

基于纹理和颜色特征的甜瓜缺陷识别*

王书志¹ 张建华² 冯全³

(1. 西北民族大学电气工程学院, 兰州 730030; 2. 中国农业大学工学院, 北京 100083;
3. 甘肃农业大学工学院, 兰州 730070)

【摘要】 为了提高硬皮甜瓜缺陷分类的正确率,提取基于纹理和颜色的综合特征,采用支持向量机分类器构造了甜瓜缺陷的自动检测系统。对甜瓜图像可疑区进行了纹理分析,提取灰度共生矩阵的4个特征参数,经过比较实验得出,对比度和角二阶矩2个参数对甜瓜瓜蒂、花萼、擦伤和霉变有明显的可区分性。在可疑区域上提取了由R、G、B分量及其算术运算组成的12种颜色特征,通过实验筛选出4种具有较好区分性的颜色特征。实验结果表明,由这些优选出的纹理与颜色特征组成的综合特征及支持向量机分类器对甜瓜缺陷的识别正确率达到92.2%。

关键词: 甜瓜 缺陷检测 纹理特征 颜色特征 支持向量机
中图分类号: TP391.41; S126 文献标识码: A 文章编号: 1000-1298(2011)03-0175-05

Defect Detection of Muskmelon Based on Texture Features and Color Features

Wang Shuzhi¹ Zhang Jianhua² Feng Quan³

(1. College of Electrical Engineering, Northwest University for Nationalities, Lanzhou 730030, China
2. College of Engineering, China Agricultural University, Beijing 100083, China
3. College of Engineering, Gansu Agricultural University, Lanzhou 730070, China)

Abstract

In order to improve the accuracy of muskmelon's defect detection, an automatic defect detection system based on support vector machine (SVM) was set up by adopting complex features of texture and color. Four textural parameters and twelve color features of combinations from RGB were tested for the discriminability in stem, calyx, bruise and mildew. Through the experiments, two textural and four color features with good discriminability were selected and treated as the complex features. The results indicated that with the complex features and SVM, the accuracy of classification on muskmelons was up to 92.2%.

Key words Melon, Defect detection, Texture features, Color features, SVM

引言

在水果品质自动分级中,梗蒂与缺陷在计算机图像处理过程中容易混淆,因此,对梗蒂和缺陷进行识别分类是关键问题,国内外学者为此开展了大量研究。Wolfe等^[1]提出了采用黑白摄像头基于苹果图像灰度的缺陷检测方法;徐娟等^[2]提出了利用形

态学识别梗蒂的方法;应义斌等^[3]提出了黄花梨果面缺陷检测方法;何东健等^[4]进行了果实缺陷面积的计算机视觉测定研究;李庆中等^[5]研究了基于分形特征的水果缺陷快速识别方法;冯斌等^[6]利用傅里叶算子识别水果缺陷;蒋益女等^[7]结合纹理分析和SVM实现了苹果梗蒂和缺陷的识别。已有研究都没有对甜瓜缺陷作系统的分类。

收稿日期: 2010-07-07 修回日期: 2010-08-19
* 西北民族大学中青年科研基金资助项目(X2007-008)
作者简介: 王书志,副教授,主要从事电子信息技术研究,E-mail: wmszl@sohu.com
通讯作者: 冯全,教授,主要从事信息处理技术研究,E-mail: fquan@sina.com

本文对甜瓜的瓜蒂 (stem) 和花萼 (calyx) 区域 (以下简称 SC 区域)、缺陷进行分类,其中缺陷又细分为擦伤和霉变,擦伤是甜瓜受外力而表皮受损,可食用,而霉变是甜瓜内部果肉腐烂引起,不可食用。由于梗蒂与缺陷的相似性,本研究提取纹理特征和颜色特征组成综合特征,结合支持向量机设计甜瓜缺陷分类器。

