

基于改进随机 Hough 变换的椭圆形农产品大小分级

符子祥 崔艳*

(天津工业大学 理学院,天津 300387)

摘要 针对椭圆形农产品的分级问题,采用最近邻分类算法和随机 Hough 变换理论,对哈密瓜这类椭圆形农产品的大小分级方法进行研究。结果表明:1)哈密瓜边缘轮廓近似椭圆形,所测出的长短轴半径,可以作为椭圆形哈密瓜大小分级的新标准;2)通过试验测定,对于白色背景的哈密瓜,最近邻分类算法可以提取出较为完整的边缘轮廓;3)随机 Hough 变换可以在边缘轮廓不完整且有随机噪声干扰的情况下,检测出任意曲率的哈密瓜边缘轮廓的近似椭圆;4)与椭圆形哈密瓜的半径的真实值相比,本改进算法识别值的相对误差小于6%;5)当哈密瓜处在不同倾斜状态时,如 0° 、 45° 、 90° 、 135° ,本改进算法仍可以准确测得其长短轴半径。本改进算法还可以推广到其他椭圆形和类椭圆形农产品的大小分级中。

关键词 最近邻分类;随机霍夫变换;大小分级;椭圆形

中图分类号 TN957.52

文章编号 1007-4333(2017)10-0133-06

文献标志码 A

Size classification of elliptical agricultural products based on improved Random Hough Transform

FU Zixiang, CUI Yan*

(School of Science, Tianjin Polytechnic University, Tianjin 300387, China)

Abstract In order to solve the classification problem of elliptical agricultural products, Nearest Neighbor classification algorithm and Random Hough Transform theory are used to study the size classifying method of ellipse agricultural products, such as cantaloupe. The results show that: 1) The radius of the long and short axis measured by the approximate ellipse of the cantaloupe edge contour can be used as a new standard for the size classification of the elliptical cantaloupe; 2) Through experiment testing, for the cantaloupe with white background, the Nearest Neighbor classification algorithm can extract the almost complete edge contour; 3) Random Hough Transform can detect the approximate ellipse of elliptical cantaloupe edge contour with arbitrary curvature under incomplete edge contour and random noise interference; 4) Compared with the true value of the elliptical shape agricultural product, the relative error of the improved algorithm is less than 6%; 5) When the cantaloupe is in different inclination states (such as 0° , 45° , 90° and 135°), the improved algorithm still can accurately measure its long and short axis radius. The improved algorithm can also be extended for the size classification of other elliptical and approximately oval agricultural products.

Keywords Nearest Neighbor Clustering; Random Hough Transform; size grading; ellipse

在水果等农产品的分级中,除品质^[1]外,大小是一个重要的指标。目前针对水果大小的分级主要有机械式^[2]和机器视觉类2种方法。机械式的分级是

使水果逐一通过由大到小的网格,然后依次筛选出不同大小的水果。这种分级方式,在分级时易造成水果的损伤。而机器视觉类的分级则可以避免此类

收稿日期:2016-09-13

基金项目:天津市应用基础及前沿技术研究计划重点项目(12JCZDJC27800)

第一作者:符子祥,硕士研究生,E-mail:fzxsxgz2016@163.com

通讯作者:崔艳,副教授,主要从事图像算法研究,E-mail:cuiyan@tjpu.edu.cn

问题的发生。

对于机器视觉类的水果大小分级,目前提出许多方法。已有研究大致可以归为2类。一类是从面积角度来分析,其分级依据是:1)横截面积的大小^[3];2)由边缘长度得到的面积^[4];3)投影面积^[5]。另一类是从边缘的角度进行分析,其分级依据是:1)边缘点到质心的平均距离^[6];2)边缘的最大纵径^[7];3)最小二乘法拟合的边缘曲率^[8];4)边缘外接矩形的长和宽^[9];5)不同颜色下边缘轮廓的最大值^[10]。

现有大小分级方法,没有从形状的角度考虑,其分类标准也不是以直接测量农产品得到的实际尺寸为依据;而是从检测目标的实际尺寸出发建立相应的参考量,间接地给出了一种分级标准。其实质仅是对大小给予较客观的估计和比较;而且利用面积和求平均的方法也不适合对较大长宽比的农产品进行分级。

