

基于机器视觉的辣椒外部品质检测

Study on external quality detection of pepper based on machine vision

任 锐

张淑娟

赵华民

REN Rui ZHANG Shu-juan ZHAO Hua-min

孙海霞

李成吉

廉孟茹

SUN Hai-xia LI Cheng-ji LIAN Meng-ru

(山西农业大学工学院, 山西 晋中 030800)

(College of Engineering, Shanxi Agriculture University, Jingzhong, Shanxi 030800, China)

摘要:为实现正常辣椒和缺陷辣椒(黑斑、虫蚀和花皮)的快速检测,通过图像采集,利用中值滤波去噪;提取样本图像 24 个颜色特征值,20 个纹理特征值,利用 Otsu 算法阈值分割,提取 2 个形态特征值;选用连续投影法(Successive Projections Algorithm, SPA)优选 14 个特征值,结合最小二乘支持向量机(Least Squares Support Vector Machine, LS-SVM)方法进行建模,SPA-LS-SVM 模型检测结果为 98.67%,模型最优。

关键词:辣椒;品质检测;机器视觉;连续投影法;最小二乘支持向量机

Abstract: To realize the rapid detection of normal peppers and defective peppers (dark spots, insect marks and stripe skin), image acquisition and median filtering were used for denoising. 24 color feature variables and 20 texture feature variables of sample images were extracted. Threshold segmentation in the Otsu algorithm was used to extract two morphological feature variables. The Successive Projections Algorithm (SPA) was used to optimize 14 feature variables. Modeling was done based on the Least Squares Support Vector Machine (LS-SVM) method. The detection result of SPA-LS-SVM model is 98.67%, which is the best.

Keywords: pepper; quality detection; machine vision; successive projections algorithm; least squares support vector machine

辣椒是世界最大的调味品作物,主要以果实为食,营养丰富,维生素 C 含量在蔬菜中占首位。辣椒市场价值

主要受其外观质量的影响,依据 NY/T 944—2006 标准,正常辣椒形状均匀,色泽油亮光洁呈红色,部分辣椒出现黑斑、虫蚀和花皮等缺陷,需分选出来。

目前的分选主要靠人工分选,效率低、主观性强且工作量大,机器视觉利用计算机技术代替人眼获取图像的信息,具有无损、检测精度高、节省劳动力等优点,科研人员在机器视觉方面对农产品做了大量的研究,取得较大的发展。毕智健等^[1]利用番茄表面的颜色特征对不同成熟度的番茄进行判别分类,完熟番茄训练集与验证集最高均为 90%。解博等^[2]通过机器视觉对杏核纹理特征进行识别,杂质杏核识别率可达 94%。项辉宇等^[3]采用基于 Halcon 的图像处理方法,依据 RGB 颜色模型,采用模板匹配的方法,对苹果品质的大小、缺陷、颜色检测判别苹果是否合格。Xu^[4]提出利用最小核值相似区和脉冲耦合神经网络识别苹果果实,在测试的 50 幅图像中,识别率为 93%。Arakeri^[5]利用计算机视觉技术,运用图像处理技术分析番茄的缺陷和成熟度,对番茄的质量评估的准确率达到 96.47%。Habib^[6]利用图像提取木瓜色彩特征,解决木瓜病害识别,通过支持向量机对疾病特征分类,准确率达到 90.15%。综上所述,机器视觉研究主要利用颜色特征、纹理特征、形态特征和直接利用图像进行判别,其他学者的研究大都采取单一特征判别。

研究拟利用机器视觉检测辣椒外部品质,通过机器视觉系统进行图像采集,图像增强,将提取颜色特征、纹理特征和形态特征作为特征值。通过回归系数法和连续投影法优选特征值,利用偏最小二乘回归法和最小二乘支持向量机建模分析,验证研究中选择特征值和模型的有效性,为今后的辣椒在线检测提供参考和技术支持。

1 材料和方法

1.1 试验样本

试验所采用的干辣椒购买自山西省襄汾县赵雄村的

基金项目:山西省自然科学基金项目(编号:201801D121252);山西省高校科技创新项目(编号:2019L0402)

作者简介:任锐,男,山西农业大学在读硕士研究生。

通信作者:张淑娟(1963—),女,山西农业大学教授,博士。

E-mail: zsujuan1@163.com

收稿日期:2020-06-16

三辣椒。利用 K-S(Kennard-Stone)算法^[7]按 3 : 1 划分训练集和预测集,其中正常样本训练集 225 组,预测集 75 组;缺陷样本训练集 225 组,预测集 75 组。