1 实验材料与方法

1.1 实验材料

实验甜瓜的品种是红金宝,由市场采购,甜瓜上的霉变是采购存放一段时间后自然产生的。实验设备由计算机视觉检测系统、光电触发电路与机械传输带组成。计算机视觉系统包含光照箱、光源、CCD 摄像头、图像采集卡、计算机。其中,光照箱是 60 cm × 60 cm × 60 cm 的自制木板立方箱体;光源为 4 盏白炽灯,分别固定在光照箱 4 个角;摄像头为 OK_AC1300 高分辨彩色摄像机 (最大分辨率为 1 300 × 1 024,局部行输出时,最大帧频可高达 230 帧/s,支持外触发功能),图像采集卡为 OK_RGB10B 彩色采集卡;计算机为 PC 机 Pentium4/2.6 G 处理器,512 MB 内存。

1.2 实验方法

设计了针对甜瓜的瓜蒂、花萼和缺陷的区分方法,首先对采集到的一帧甜瓜图像预处理,将甜瓜从背景中分割出来;然后运用 Otsu 算法对甜瓜进行可疑区域 (SC、缺陷区域) 分割;在可疑区域上提取灰度共生矩阵,并计算出灰度纹理特征;颜色特征的计算也是在可疑区域上进行的;最后采用支持向量机分类器进行模式识别,如图 1 所示。

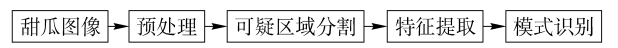


图 1 甜瓜缺陷检测分类算法框图

Fig.1 Chart of muskmelon defect detection and classification algorithm

2 可疑区域的特征提取与分类

2.1 图像预处理与可疑区域分割

甜瓜图像采集时,其大小设定为 800 × 600 像素。图像预处理的灰度变换、快速中值滤波、背景分割、形态学开运算组成。灰度变换把彩色甜瓜图像进行灰度化处理,便于提高图像处理速度。由于是在可控的均匀照明环境中,背景分割采用全局阈值法,可以快速将甜瓜从背景中分割出来。形态学开运算即先进行腐蚀运算,再进行膨胀运算,其作用是平滑甜瓜边缘,消除甜瓜内部小空洞。

在分割后的甜瓜区域上采用 Otsu 算法自动提取可疑区域。Otsu 算法的基本思想为确定一个最佳阈值,使图像二值化处理后,可疑区域与甜瓜背景两个像素类间方差最大。

2.2 灰度纹理特征

灰度共生矩阵建立在估计图像二阶组合条件概率密度函数基础上,是一种重要的纹理分析方法。灰度共生矩阵反映了图像灰度关于方向、相邻间隔、变化幅度的综合信息,为了能更直观地以共生矩阵描述纹理状况,可以从共生矩阵导出一些反映矩阵状况的参数。Haralick^[8]曾提出 14 种由灰度共生矩阵计算出的参数,本文选择其中的对比度、熵、角二阶矩、局部均匀性 4 种参数作对比。这 4 个参数所表达的纹理意义和变化规律如下:对比度也称惯性矩,反映纹理变化、周期性的物理量,其值越大,表明纹理周期性越大;熵是图像所具有信息量的度量,图像若没有纹理信息,则熵值也接近为零,图像若充满着细纹理,则该图像的熵值也大;角二阶矩也称能量,是图像灰度分布均匀程度和纹理粗细的一个度量,当图像较细致、均匀时,角二阶矩值较大;局部均匀性反映纹理的规则程度,纹理杂乱无章、难以描述的,局部均匀性值较小,规律性较强、易于描述的,局部均匀性值较大。

2.3 颜色特征

SC 及缺陷区域的每个像素点都包含着色度、亮度及饱和度等丰富的颜色特征信息。大量实验表明,利用像素点的 R、G、B 三分量及其不同组合形式及数学运算形成的一些特征参数,对缺陷分类识别十分有效^[9-10]。根据现有文献,选取了 12 个颜色参数,分别是: $I_1 = R, I_2 = G, I_3 = B, I_4 = (2R - G - B)/4, I_5 = (2G - R - B)/4, I_6 = (2B - G - R)/4, I_7 = G - B, I_8 = G - R, I_9 = R - B, I_{10} = 224(G - B)/(G + B), I_{11} = (R + G + B)/3, I_{12} = 224B/R$, 其中 I_{10} 和 I_{12} 中的常系数是通过实验确定的。由于以上参数数量较多,故对训练样本中的 SC、擦伤和霉变区域中各像素点进行统计实验,从中进一步筛选出若干个区分性较好的参数作为分类颜色特征。

