

基于机器视觉的番茄成熟度颜色判别

Tomato maturity color discrimination based on machine vision

毕智健¹ 张若宇¹ 齐妍杰² 吴业北²

BI Zhi-jian¹ ZHANG Ruo-yu¹ QI Yan-jie² WU Ye-bei²

(1. 石河子大学机械电气工程学院, 新疆 石河子 832000; 2. 新疆兵团农业机械重点实验室, 新疆 石河子 833200)

(1. Mechanical and Electrical Engineering College, Shihezi University, Shihezi, Xinjiang 832000, China;

2. The Key Laboratory of Xinjiang Production and Construction Group, Shihezi University, Shihezi, Xinjiang 832000, China)

摘要: 提出一种颜色分析方法用于新鲜番茄分类, 以 GB 8852—88 标准为参考, 定义番茄成熟度的分类标准(在研究中将其分成四类: 完熟、成熟、半熟、绿熟), 将采集到的番茄 RGB 图像, 去除背景后, 滤波去噪, 转换成 HIS 颜色模型和 HSV 颜色模型。通过 Matlab 编程获取 R、G、B、H、S、V、I 各颜色分量的均值, 运用 SPSS 软件进行判别筛选组合特征分量, 运用 Matlab 进行判别分析。分析结果显示, 绿熟番茄在 3 种不同判别函数下训练集与验证集判别率均达到了 100.00%; 半熟番茄训练集判别率最高为 94.74%, 同时验证集判别率最高达到 100%; 成熟番茄训练集与验证集判别率最低, 分别为 76.67% 和 70.00%; 完熟番茄训练集与验证集最高均为 90.00%。总体上实现了不同成熟度番茄的判别分类。

关键词: 机器视觉; 番茄; 成熟度; 颜色分量均值; 判别分析

Abstract: Tomato quality is one of the most important factors ensured the consistency of tomato market factors. A color analysis method was proposed for classifying the fresh tomato, with reference to the national standard GB 8852—88, defining the classification standards of tomato maturity. In this study, tomato was divided into the four categories, full ripe, ripe, half ripe, and green ripe. RGB images of tomato were collected, removing the background and filtering de-noising, and then they were converted to HIS and HSV color models. Through the MATLAB programming, the mean values of the color components R, G, B, H, S, V, and I were obtained, and the determination and selection of combination components were carried on using SPSS software. Moreover, the discriminant analyses were then performed using Matlab. The results showed that the identification

rates of the green ripe and validation sets were the best of all the three different discriminant functions, reached 100.00%, and the highest discrimination rates of half ripe tomato set was 94.74%. However, the identification rates of training and validation sets of ripe tomato were the lowest, identified as 76.67% and 70.00%, respectively. Those of the training and validation sets of full ripe were the highest, about 90.00%. In general, the discrimination and classification of tomato with different maturity were realized using the machine recognition system in the present study.

Keywords: machine vision; tomato; maturity; the mean color component; discriminant analysis

成熟度被看做是与果蔬质量评估密切相关的最重要的因素之一^[1]。而在判定果蔬成熟方面, 表面颜色特征与成熟度有着极其密切的关系, 颜色是一个重要的参数^[2]。

目前机器视觉技术已经在番茄识别领域得到了较为广泛的研究^[3]。张瑞合等^[4]利用图像分割对番茄进行标定, 继而采用面积配准, 采用体视成像进行三维定位。蒋焕煜等^[5]利用形心匹配和区域匹配的方法, 获取番茄位置信息实现识别。赵杰文等^[6]利用 HIS 颜色特征, 分割出成熟番茄区域, 实现番茄的识别。纪平等^[7]利用 Canny 算子对番茄进行轮廓提取, 再用圆对轮廓进行拟合, 实现了番茄的识别。

传统方法判别番茄成熟度不能定量进行, 且受人为主观感觉影响有较大的误差。本研究拟基于机器视觉理论实现对番茄不同成熟阶段进行研究分析, 以颜色判别为目的, 在研究不同成熟阶段的加工番茄颜色特征分量均值的基础上, 获取加工番茄表面颜色特征分量^[8]和组合颜色特征分量, 探求一种新的颜色判别分析和颜色量化方法^[9], 进而实现对不同成熟阶段加工番茄的分类。弥补人工辅助判别加工番茄成熟度的主观缺陷, 为实现加工番茄选择性收获提供一种依据。

1 材料与方法

1.1 加工番茄样本

试验所用加工番茄样本于 2015 年 9 月 11 日采自石河

基金项目: 国家自然科学基金项目(编号: 61565016); 兵团国际合作项目(编号: 2015AH003)