本研究拟以椭圆形的哈密瓜为例,针对新皇后、乙型新密这2个品种,通过最近邻算法提取其边缘轮廓,随机 Hough 变换提取其边缘的近似椭圆,探索一种基于农产品轮廓外形的大小分级方法。

1 算法描述

1.1 哈密瓜形状提取

1.1.1 最近邻分类算法

最近邻(Nearest Neighbor, NN)分类是一种特殊的 K 最近邻(k NN, k -Nearest Neighbor)分类算法。目前,最近邻分类已经在文本分类、模式识别等领域得到广泛的应用^[11-13]。

设存在集合 $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_d, y_d)\}$ 有 d 个样本,将其划分成 L 个类。设有另一个集合 $E = \{(\omega_1, z_1), (\omega_2, z_2), \dots, (\omega_e, z_e)\}$ 有 e 个样本。对于 E 中的每1个分量 (ω_j, z_j) ,计算其与 D 中每1个分量的欧式距离,找到 D 中与 (ω_j, z_j) 距离最小(或为某个设定阈值)的 (x_i, y_i) 。此时, (x_i, y_i) 对应的分类就是 (ω_j, z_j) 的分类,即:

$$G_i = \sqrt{(x_i - \omega_j)^2 + (y_i - z_j)^2} \quad (1)$$

式中: G_i 为 D 中第 i 个样本 (x_i, y_i) 与 (ω_j, z_j) 的欧式距离。当 G_i 取得最小值时, (x_i, y_i) 就是 (ω_j, z_j) 的最近邻。 (ω_j, z_j) 属于 (x_i, y_i) 所在的那个分类。

NN 算法的优点在于只需在分类之前提供一个训练的样本即可,并可以避免参数过拟合的发生。

在运用 NN 算法提取新皇后、乙型新密这2个品种的哈密瓜时,均取得满意的效果。本研究使用 NN 算法使得哈密瓜与背景有效的分离,缩小了随机 Hough 变换检测的范围,避免了大量无效计算,提高了检测的效率。

1.1.2 最近邻分类提取哈密瓜轮廓的步骤

1)分类训练样本的获得。在白色背景下拍摄图像,获取部分图像的颜色值,并计算出这些颜色转化到 Lab 空间后的平均值。每种颜色对应 Lab 空间中的2个颜色分量,将这些颜色分成6类,这就是训练样本 D 。

2)颜色信息分类。首先,将待处理图像的所有颜色值都转化到 Lab 空间。然后,采用 NN 算法对其进行分类,即:计算出图像中每个颜色值与6类训练样本的欧式距离。欧式距离最小的颜色分类就是该点对应的分类。

3)边缘提取。根据分类结果,提取其边缘轮廓。采用随机 Hough 变换提取边缘的近似椭圆。图1示出哈密瓜原图和 NN 算法提取的轮廓以及随机 Hough 变换检测结果。

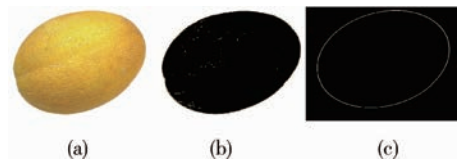


图1 哈密瓜原图(a)和 NN 算法提取的轮廓(b)以及随机 Hough 变换检测结果(c)

Fig.1 Cantaloupe original picture (a), NN extraction outline (b) and Random Hough Transform detection result (c)