2.4 模式分类

分类器采用支持向量机 (support vector machine, 简称 SVM)^[11]。支持向量机是以统计学习理论为基础的通用机器学习技术,能够较好地解决小样本、非线性、高维数和局部极小点等模式识别问题。与传统的神经网络相比,支持向量机最终将转换为一个二次寻优问题,从理论上讲得到的是全局最优点,解决了神经网络中无法避免的局部极小值问题。

支持向量机的基本思想可以概括为:首先通过非线性变换将输入空间变换到一个高维空间,然后在这个新空间中求取最优线性分类面,而这种非线性变换是通过定义适当的核函数(内积函数)实现的。常用的核函数有:

- (1) 多项式核函数: $K(x,y)=(xy+1)^q$ 。
- (2) 径向基核函数: $K(x,y)=\exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right)$ 。
- (3) Sigmoid 核函数: $K(x,y)=\tanh[v(xy)+c]$ 。

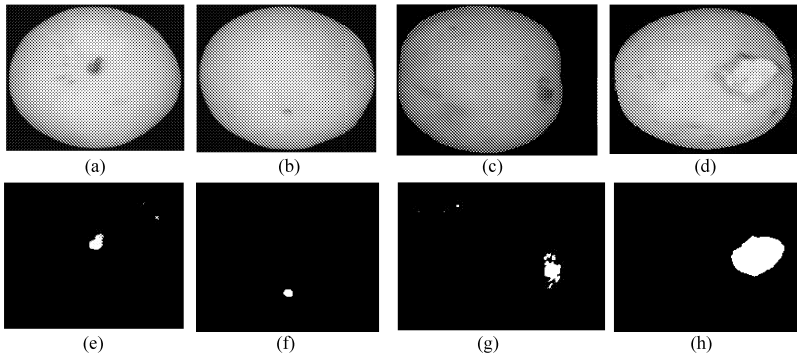


图 2 甜瓜可疑区域的分割

Fig. 2 Segmentation of interesting region on muskmelon images

3.2 灰度纹理特征的实验结果

从训练图像中得到瓜蒂共 28 处、花萼 25 处,擦伤 24 处、霉变 27 处。经过滑动窗口在这些区域的移动生成灰度共生矩阵,然后计算灰度共生矩阵的角二阶矩阵、相关性、均匀性、惯性等 4 个特征向量,经过实验比较得出,其中滑动窗口大小为 9×9 、灰度等级 $L=32$ 、步距 $d=1$,方向 θ 取 0° 、 45° 、 90° 和 135° 上的平均值为最佳。

对以上 4 种类型的 4 个灰度纹理特征参数的分布区间进行统计,结果如图 3 所示,其中, Q_1 、 Q_2 、 Q_3 和 Q_4 分别代表对比度、熵、角二阶矩和局部均匀性,图中工字形代表参数的数值分布范围。从图中可以看出,4 种参数关于 4 种检测类型不都具有明显的可区分性,其中熵和局部均匀性特征的分布范围彼此有相当多的交叠,无法将 4 种检测类型区分开,故舍弃。而 4 种检测类型在对比度和角二阶矩特征的分布上有明显的差别,故选择它们作为分类器的纹理特征。

图 4 是对比度和角二阶矩关于 4 种检测类型的样本分布,从图中可以看出,基于对比度和角二阶矩两个特征参数对瓜蒂、花萼、擦伤和霉变有一定的聚类作用,各自样本成团分布,其中,花萼和霉变分布较集中,有明显的分界线,聚类效果较好;但擦伤和瓜蒂有交叉重叠的地方,没有明显的分界线,聚类效果不佳,需要引入其他特征变量。

3 实验结果与分析

3.1 可疑区分割

通过计算机视觉检测系统采集了 200 幅含有甜瓜可疑区域的图像作为样本集,其中 100 幅作为训练样本,另外 100 副作为测试样本。图 2 给出了 Otsu 算法分割不同可疑区域(瓜蒂、花萼、擦伤和霉变)的图片,其中,图 2a、2b、2c、2d 分别为带有瓜蒂、花萼、擦伤、霉变的甜瓜图像;图 2e、2f、2g、2h 分别为通过 Otsu 算法分割出的可疑区域。