1.2 机器视觉系统和数据采集

机器视觉系统包括暗箱(440 mm×335 mm×425 mm)内壁使用黑色拷贝纸进行贴壁处理,屏蔽外界自然光;将相机(尼康 D3300)和镜头(尼康尼克尔镜头焦距:18~140 m,光圈:f/3.5-5.6G ED VR)安装于暗箱顶部,采集辣椒图像信息;暗箱顶部布置环形荧光灯(YH327200KE)提供光源,暗箱下部安装可调节载物台,载物台顶部位于镜头正下方 24 cm 处。通过存储卡(SDSQUNC-032G-ZN6MA)将图像传入计算机(处理器:Y50P15-4210H, CPU: 2.90 HZ, 显卡: NVIDIA GTX 960M)中,采用 MATLAB R2010b 软件进行图像处理。

2 结果与分析

2.1 图像增强

图像增强是为了去除噪声,提高处理图像的质量和速度。中值滤波法为线性滤波器,可以减少噪声,保护图像边缘不受影响。对图像进行裁剪和灰度处理,利用中值滤波去噪,达到增强图像的目的,为后续提取特征值作准备,图像增强结果如图 1 所示。



图 1 图像增强

Figure 1 Image enhancement

2.2 特征值提取

2.2.1 颜色特征值提取 颜色特征描述图像区域所对应的目标物的表面性质^[8]。RGB 颜色空间是最基础的模型;HSI 颜色空间接近于人眼视觉特性;HSV 颜色空间由色调、饱和度和亮度组成;Lab 颜色空间是范围最大的色彩模式。从 RGB、HSV、HSI、Lab 颜色空间中分别提取 R、G、B、H、S、V、H、S、I、L、a、b 颜色特征值的最大值、最小值、均值和方差,提取样本各颜色特征值如表 1 所示。

2.2.2 纹理特征值提取 纹理特征描述图像或图像区域所对应的表面性质。主要从灰度差分统计法和灰度共生矩阵来提取能量、熵、惯性矩、相关性、逆差距的 4 个方向(0°,45°,90°,135°)的特征值的均值^[9],提取结果如表 2 所示。

表 1 样本各颜色特征值

Table 1 Sample color eigenvalue

颜色空间	特征值	正常辣椒				缺陷辣椒			
		最大值	最小值	均值	方差	最大值	最小值	均值	方差
RGB	R 值	147.00	21.00	84.44	14.58	165.00	23.00	87.73	11.92
	G 值	154.00	0.00	73.65	21.27	169.00	0.00	69.74	16.77
	B 值	164.00	0.00	79.65	22.26	188.00	0.00	78.50	18.65
HSI	H 值	0.99	0.00	0.75	0.10	0.99	0.00	0.73	0.13
	S 值	1.00	0.00	0.42	0.14	1.00	0.01	0.47	0.11
	I 值	0.60	0.03	0.36	0.07	0.68	0.04	0.38	0.06
HSV	H 值	0.99	0.00	0.75	0.10	0.99	0.00	0.73	0.13
	S 值	1.00	0.00	0.31	0.13	1.00	0.02	0.35	0.12
	V 值	0.64	0.08	0.31	0.08	0.74	0.09	0.31	0.06
Lab	L 值	63.33	1.81	31.68	8.43	69.25	1.91	31.04	6.72
	a 值	32.57	-5.48	6.64	2.89	32.91	-1.49	10.47	2.51
	b 值	29.25	-8.71	-0.94	1.74	33.87	-15.20	-0.91	2.66

表 2 样本纹理特征值均值

Table 2 Mean of texture feature values of sample

纹理特征	正常辣椒				缺陷辣椒			
	0°	45°	90°	135°	0°	45°	90°	135°
能量	0.09	0.08	0.13	0.08	0.04	0.04	0.08	0.04
熵	3.04	3.09	2.65	3.12	3.51	3.57	3.07	3.59
惯性矩	5.23	4.96	0.78	6.49	4.16	4.18	0.82	5.24
相关性	0.07	0.07	0.08	0.06	0.07	0.07	0.07	0.07
逆差距	0.08	0.07	0.12	0.07	0.03	0.03	0.08	0.03

2.2.3 形态特征值提取 图像形态特征是在物体从图像中分割出来后进行分析,在机器视觉中起着十分重要的作用^[10]。通过计算缺陷面积比值和欧拉数确定辣椒类别。利用 Otsu 算法阈值分割,二值化分别提取模板图像和缺陷图像,如图 2、3 所示。

$$S = \frac{s_1}{s_2}, \quad (1)$$

式中:

S ——面积比;

s_1 ——模板图像像素面积;

s_2 ——缺陷图像像素面积。

$$EUL = C - H, \quad (2)$$

式中:

EUL ——欧拉数;

C ——代表对象总数;

H ——孔洞数。

2.3 优选特征值

研究中,通过提取出 600 组图像,每幅图像 46 个特征值。由于特征值提取数据量大,每一组特征值对辣椒检测影响不同,因此需要提取影响效果显著的特征值建模检测。

2.3.1 RC 方法优选特征值 回归系数法(RC)通过建立偏最小二乘回归法(PLS)模型的回归系数中的局部极值进行提取,其表达了 PLS 方法建模中每个特征值对此次建模相关性的情况。RC 方法优选特征值如图 4 所示,优选出 21 个特征值,分别为 5(0°方向的熵),6(45°方向的熵),7(90°方向的熵),8(135°方向的熵),9(0°方向的惯性矩),11(90°方向的惯性矩),20(135°方向的逆差距),21(比值),24(RGB 中 R 值方差),25(RGB 中 G 值均值),26(RGB 中 G 值方差),27(RGB 中 B 值均值),29(HSV 中 H 值均值),30(HSV 中 H 值方差),31(HSV 中 S 值均值),35(HSV 中 H 值均值),36(HSV 中 H 值方差),



图 2 提取模板图像

Figure 2 Extraction of template image



图 3 提取缺陷图像

Figure 3 Extraction of defect image

37(HSV 中 S 值均值),41(Lab 中 L 值均值),42(Lab 中 L 值方差),44(Lab 中 a 值方差)。

2.3.2 SPA 方法优选特征值 连续投影法(SPA)是一种研究系统确定性与不确定性相互作用的分析理论^[11]。SPA 方法优选特征值如图 5 所示,可得 14 个优选特征值最优解依次排列为 42(Lab 中 L 值方差),26(RGB 中 G 值方差),34(HSV 中 V 值方差),21(比值),18(45°方向的逆差距),5(0°方向的熵),6(45°方向的熵),17(0°方向的逆差距),8(135°方向的熵),32(HSV 中 S 值方差),45(Lab 中 b 值均值),46(Lab 中 b 值方差),7(90°方向的熵),30(HSV 中 H 值方差)。

2.4 建立检测模型

偏最小二乘回归法(PLS)是一种多元统计数据分析方法^[12]。利用 PLS 方法与 RC 方法优选特征值和 SPA 方法优选特征值分别建模检测,判别结果如表 3 所示。

最小二乘支持向量机(LS-SVM)是支持向量机(SVM)的一种扩展,解决问题的线性性能。利用 LS-SVM 方法与 RC 方法优选特征值和 SPA 方法优选特征值分别建模检测,判别结果如表 3 所示。

由于模型预测出的结果非整数型,因此将 0.5 作为最大偏离值,预测类别值和假设类别值相差小于 0.5 的判定为此类样本^[13]。由表 3 可知,利用 PLS 方法和 LS-SVM 方法分别结合 RC 方法和 SPA 方法优选特征值分别建模