作者简介: 毕智健, 男, 石河子大学在读硕士研究生。

通信作者: 张若宇(1980—), 男, 石河子大学副教授, 硕士生导师。

E-mail: ry248@163.com

收稿日期: 2016—09—05

子阜康农场,加工番茄的品种是金红果 1 号,共采集番茄样本 150 组,用于试验的有效番茄样本共有 109 组。其中,完熟番茄 30 组,成熟番茄 30 组,半熟番茄 19 组,绿熟番茄 30 组。完熟番茄编号为 F1-F30,成熟番茄编号为 M1-M20,半熟番茄编号为 H1-H19,绿熟番茄编号为 G1-G30。每个番茄从花萼区等分成 3 个区域,每个区域采集一幅 RGB 图像,顶部采集一幅 RGB 图像,总共采集 RGB 图像 436 幅。

1.2 机器视觉系统

图像采集系统主要由工业相机 (point gray FL2G-13S2C-C, Point Grey, Canada)、镜头 (HF6M-2, Spacecom CCTV, Japan) 主要获取加工番茄表面的颜色特征信息、LED 面光源等组成。考虑到后续图像数据处理的问题,将采集的图像信息通过图像采集卡 (FirePRO1394b PCI-e, Point Grey, Canada) 传输到计算机,计算机用于试验数据的存储和处理,系统整体架构图见图 1。

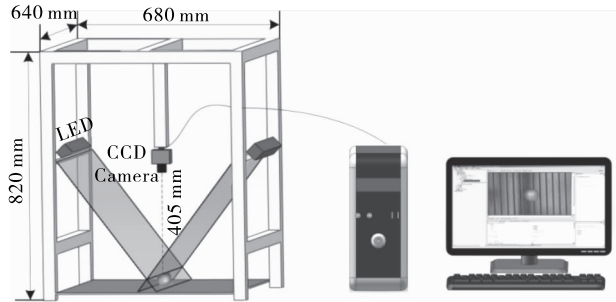


图 1 系统整体架构图

Figure 1 The architecture diagram of entirety system

1.3 颜色模型

颜色模型主要有 HSV、RGB、HSI、CHL、LAB、CMY 等。RGB 模型是数字图像处理的重要模型,计算机利用 RGB 模型表示像素的颜色,其中 $R[0,255], G[0,255], B[0,255]$ 。RGB 的分量与人对颜色的感知并没有直接的联系,这显然不适应人的视觉特点。HIS 颜色空间通常使用非常接近于人类对彩色感知的方法来定义彩色,是一种很直观的方法,用色调 H 、饱和度 S 、光强 I 3 个参数来描述颜色特征, HIS 颜色模型对应于圆柱坐标系的双圆锥子集上。HSV 颜色模型的每一种颜色都是由色相 H 、饱和度 S 、色明度 V 3 个参数来描述颜色特征,其是人类对色彩感知的最好的颜色模型, HSV 颜色模型对应于圆柱坐标系中的一个圆锥形子集。因此,在本研究中选取了 HSV、HIS 两种颜色模型。

1.4 图像处理及获取有用颜色特征分量

运用 Matlab 软件将采集到的 436 幅不同成熟阶段番茄图像进行预处理,减少由于光线、曝光的不均匀、电子器件的误差引起的噪声干扰,提高图像的质量^[10]。如图 2 所示,首先,将原始 RGB 图像转换成灰度图像,灰度图像二值化,并进行数据类型转换,再将原始 RGB 图像进行背景分割,提取 RGB 图像的 3 个单分量图像,然后用得到的 $R、G、B$ 单分量分别点乘二值化图像,利用 cat 函数将原始 RGB 图像进行复原,滤除噪声,便得到去背景之后的 RGB 图像,再提取去除背景后的 $R、G、B$ 分量图像,获取单分量的值。其次,将去背

景的番茄 RGB 图像转换到 HSV、HIS 颜色模型下,提取 $H、S、V、I$ 各单分量图像,获取单分量的值,然后选取合适的单分量进行自由组合^[11],运用 SPSS 软件进行判别筛选,最终筛选出来的有价值的组合特征分量有: $(R-G)/(G-B)、(R+G)/(G-B)、(R-G)/(R+G)、G/B、R/B、B/R$ 。这些特征分量被作为线性判别函数的每一项,用来判别不同成熟阶段的番茄样本。不同成熟度番茄样本图像去背景后灰度图对比见图 3。