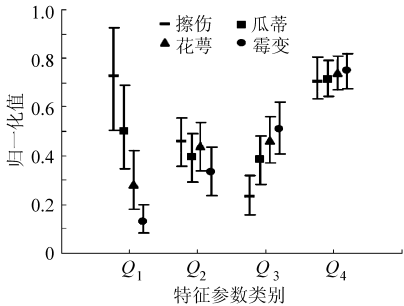


图 3 瓜蒂、花萼、擦伤、霉变分别在 4 个特征参数统计分布

Fig. 3 Distribution regions of four textural parameters of stem, calyx, bruise and mildew

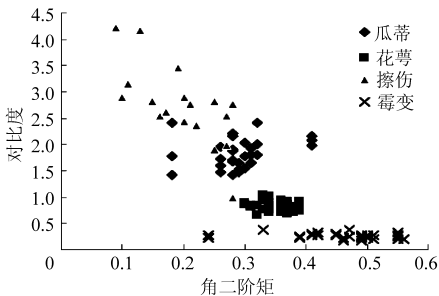


图 4 对比度和角二阶矩特征参数分布

Fig. 4 Samples distribution of contrast and energy of stem, calyx, bruise and mildew

3.3 颜色特征实验结果

基于对比度和角二阶矩两个特征参数的聚类分析得出的擦伤和瓜蒂聚类效果不佳,现引入颜

色特征。提取训练样本中可疑区每个像素点的 $I_1 \sim I_{12}$ 共 12 个颜色参数,利用直方图分别统计瓜蒂、花萼、擦伤、霉变分布情况。统计结果表明,参数 I_1 、 I_2 、 I_4 、 I_5 、 I_6 、 I_7 、 I_8 和 I_9 直方图中甜瓜各个部分之间峰值距离较近,且有较大部分的重合交叠。

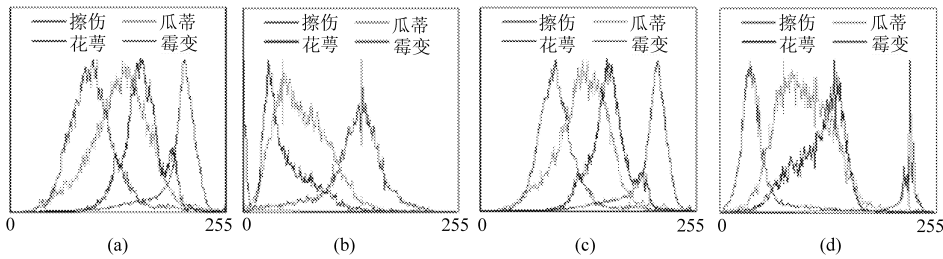


图5 甜瓜图片中擦伤、瓜蒂、花萼、霉变统计直方图
Fig. 5 Histograms of stem, calyx, bruise and mildew
(a) I_3 (b) I_{10} (c) I_{11} (d) I_{12}

3.4 分类识别效果

分类实验在 Matlab 2007b 环境进行,利用了 LibSVM 工具箱。进行分类时主要用到两个函数: svmtrain 函数根据训练样本建立分类器模型, svmpredict 函数对测试样本进行分类。以上 2 个纹理特征与 4 个颜色特征组成的特征向量作为分类器的输入。测试样本中含有 SC 区域共 65 处、擦伤 35 处、霉变 18 处。实验发现在均方差 (MSE) 最小时,惩罚因子 $C = 100$,松弛系数 $\xi = 0.001$ 。选择适当的核函数,可以提高系统的识别率和识别速度,支持向量机 3 种核函数的比较,如表 1 所示。