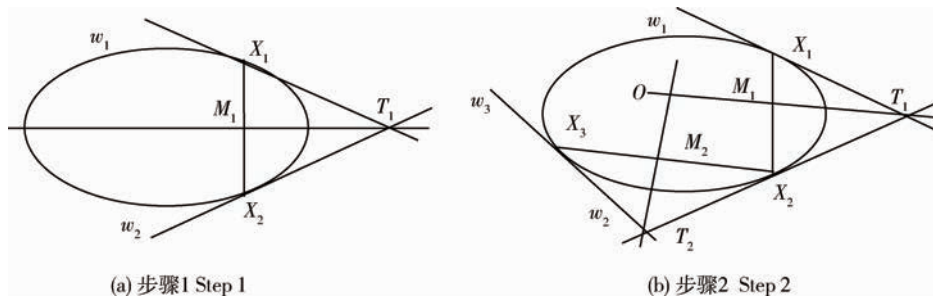
1.2 哈密瓜长短轴半径的检测

1.2.1 随机 Hough 变换

随机 Hough 变换理论具有对局部残缺不敏感和对随机噪声鲁棒性好等突出优点,可以有效地提取几何图形及其特征,在农业技术上的应用十分广泛^[14-16]。本研究通过随机 Hough 变换检测出哈密瓜轮廓边缘的近似椭圆,以获取的椭圆形轮廓的长、短轴半径作为分级依据。

1.2.2 随机 Hough 变换检测的步骤

1)随机取点和确定椭圆中心。在得到的二值化边缘图上,按照图2所示随机取点。为提高检出椭圆的成功率,在随机取点时为避免无效取点和无效



(a) 步骤1 Step 1

(b) 步骤2 Step 2

步骤 1: 在图像上随机选取 2 点 X_1, X_2 , 根据其梯度得到对应的切线 w_1, w_2 及切线的交点 T_1 ; 计算出线段 X_1X_2 的中点 M_1 。根据椭圆的性质, 直线 T_1M_1 一定经过椭圆中心。步骤 2: 另选 1 个点 X_3 , 重复前面的过程, 得到另一条过椭圆中心的直线 T_2M_2 , T_1M_1 和 T_2M_2 的交点就是椭圆中心。

Step 1: Select two points X_1, X_2 randomly on the image, and use their gradient to obtain the corresponding tangent line w_1, w_2 and the intersection T_1 of them. Then obtain the midpoint M_1 of the segment X_1X_2 . According to the nature of ellipse, line T_1M_1 must pass through the center of the ellipse. Step 2: Reselect a point X_3 to repeat the previous process. Acquiring another line through the center of the ellipse, T_2M_2 , T_1M_1 and T_2M_2 intersection is the center.

图 2 椭圆中心的确定过程

Fig. 2 Process of determining the center of the ellipse

计算, 限定点与点的最大和最小距离, 并且限定所得椭圆的最小面积。如果找到的点没有满足限定条件就重新选取。在满足限定条件之后, 确定椭圆中心。

2) 判断 3 点是否在椭圆上。得到椭圆中心之后, 将椭圆中心移动到坐标原点 $(0, 0)$ 处。根据 X_1, X_2, X_3 3 个点的坐标和式 (2), 计算出椭圆的 3 个参数 a, b, c 。如果式 (2) 无解, 返回第一步。如果有解, 分析是否满足椭圆判断条件: $(ac - b^2) > 0$ 。如果满足判断条件, 那么就检测到 1 个椭圆。

$$ax^2 + 2bxy + cy^2 = 1 \quad (2)$$

3) 转换椭圆方程到新的参数空间。由已知的参数 p, q, a, b, c 得到 Hough 变换空间的新参数 p, q, r_1, r_2 和 θ 。新参数满足式 (3)。

$$\frac{[(y - q)\sin\theta + (x - p)\cos\theta]^2}{r_1^2} + \frac{[(y - q)\cos\theta - (x - p)\sin\theta]^2}{r_2^2} = 1 \quad (3)$$

式中: r_1, r_2 分别为椭圆长、短轴半径 (长、短轴长度的一半); (p, q) 为椭圆中心的坐标; θ 为偏离长轴的角度。

4) 筛选满足条件的椭圆并获得长短轴半径。每当得到 1 组新的参数 (p, q, r_1, r_2 和 θ) 时, 就判断累加器中是否有这样 1 组参数。如果检测出相同的参数或者差别很小的参数都将其归为一组, 对应的数量加 1。如果没有这组参数, 就新增 1 个空间存放它, 数量记为 1。重复上面 1)~3) 的过程, 直到获得足够的参数。检测结束后, 数量最多的椭圆就是所要检测的边缘。输出该椭圆的长短轴半径。