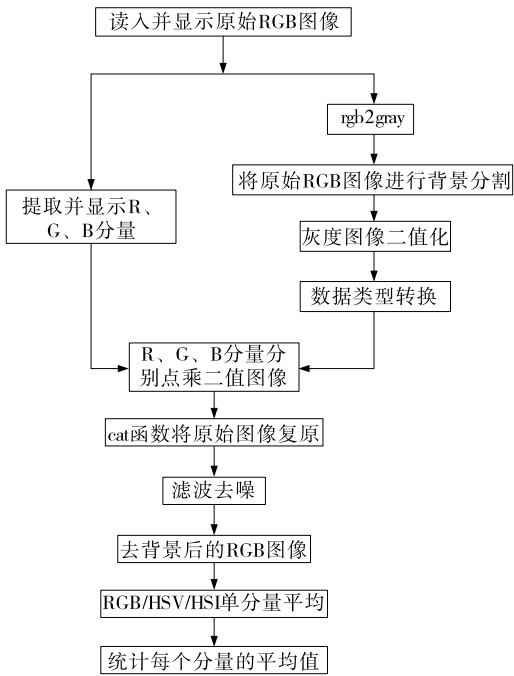


图 2 图像处理流程图

Figure 2 The flow chart of image processing

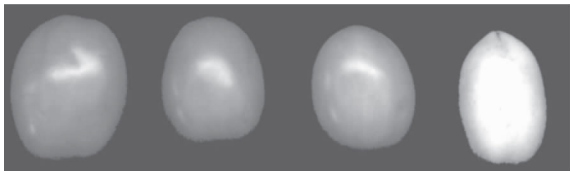


图 3 不同成熟度番茄样本图像去背景后灰度图对比

Figure 3 The grayscale contrast of image background removal of different maturity tomato sample

1.5 数据分析方法

对 SPSS 筛选出来的组合特征分量 $(R-G)/(G-B)、(R+G)/(G-B)、(R-G)/(R+G)、G/B、R/B、B/R$, 包括单分量特征分量 $R、G、B、H、S、V、I$ 在内的,运用 Matlab 进行数据预处理,例如平滑处理(或去噪)、标准变换和极差归一化变换等,本研究采用极差归一化处理特征分量均值原始数据,然后再利用 Matlab 进行距离判别分析。

数据的极差归一化处理^[12],即将数据统一映射到 $[0,1]$ 区间上。设 p 维向量 $\mathbf{X} = (X_1, X_2, \dots, X_p)$ 的观测值矩阵为

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \quad (1)$$

极差归一化变换后的矩阵为

$$\mathbf{X}^R = \begin{bmatrix} x_{11}^R & x_{12}^R & \cdots & x_{1p}^R \\ x_{21}^R & x_{22}^R & \cdots & x_{2p}^R \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1}^R & x_{n2}^R & \cdots & x_{np}^R \end{bmatrix}, \tag{2}$$

其中, $x_{ij}^R = \frac{x_{ij} - \min_{1 \leq k \leq n} x_{kj}}{\max_{1 \leq k \leq n} x_{kj} - \min_{1 \leq k \leq n} x_{kj}}, i = 1, 2, \cdots, n; j = 1, 2, \cdots, p$ 。

这里 $\min_{1 \leq k \leq n} x_{kj}$ 为变量 X_j 的观测值, $\max_{1 \leq k \leq n} x_{kj} - \min_{1 \leq k \leq n} x_{kj}$ 为变量 X_j 的观测值的极差。经过归一化变换后, 矩阵 $\mathbf{X}^R$ 的每个元素的取值均在 0~1。

2 结果与讨论

2.1 获取颜色特征分量均值

运用 Matlab 软件得到去背景之后的番茄 RGB 图像, 将其

转换到 HSV 和 HIS 颜色空间下, 分别获取它们的单分量的值, 然后再求得它们的均值, 各颜色分量的单分量均值见表 1。

通过已获得的颜色单分量值, 从而可以获取组合颜色特征分量 $(R-G)/(G-B)$ 、 $(R+G)/(G-B)$ 、 $(R-G)/(R+G)$ 、 G/B 、 R/B 、 B/R 的值, 进而求得组合颜色特征分量的均值见表 2。

2.2 获取归一化均值

运用 Matlab 将原始的单分量与组合分量值进行归一化处理, 从而得到其分量的归一化均值, 见表 3、4。

2.3 判别分类结果分析

运用 SPSS 筛选出的组合特征分量对完熟样本、成熟样本、半熟样本、绿熟样本 4 类成熟度番茄进行判别分析, 将完熟、成熟、半熟、绿熟样品集按照 2:1 的比例划分成训练集和验证集。利用 Mahalanobis、quadratic、linear 3 种函数分别进行判别分析, 得到的判别分析结果见表 5~7。