表 1 不同核函数 SVM 分类实验结果
Tab. 1 Results of SVM classifications by different kernels

核函数	参数	平均运行	平均识别
		时间 /ms	正确率 /%
$(xy+1)^q$	$q=2$	440	91.2
	$q=3$	450	91.9
	$q=4$	440	92.2
	$q=5$	450	91.6
$\exp\left(-\frac{\ x-y\ ^2}{2\sigma^2}\right)$	$\sigma^2=1$	460	90.5
	$\sigma^2=3$	470	90.8
	$\sigma^2=5$	470	90.7
	$v=0.6$	390	81.2
$\tanh[v(xy)+c]$	$c=0.5$		

从表中可以看出,基于纹理特征和颜色特征为综合特征,对甜瓜瓜蒂、花萼、擦伤和霉变进行分类识别效果较好。从识别正确率看,多项式核函数最好,径向基核函数次之,Sigmiod 核函数最差;但从运行时间比较,Sigmiod 核函数比多项式核函数和径向基核函数运行时间短;从参数上看,多项式核函数在 $q=4$ 分类效果最好,径向基 $\sigma^2=3$ 时效果最好;综合考虑选用多项式核函数参数 $q=4$ 的基于综合特

而在参数 I_3 、 I_{10} 、 I_{11} 和 I_{12} 直方图中(图 5),甜瓜各个部分之间峰值距离较远,且重合部分较少,因此,筛选 I_3 、 I_{10} 、 I_{11} 和 I_{12} 等 4 个颜色参数,计算可疑区域各个像素点关于这些参数的平均值作为颜色特征。

征分类识别效果最好。 $q=4$ 时,多项式核函数 SVM 对测试样本进行分类结果如表 2 所示,SC 区域总数 65 个,错误识别数 6 个,错把瓜蒂判断为擦伤;擦伤总数 35 个,错误识别数为 3 个,错把擦伤判断为瓜蒂;霉变总数 18 个,错误识别数为 1 个,错把霉变判断为擦伤。

表 2 多项式核函数参数 $q=4$ 时 SVM 分类结果
Tab. 2 Experimental results of SVM with polynomial kernel and $q=4$

	总数	正确识别数	错误识别数	正确识别率/%
SC 区域	65	59	6	90.7
擦伤	35	32	3	91.4
霉变	18	17	1	94.4

3.5 实验分析

基于灰度与颜色为综合特征,并应用支持向量机作为分类方法对甜瓜缺陷进行识别,获得满意的识别正确率,当擦伤或者霉变区域和 SC 区域重合或者相隔距离近时,利用 Otsu 算法进行可疑区域分割时很难把两者分开,导致无法分类。可疑区域特征提取向量较多,需要分别提取纹理特征 2 个变量和颜色特征 4 个变量,尤其是纹理特征提取平均耗时 182 ms,以后可以改进算法提高特征提取速度。

4 结束语

构建了一个硬皮甜瓜的缺陷区域与 SC 区域的分类识别系统,通过 Otsu 算法提取了甜瓜上的可疑区域,然后计算可疑区灰度共生矩阵,在此基础上提取了可疑区域的角二阶矩、熵、对比度和局部均匀性等 4 个纹理参数,经过实验比较得出对比度和角二阶矩两个参数在甜瓜瓜蒂、花萼、擦伤和霉变上有明显的可区分性,但通过聚类分析得出擦伤和瓜蒂

有交叉重叠部分。为了提高分类的正确性,另外引入了 12 种颜色特征,通过颜色分量 R 、 G 、 B 及其组合的直方图分析,得出其中的 I_3 、 I_{10} 、 I_{11} 和 I_{12} 直方图中甜瓜各个部分之间峰值距离较远,有较好的区分性。最终形成的综合分类特征向量由以上 2 个灰度纹理特征与 4 个颜色特征构成。实验结果表明,基于综合特征的支持向量机识别分类器对于测试样本有较高的识别率。

参 考 文 献

1 Wolfe R, Sandler W. An algorithm for detection using digital image analysis[J]. Transactions of the ASAE, 1985, 28 (2): 641 ~ 644.