2 结果与分析

2.1 不同旋转角度下的检测结果

本研究所选取的哈密瓜品种为新皇后和乙型新密。这 2 个品种哈密瓜的特征是: 通体金黄或绿色, 带有皮纹, 外形呈椭圆形。图 3 示出哈密瓜摆放在不同倾斜角度下的检测结果。本研究有效地检测出哈密瓜的边缘轮廓, 同时得到其长短轴半径。该试验结果表明: 本研究的改进算法识别能力强, 鲁棒性好。

2.2 测量结果

本研究选取了 12 组哈密瓜样本, 分别测得其外形近似椭圆的真实长、短轴半径 (长短轴真实值长度的一半)。采用真实值与本研究的识别结果进行比较。结果表明, 其相对误差小于 6%, 获得了满意的结果 (表 1)。相对误差的计算公式为:

$$\varphi = \frac{|r - R|}{r} \times 100\% \quad (4)$$

式中: r 为哈密瓜真实长轴 (或短轴) 半径; R 为本算法测得的长轴 (或短轴) 半径; φ 为相对误差。

为比较该改进算法, 设置一个对比算法。对比算法为: 测量边缘上的最大纵径作为长轴长度, 垂直这一方向的 2 个边缘点距离作为短轴长度。将测得的数值除以 2 与本研究识别值进行比较。图 4 示出哈密瓜半径的实际测量值以及本算法和对比算法的识别值。可以看出: 相较于对比算法, 本算法的识别值与实际测量值相差较小。

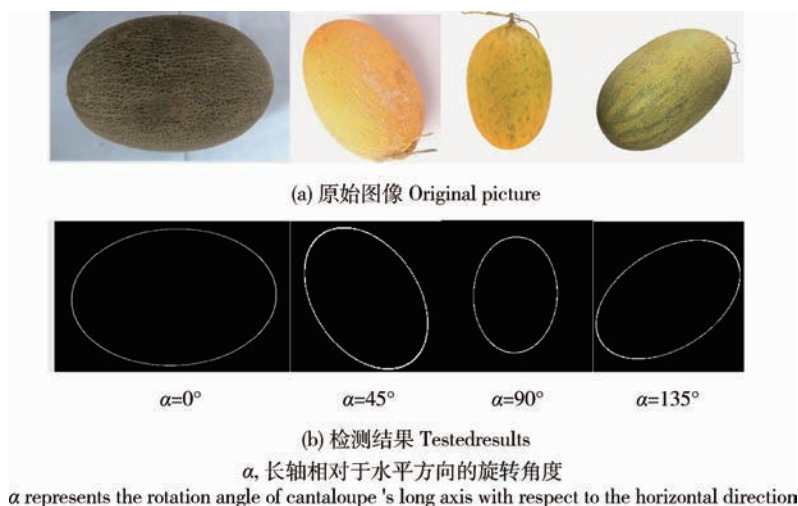


图3 哈密瓜不同旋转角度下的检测结果

Fig. 3 Cantaloupe tested results under different rotations

表1 哈密瓜半径实际测量值与识别值比较

Table 1 Cantaloupe actual measured size compared to the measured size

试验组数 Number	长轴半径 Real long axis			短轴半径 Real short axis		
	实际测量值 r_1 /像素 Actual measured size	识别值 R_1 /像素 Obtained size	相对误差 φ_1 / % Relative error rate	实际测量值 r_2 /像素 Actual measured size	识别值 R_2 /像素 Obtained size	相对误差 φ_2 / % Relative error rate
1	235.53	243.97	3.6	171.50	176.46	2.9
2	145.54	146.55	0.7	90.72	92.94	2.4
3	220.01	221.48	0.7	136.56	140.36	2.8
4	219.68	226.31	3.0	138.09	142.68	3.3
5	88.62	93.55	5.6	66.53	69.62	4.6
6	100.50	95.94	4.5	57.51	56.42	1.9
7	163.01	164.26	0.8	109.50	111.34	1.7
8	85.56	88.89	3.9	59.44	60.96	2.6
9	162.72	162.44	0.2	107.67	111.01	3.1
10	109.60	112.78	2.9	68.39	70.12	2.5
11	161.78	164.73	1.8	110.93	110.29	0.6
12	232.22	235.74	1.5	108.70	111.75	2.8