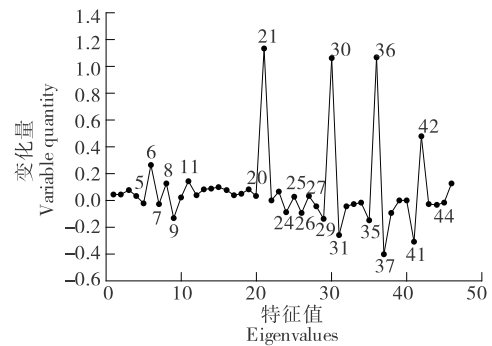


图 4 RC 曲线

Figure 4 RC curve

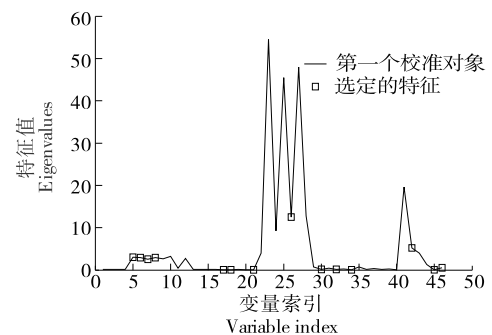


图 5 SPA 优选特征值

Figure 5 Preferred eigenvalues by SPA

表 3 检测模型判别结果分析

Table 3 Analysis of detection model discrimination results

检测模型	辣椒类别	特征值数	数量	误判数	判别率/%	总判别率/%
PLS	正常辣椒	46	75	1	98.67	97.33
	缺陷辣椒	46	75	3	96.00	
RC-PLS	正常辣椒	21	75	1	98.67	95.33
	缺陷辣椒	21	75	6	92.00	
SPA-PLS	正常辣椒	14	75	0	100.00	96.67
	缺陷辣椒	14	75	5	93.33	
LS-SVM	正常辣椒	46	75	0	100.00	98.67
	缺陷辣椒	46	75	2	97.00	
RC-PLS	正常辣椒	21	75	0	100.00	96.67
	缺陷辣椒	21	75	5	93.33	
SPA-PLS	正常辣椒	14	75	0	100.00	98.67
	缺陷辣椒	14	75	2	97.00	

分析,RC-PLS 模型判别率为 95.33%,SPA-PLS 模型和 RC-PLS 模型判别率为 96.67%,PLS 模型判别率为 97.33%,LS-SVM 模型和 SPA-LS-SVM 模型判别率为 98.67%,6 种方法检测结果都达到了 90% 以上。显然 LS-SVM 模型和 SPA-LS-SVM 模型检测结果最好,SPA-LS-SVM 方法可以通过 SPA 方法优选特征值简化模型,提高检测精度和效率,因此选用 LS-SVM 方法建模结合 SPA 方法优选特征值检测模型最优。

3 结论

研究主要利用机器视觉系统进行图像采集,利用中值滤波增强图像,减少噪声;通过提取 RGB、HSV、HSI、Lab 4 种颜色空间各颜色特征值的均值和方差,提取纹理特征的能量、熵、惯性矩、相关性、逆差距的 4 个方向(0°, 45°, 90°, 135°)的特征值,利用 Otsu 算法阈值分割图像,提取形态特征缺陷比值和欧拉数作为特征值,共提取 46 个特征值。结果表明,选用连续投影法优选 14 个特征值,结合最小二乘支持向量机法进行建模,所得模型的检测准确率为 98.67%,模型最优。

参考文献

[1] 毕智健,张若宇,齐妍杰,等. 基于机器视觉的番茄成熟度颜色判别[J]. 食品与机械, 2016, 32(12): 133-136.
[2] 解博,周律. 基于机器视觉的杏核图像识别系统设计[J]. 电子科技, 2016, 29(10): 97-100.
[3] 项辉宇,薛真,冷崇杰,等. 基于 Halcon 的苹果品质视觉检测试验研究[J]. 食品与机械, 2016, 32(10): 123-126.
[4] XU Li-ming. Recognition method for apple fruit based on SUSAN and PCNN[J]. Multimedia Tools and Applications, 2018, 77(6): 7 205.

[5] ARAKERI Megha P. Computer vision based fruit grading system for quality evaluation of tomato in agriculture industry[J]. Procedia Computer Science, 2016, 79: 426-433.
[6] HABIB Md Tarek. Machine vision based papaya disease recognition[J]. Journal of King Saud University - Computer and Information Sciences, 2020, 32(3): 300-309.
[7] 陈奕云,赵瑞瑛,齐天赐,等. 结合光谱变换和 Kennard-Stone 算法的水稻土全氮光谱估算模型校正集构建策略研究[J]. 光谱学与光谱分析, 2017, 37(7): 2 133-2 139.
[8] 宋西平,李国琴,罗陆锋,等. 基于 HSI 色彩空间与 FFCM 聚类的葡萄图像分割[J]. 农机化研究, 2015, 37(10): 40-44.
[9] 薛建新,张淑娟,张晶晶. 高光谱成像技术对鲜枣内外部品质检测的研究[J]. 光谱学与光谱分析, 2015, 35(8): 2 297-2 302.
[10] 马志宏,贡亮,林可,等. 基于稻穗几何形态模式识别的在穗籽粒数估测[J]. 上海交通大学学报, 2019, 53(2): 239-246.
[11] 李军宇,张淑娟,张学豪,等. 李果实成熟度的高光谱成像判别研究[J]. 农机化研究, 2017, 39(12): 141-145.
[12] 孙海霞,张淑娟,薛建新,等. 变量优选校正算法的鲜枣可溶性固形物检测模型传递方法研究[J]. 光谱学与光谱分析, 2019, 39(4): 1 041-1 046.
[13] 赵旭婷,张淑娟,刘蒋龙,等. 高光谱技术结合 CARS-ELM 的油桃品种判别研究[J]. 现代食品科技, 2017, 33(10): 281-287.