍油菜花图像的分割。

(2)使用经典 K-means 算法对油菜花图像进行聚类分割,虽然能够很好地将油菜花与油菜的杆径分离,聚类区域得到完整的油菜花,但同时油菜花的旁边也会出现很多黄色块(非油菜花),即过分割现象。若光照太强,此分割方法也会出现欠分割现象。

(3)基于 Lab 空间利用 K-means 与 HSV 空间颜

色分割相结合的算法能够较为理想地从油菜花图像中分割出油菜花,有效地解决了因光照影响而带来的欠分割现象,同时也避免了过分割的问题,适合解决复杂背景下的油菜图像分割问题。此次大田油菜实验结果能够达到研究后期研究油菜花期识别对图像预处理的要求。

参考文献:

- [1] 胡代玉,付茂宁. 中国油菜生产技术供需的经济学分析[J]. 农业与技术,2018,38(12):80.
- [2] 李辉,滕桂法,张燕. 数字图像处理技术及其在农业中的应用研究[J]. 安徽农业科学,2013,41(13):6060-6061.
- [3] 刘国成,张杨,黄建华. 基于 K-means 聚类算法的叶螨图像分割与识别[J]. 昆虫学报,2015,58(12):1338-1343.
- [4] 王雪光,陈淑红. 基于 K 均值聚类的成熟草莓图像分割算法[J]. 农机化研究,2013(6):51-54.
- [5] 徐黎明,吕继东. 基于同态滤波和 K 均值聚类算法的杨梅图像分割[J]. 农业工程学报,2015,31(14):202-208.
- [6] 赵文昌,李忠木. 融合改进人工蜂群和 K 均值聚类的图像分割[J]. 液晶与显示,2017(9).
- [7] 刘智杭,于鸣,任洪娥. 基于改进 K 均值聚类的葡萄果穗图像分割[J]. 江苏农业科学,2018,46(24):239-244.
- [8] 王雪,尹来武,郭鑫鑫. 室外多变光照条件下农田绿色作物的图像分割方法[J]. 吉林大学学报(理学版),2018,56(05):179-184.
- [9] Sima H, Liu L, Guo P. Color Image Segmentation Based on Blocks Clustering and Region Growing [C]// International Conference on Neural Information Processing. Springer, Berlin, Heidelberg, 2011.
- [10] 王丹丹,徐越,宋怀波,et al. 融合 K-means 与 Ncut 算法的无遮挡双重叠苹果目标分割与重建[J]. 农业工程学报,2015,31(10):227-234.
- [11] 帅杜娟. 基于图像分割的大田油菜花期识别研究. [D]. 武汉:武汉轻工大学,2019.
- [12] 李伟斌,马洪林,易贤,等. 基于色彩空间变换的彩色图像分割方法[J]. 计算机工程与应用,2019,55(9):162-167.

(下转第 21 页)

图3所示的模拟结果可见,通过川北方程建立的数学模型与实验数据是很接近的,可见用川北方程来拟合是比较可靠的。

4 结论

本文通过对油桐籽进行压榨实验,揭示了其压榨过程中的变形规律。结果表明:

(1)油桐籽的压榨过程加载速率对油桐籽压榨过程中的应力-应变关系有较大影响,较小的加载速率在相同的压榨应力下所对应的应变更大。

(2)总体上加载速率越小,同一应力水平下应力下的应变越大,出油率越高。

(3)基于川北方程建立了油桐籽在侧限排油的一维压榨应力-应变的数学模型,结果显示,它能够较好地建模脱壳油桐籽压榨过程中的应力-应变关系。

参考文献:

- [1] 杨焰,廖有为,谭晓风. 我国油桐产业与未来环保型涂料产业协同发展之探讨[J]. 经济林研究,2018,36(04):188-192.
- [2] 苏亮. 一种桐油的制备方法:CN201310622254.2[P]. 2015-06-03.
- [3] 贵州聚恒元化工有限公司. 一种桐油制备方法:CN201711036483.0[P]. 2018-01-19.
- [4] 刘汝宽,肖志红,李昌珠,王红建,张爱华,吴红. 油桐籽破壳力学特性及主要影响因素研究[J]. 中国油脂,2018,43(09):132-134.
- [5] 郑晓,林国祥,王少梅. 油料压榨的应力-应变关系研究[J]. 中国油脂,2005(10):37-39.
- [6] 郑晓,李智,林国祥,王少梅. 菜籽与脱皮菜籽压榨的应力-应变关系研究[J]. 农机化研究,2004(06):187-189+192.

(上接第17页)

- [13] 张凯兵,章爱群,李春生. 基于HSV空间颜色直方图的油菜叶片缺素诊断[J]. 农业工程学报,2016,32(19):179-187.
- [14] LIU Huan-xi, ZHU Tian-hong. Infrared and

Visible Image Fusion Based on Region of Interest Detection and Nonsampled Contourlet Transform[J]. Journal of Shanghai Jiaotong University(Science),2013,18(5):526-534.