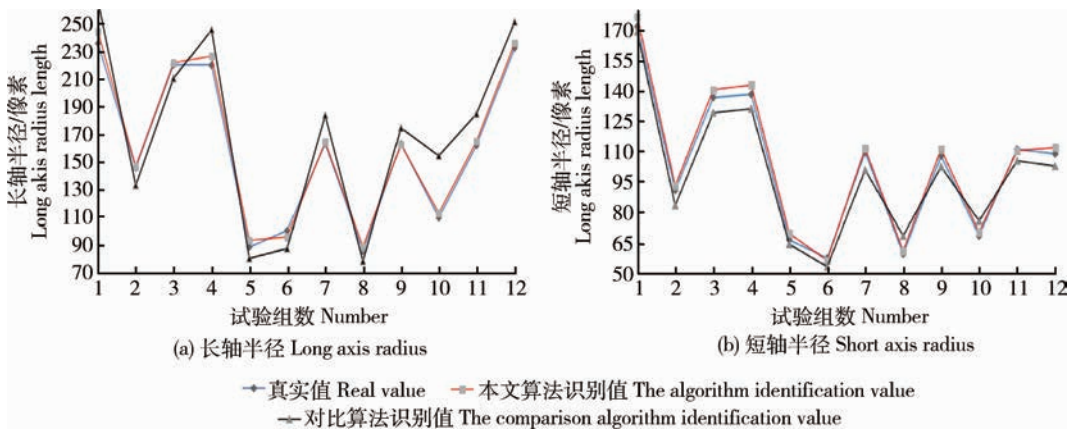


图 4 哈密瓜半径实际测量和 2 种识别值的对比

Fig. 4 Comparison between cantaloupe radius' actual measured value and two kinds of identified values

目前针对哈密瓜大小分级,国家没有制定相应的标准。本研究根据以上检测结果,提出了哈密瓜大小的分级标准(表 2)。据此标准,12 组测试样本的分级结果为:一级 4 个,二级 3 个,三级 5 个。

表 2 哈密瓜大小分级标准

Table 2 Cantaloupe size grading standards

	像素 pixel	
分级 Classification	长轴半径 Long axis radius	短轴半径 Short axis radius
一级 High quality	≥200	≥105
二级 Good quality	>150	>100
三级 Ordinary quality	≤150	≤100

3 结 论

以椭圆形的哈密瓜为例,本研究提出了一种基于边缘轮廓的大小分级方法。试验结果表明:通过 NN 算法提取边缘,以及随机 Hough 变换取点时对距离和范围的限制,极大地提高了随机 Hough 变换的检测速度和效率。当哈密瓜边缘不完整且倾斜状态各异时,本研究仍然能将任意曲率的椭圆长短轴半径检测出来,并且识别值的相对误差 < 6%。考虑到椭圆的提取本身就具有一定的难度,哈密瓜表面又有纹理,因此在提取检测时的情况较复杂。而冬瓜等农产品的表面并无纹理,橘子、苹果等农产品的外形呈圆形(圆比椭圆分析起来要简单),本算法完全可以推广到其他椭圆形和类椭圆形农产品的大小分级中。

参考文献 References

[1] 胡发焕,董增文,匡以顺. 基于机器视觉的脐橙品质在线分级检测系统[J]. 中国农业大学学报, 2016, 21(3): 112-118
 Hu F H, Dong Z W, Kuang Y S. Orange online grading detection system based on machine vision [J]. *Journal of China Agricultural University*, 2016, 21 (3): 112-118 (in Chinese)

[2] 杨松夏,陆华忠,吕恩利,许锦锋. 球形水果机械式大小分级技术研究现状[J]. 农机化研究, 2012(8): 235-239
 Yang S X, Lu H Z, LU E L, Xu J F. Research status on spherical fruits mechanical grader[J]. *Journal of Agricultural Mechanization Research*, 2012(8): 235-239 (in Chinese)

[3] 莫亚子,段佳欢,沈斌,张梦珂,周敬. 基于面积的猕猴桃大小分级检测算法[J]. 电子世界, 2014(10): 257-258
 Mo Y Z, Duan J H, Shen B, Zhang M G, Zhou J. Grading algorithm of kiwifruit size based on area [J]. *Electronics World*, 2014(10): 257-258 (in Chinese)