表 1 不同成熟度番茄颜色单分量均值

Table 1 The average of single color component of different maturity tomato							
成熟度	R 分量	G 分量	B 分量	H 分量	S 分量	V 分量	I 分量
完熟	17.010 8	10.685 3	6.701 7	0.009 7	0.209 8	0.148 4	14.571 9
成熟	16.218 6	11.106 7	6.674 6	0.008 6	0.196 2	0.145 9	14.085 6
半熟	16.768 8	13.646 9	8.786 7	0.011 2	0.190 7	0.163 9	15.191 9
绿熟	14.858 4	14.944 7	13.163 4	0.040 0	0.099 5	0.140 8	14.508 7

表 2 不同成熟度番茄颜色组合分量均值

Table 2 The average of combination color components of different maturity tomato						
成熟度	$(R-G)/(G-B)$	$(R+G)/(G-B)$	$(R-G)/(R+G)$	G/B	R/B	B/R
完熟	1.609 6	7.005 6	0.229 5	1.600 9	2.556 0	0.392 8
成熟	1.182 2	6.222 8	0.188 3	1.669 7	2.446 3	0.410 3
半熟	0.705 7	6.540 6	0.103 9	1.587 3	1.960 5	0.518 0
绿熟	-0.081 2	20.114 1	-0.003 6	1.139 0	1.131 6	0.888 0

表 3 不同成熟度番茄颜色单分量归一化均值

Table 3 The normalized mean of single color component of different maturity tomato							
成熟度	R 分量	G 分量	B 分量	H 分量	S 分量	V 分量	I 分量
完熟	0.496 1	0.297 1	0.205 2	0.087 4	0.664 4	0.445 7	0.476 1
成熟	0.442 0	0.322 7	0.202 5	0.066 6	0.599 4	0.426 9	0.437 7
半熟	0.481 4	0.488 5	0.336 0	0.117 0	0.576 9	0.551 1	0.525 1
绿熟	0.353 4	0.569 4	0.607 8	0.670 5	0.152 3	0.392 8	0.471 1

表 4 不同成熟度番茄颜色组合分量归一化均值

Table 4 The normalized mean of combination color components of different maturity tomato						
成熟度	$(R-G)/(G-B)$	$(R+G)/(G-B)$	$(R-G)/(R+G)$	G/B	R/B	B/R
完熟	0.606 5	0.048 8	0.894 4	0.702 9	0.779 5	0.092 0
成熟	0.465 9	0.030 0	0.746 8	0.789 9	0.723 2	0.119 9
半熟	0.309 1	0.037 7	0.444 4	0.685 7	0.473 7	0.291 4
绿熟	0.050 2	0.363 8	0.059 2	0.119 1	0.047 9	0.880 4

表 5 Mahalanobis 判别分析结果

Table 5 Discriminant analysis of Mahalanobis

成熟度	样品集		判别率/%	
	训练集	验证集	训练集	验证集
完熟	20	10	80.00	60.00
成熟	20	10	76.67	50.00
半熟	14	5	94.74	80.00
绿熟	20	10	100.00	100.00

表 6 quadratic 判别分析结果

Table 6 Discriminant analysis ofquadratic

成熟度	样品集		判别率/%	
	训练集	验证集	训练集	验证集
完熟	20	10	90.00	80.00
成熟	20	10	73.33	60.00
半熟	14	5	89.47	80.00
绿熟	20	10	100.00	100.00

表 7 linear 判别分析结果

Table 7 Discriminant analysis oflinear

成熟度	样品集		判别率/%	
	训练集	验证集	训练集	验证集
完熟	20	10	86.67	90.00
成熟	20	10	76.67	70.00
半熟	14	5	89.47	100.00
绿熟	20	10	96.67	100.00

表 8 Mahalanobis、quadratic、linear 判别分析结果对比

Table 8 Discriminant analysis results contrast of Mahalanobis, quadratic, linear

判别函数	训练集	验证集	综合判别率/%
Mahalanobis	74	35	77.14
quadratic	74	35	80.00
linear	74	35	94.29