2 徐娟. 水果分级中计算机视觉信息并行处理技术的研究[D]. 北京:中国农业大学,1997.
Xu Juan. Study on parallel processing for computer vision information in fruit grading[D]. Beijing: China Agricultural University, 1997. (in Chinese)

3 应义斌, 景寒松, 马俊福, 等. 机器视觉技术在黄花梨尺寸和果面缺陷检测中的应用[J]. 农业工程学报, 1999, 15(1): 197 ~ 200.
Ying Yibin, Jing Hansong, Ma Junfu, et al. Application of machine vision to detecting size and surface defect of Huanghua pear[J]. Transactions of the CSAE, 1999, 15(1): 197 ~ 200. (in Chinese)

4 何东健, 杨青, 薛少平, 等. 果实缺陷面积的计算机视觉测定研究[J]. 农业工程学报, 1997, 12(4): 156 ~ 160.
He Dongjian, Yang Qing, Xue Shaoping, et al. Computer vision method for measuring defect area of sphere-fruits[J]. Transactions of the CSAE, 1997, 12(4): 156 ~ 160. (in Chinese)

5 李庆中, 汪懋华. 基于分形特征的水果缺陷快速识别方法[J]. 中国图象图形学报, 2000, 5(2): 144 ~ 148.
Li Qingzhong, Wang Maohua. A fast identification method for fruit surface defect based on fractal characters[J]. Journal of Image and Graphics, 2000, 5(2): 144 ~ 148. (in Chinese)

6 冯斌, 汪懋华. 计算机视觉识别水果缺陷新方法研究[J]. 中国农业大学学报, 2002, 7(4): 73 ~ 76.
Feng Bin, Wang Maohua. Study on identifying measurement about default of fruit in computer vision[J]. Journal of China Agricultural University, 2002, 7(4): 73 ~ 76. (in Chinese)

7 蒋益女, 徐从富. 结合纹理分析和 SVM 的苹果梗蒂和缺陷识别方法[J]. 计算机工程与应用, 2008, 44(1): 235 ~ 237.
Jiang Yinü, Xu Congfu. Apple stem and calyx identification based on texture analysis and SVM[J]. Computer Engineering and Applications, 2008, 44(1): 235 ~ 237. (in Chinese)

8 Haralick R M, Shanmugam K, Din Stein I. Texture features for image classification[J]. IEEE Transactions on Systems, Man and Cybernetics, 1973, 3(6): 610 ~ 621.

9 王树文, 张长利, 房俊龙. 基于计算机视觉的番茄损伤自动检测与分类研究[J]. 农业工程学报, 2005, 21(8): 98 ~ 101.
Wang Shuwen, Zhang Changli, Fang Junlong. Automatic identification and classification of tomatoes with bruise using computer vision[J]. Transactions of the CSAE, 2005, 21(8): 98 ~ 101. (in Chinese)

10 Leemans V, Destain M F. A real-time grading method of apples based on features extracted from defects[J]. Journal of Food Engineering, 2004 (61): 83 ~ 89.

11 邓乃扬, 田英杰. 数据挖掘中的新方法: 支持向量机[M]. 北京: 科学出版社, 2004.

12 张建华, 王书志, 冯全. 一种计算机视觉识别硬皮甜瓜瓜蒂新方法[J]. 计算机系统应用, 2009(7): 186 ~ 189.
Zhang Jianhua, Wang Shuzhi, Feng Quan. A new algorithm for recognizing stem of melon with computer vision[J]. Computer Systems Applications, 2009(7): 186 ~ 189. (in Chinese)

13 陈树人, 沈宝国, 毛罕平, 等. 基于颜色特征的棉田中铁苋菜识别技术[J]. 农业机械学报, 2009, 40(5): 149 ~ 152.
Chen Shuren, Shen Baoguo, Mao Hanping, et al. Copperleaf herb detection from cotton field based on color feature[J]. Transactions of the Chinese Society for Agricultural Machinery, 2009, 40(5): 149 ~ 152. (in Chinese)

14 杨福增, 杨亮亮, 田艳娜, 等. 基于颜色和形状特征的茶叶嫩芽识别方法[J]. 农业机械学报, 2009, 40(增刊): 119 ~ 123.
Yang Fuzeng, Yang Liangliang, Tian Yanna, et al. Recognition of the tea sprout based on color and shape features[J]. Transactions of the Chinese Society for Agricultural Machinery, 2009, 40(Supp.): 119 ~ 123. (in Chinese)