[4] 姚娜,吴刚,陈杰. 基于图像边缘检测的红枣大小分级[J]. 湖北农业科学, 2014, 53(10): 2427-2430
 Yao N, Wu G, Chen J. Grading red dates based on the size from image edge detection[J]. *Hubei Agricultural Sciences*, 2014, 53(10): 2427-2430 (in Chinese)

[5] 陈英,廖涛,林初靠,万虎,李伟. 基于计算机视觉的葡萄检测分级系统[J]. 农业机械学报, 2010, 41(3): 169-172
 Chen Y, Liao T, Lin C K, Wan H, Li W. Grape detection classification system based on computer vision [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2010, 41(3): 169-172 (in Chinese)

[6] 杜永忠,平雪良,何佳唯. 圣女果表面缺陷检测与分级系统研究[J]. 农业机械学报, 2013, 44(S1): 194-199
 Du Y Z, Ping X L, He J W. Surface defect detection and classification system for cherry tomatoes[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2013, 44

- (S1):194-199 (in Chinese)
- [7] 施健,何建国,张冬,马翔. 基于计算机视觉鲜枣大小分级系统研究[J]. 食品与机械, 2013, 29(5):134-137
Shi J, He J G, Zhang D, Ma X. Study on jujube size-grading system based on computer vision[J]. *Food & Machinery*, 2013, 29(5):134-137 (in Chinese)
- [8] 刘启全,郭俊先,刘明涛,刘航,王艳. 基于机器视觉的哈密瓜大小分级研究[J]. 新疆农机化, 2016(2):16-18
Liu Q Q, Guo J X, Liu M T, Liu H, Wang Y. Cantaloupe size grading based on machine vision[J]. *Xinjiang Agricultural Mechanization*, 2016(2):16-18 (in Chinese)
- [9] 周竹,黄懿,李小昱,文东东,汪成龙,陶海龙. 基于机器视觉的马铃薯自动分级方法[J]. 农业工程学报, 2012, 28(7):178-183
Zhou Z, Huang Y, Li X Y, Wen D D, Wang C L, Tao H L. Automatic detecting and grading method of potatoes based on machine vision[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2012, 28(7):178-183 (in Chinese)
- [10] Mhaski R R, Chopade P B, Dale M P. Determination of ripeness and grading of tomato using image analysis on raspberry Pi [C]. In: 2015 Communication, Control and Intelligent Systems (CCIS), Mathura, 2015:214-220
- [11] Khine T T Y, Araki K, Mitsunaga D, Zhao H A. A Nearest Neighbor search algorithm for LR-LD on high SNR[C]. In: 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS), Okayama, Japan, 2016:1-5
- [12] Touazi A, Bouchaffra D. A k-Nearest Neighbor approach to improve change detection from remote sensing: Application to optical aerial images[C]. In: 15th International Conference on Intelligent Systems Design and Applications (ISDA), Marrakech, 2015:98-103
- [13] Gunale K G, Mukherji P. Fall detection using k-Nearest Neighbor classification for patient monitoring [C]. In: 2015 International Conference on Information Processing, Pune, 2015:520-524
- [14] 张卫,杜尚丰. 机器视觉对农田中定位基准线的识别[J]. 中国农业大学学报, 2006, 11(4):75-77, 88
Zhang W, Du S P. Machine vision recognizing position baseline in cropland[J]. *Journal of China Agricultural University*, 2006, 11(4):75-77, 88 (in Chinese)
- [15] 丁幼春,王书茂. 基于 RHT 的多圆检测改进算法[J]. 中国农业大学学报, 2008, 13(4):121-125
Ding Y C, Wang S M. An improved arithmetic for multi-circle detection based on randomized Hough transformation [J]. *Journal of China Agricultural University*, 2008, 13(4):121-125 (in Chinese)
- [16] 黄士凯,祁力钧,张建华,王俊. 基于行宽的玉米行间杂草识别算法[J]. 中国农业大学学报, 2013, 18(1):165-171
Huang S K, Qi L J, Zhang J H, Wang J. An identification algorithm of weeds among multi-row corn based on the mapping of the corn row's width [J]. *Journal of China Agricultural University*, 2013, 18(1):165-171 (in Chinese)

责任编辑:刘迎春