由表 5 可知,使用 Mahalanobis 函数判别的情况下,绿熟番茄训练集和验证集判别率最好,都达到了 100.00%。由表 6 可知,使用 quadratic 函数判别的情况下,绿熟番茄训练集和验证集判别率也是最好,均为 100.00%,并且 4 种成熟度番茄验证集判别率都有所提高。由表 7 可知,使用 linear 函数判别的情况下,绿熟番茄训练集判别率达到了 96.67%,半熟和绿熟番茄验证集判别率均为 100.00%,4 种成熟度番茄验证集判别率均达到了最好。从表 8 可以得出,使用 Mahalanobis 判别函数的情况下,综合判别率为 77.14%,使用 quadratic 判别函数的情况下,综合判别率为 80.00%,使用 linear 判别函数的情况下,综合判别率为 94.29%,因此,通过综合对比得出,使用 linear 判别函数的判别率最佳。

综上所述,当利用 Matlab 进行距离判别分析时,3 种判别函数下,半熟和绿熟番茄的判别率与完熟和成熟番茄相比会较好。综合对比下,使用 linear 判别函数的判别率最佳。

这是因为在数据极差归一化过程中,半熟和绿熟番茄的数据离散点误差明显会少于完熟和成熟番茄,数据的线性度较好,因此,会获得较准确的判别率。另外可能的原因是颜色分量均值受到样本的影响,当实际的颜色分量均值与判别函数的特征值不吻合时,使得出现误判的可能性较大。

3 结论

(1) 本研究利用机器视觉的方法,通过番茄表面的颜色特征对不同成熟度的番茄进行判别分类,并将样品集按照 2:1 划分为训练集和验证集,从目前分类结果来看,较好地实现了对 4 种不同成熟度番茄的判别分类。

(2) 从判别分析结果来看,绿熟番茄在 3 种不同判别函数下训练集与验证集判别率均最好,达到了 100.00%;半熟番茄训练集判别率最高为 94.74%,同时验证集判别率最高达到 100.00%;成熟番茄训练集与验证集判别率最低,分别为 76.67%和 70.00%;完熟番茄训练集与验证集最高均为 90.00%,综合对比下,使用 linear 判别函数的判别率最佳。

(3) 通过判别分析函数总体实现不同成熟度番茄的判别分类,但是对于完熟与成熟番茄的判别效果不理想。在后续的研究中,将改进并优化图像处理算法,增强算法的鲁棒性以适应不同的对象;选择新的特征信息并结合颜色特征分量,以此提高判别率。

参考文献

[1] CHOI K, LEE G, HAN Y J, et al. Tomato maturity evaluation using color image analysis[J]. Transactions of the ASABE, 1995 (1): 171-176.

[2] FELIPE Avila, MARCO Mora, MIGUEL Oyarce, et al. A method to construct fruit maturity color scales based on support machines for regression; Application to olives and grape seeds [J]. Journal of Food Engineering, 2015(2): 9-17.

[3] 方政, 胡晓辉, 陈永, 等. 基于计算机视觉的成熟番茄识别研究 [J]. 农机化研究, 2016, 38(8): 31-35.

[4] 张瑞合, 姬长英, 沈明霞, 等. 计算机视觉技术在番茄收获中的应用[J]. 农业机械学报, 2001, 32(5): 50-52.

[5] 蒋焕煜, 彭永石, 申川, 等. 基于双目立体视觉技术的成熟番茄识别与定位[J]. 农业工程学报, 2008, 24(8): 279-283.

[6] 赵杰文, 刘木华, 杨国彬. 基于 HIS 颜色特征的田间成熟番茄识别技术[J]. 农业机械学报, 2004, 35(5): 122-124, 135.

[7] 纪平, 王俊, 陈鹤碧. 基于图像的番茄识别与几何尺寸测量[J]. 安徽农业科学, 2012, 40(33): 16 426-16 428, 16 434.

[8] 刘宇飞, 汤晓华. 基于机器视觉的番茄颜色分级方法的研究[J]. 机电产品开发与创新, 2013, 26(2): 99-100.

[9] 彭波, 李旭宇. TCS230 颜色传感器结合 BP 神经网络在颜色识别系统上的研究[J]. 食品与机械, 2010, 26(3): 108-112, 123.

[10] 董坦坦. 成熟番茄的图像识别及其位姿的获取研究[D]. 南京: 南京农业大学, 2009: 8-9.

[11] 尹建军, 王新忠, 毛罕平, 等. RGB 与 HSI 颜色空间下番茄图像分割的对比研究[J]. 农机化研究, 2006(11): 171-174.

[12] 谢中华. MATLAB 统计分析与应用 40 个案例分析[M]. 北京: 北京航空航天大学出版社, 2010: 345-350.