

基于亮度校正和 AdaBoost 的苹果缺陷在线识别*

张保华¹ 黄文倩² 李江波² 赵春江² 刘成良¹ 黄丹枫¹

(1. 上海交通大学机械系统与振动国家重点实验室, 上海 200240;

2. 北京市农林科学院北京农业智能装备技术研究中心, 北京 100097)

摘要: 提出了一种基于亮度校正和 AdaBoost 的苹果缺陷与果梗-花萼在线识别方法。以富士苹果为研究对象, 首先在线采集苹果的 RGB 图像和 NIR 图像, 并分割 NIR 图像获得苹果二值掩模; 其次利用亮度校正算法对 R 分量图像进行亮度校正, 并分割校正图像获得缺陷候选区(果梗、花萼和缺陷); 然后以每个候选区域为掩模, 随机提取其内部 7 个像素的信息分别代表所在候选区的特征, 将 7 组特征送入 AdaBoost 分类器进行分类、投票, 并以最终投票结果确定候选区的类别。实验结果表明, 该算法检测速度为 3 个/s, 满足分选设备的实时性要求, 且总体正确识别率达 95.7%。

关键词: 苹果缺陷 机器视觉 亮度校正 AdaBoost 在线识别 果梗-花萼

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 1000-1298(2014)06-0221-06

引言

苹果的外观特征是其品质最直接的反映, 且表面缺陷在一定程度上还会影响内部品质。在交易或存储前, 按照苹果外观品质进行分选不仅可以实现优质优价, 而且剔除缺陷苹果可以有效防止其感染其他优质水果。现在已经有运用机器视觉技术对水果的大小、质量、颜色等进行分选的商业化水果分选生产线。苹果的边缘、果梗、花萼、缺陷在图像中都呈现较低的亮度, 容易造成缺陷误识别, 因此对苹果缺陷的检测还是以人工为主^[1-6]。这些方法中, 基于高光谱、多光谱成像技术的缺陷识别方法不仅相机结构复杂、价格昂贵, 而且图像采集速度慢, 处理方法也相对复杂; 逐像素对水果表面图像缺陷和正常区域进行识别分类不适合水果品质快速在线检测; 从水果自身而言, 由于朗伯现象的存在, 水果边缘区域的亮度较低, 不利于缺陷分割。

针对以上问题, 本文提出一种基于亮度校正和 AdaBoost 的苹果缺陷与果梗-花萼在线识别方法, 首先对图像进行亮度校正, 然后单阈值分割校正图像获得缺陷候选区, 然后以每个缺陷候选区为一个整体, 提取其特征, 运用 AdaBoost 分类器实现缺陷与果梗-花萼的识别。

1 材料与方法

1.1 实验材料

以双色的红富士苹果为研究对象, 苹果样本购自北京当地水果市场, 共计 152 个, 挑选其中 140 个作为实验样本, 挑选过程中尽可能使样本的形状、大小、缺陷分布均匀, 其中正常果 80 个, 其余 60 个为带有各类常见缺陷的缺陷果。样品中的缺陷种类主要包括: 腐烂、风伤、虫伤、疤痕、雹伤等。

1.2 在线检测系统

自动检测所用的分选机示意图如图 1 所示, 系统主要包括: 上料装置、传输机构、计算机控制模块、照明装置、视觉检测系统和卸料装置等部分。

检测流程为: 苹果上料, 单列置于传输滚子上随传送链条传动, 光电开关检测到苹果, 发出信号并通过串口传送至计算机, 计算机检测软件控制相机采集苹果图像, 并运用本文算法完成苹果缺陷与果梗-花萼的识别。

其中视觉检测系统由计算机、相机和 VS2008 开发的检测与控制软件组成, 相机为可见-近红外双 CCD 相机, 型号为: AD-080 GE 2 CCD Multi-spectral, 计算机配置为: Dell, Inter Core i5-2400 CPU @ 3.10 GHz, 内存 4.0 GB; 照明装置由一对可

收稿日期: 2013-05-18 修回日期: 2013-06-16

* 国家自然科学基金资助项目(31301236)、国家高技术研究发展计划(863 计划)资助项目(2013AA100307)和 2012 年北京市农林科学院博士后基金资助项目

作者简介: 张保华, 博士生, 主要从事农产品品质与安全无损检测研究, E-mail: zhangbaohua@sjtu.edu.cn

通讯作者: 赵春江, 研究员, 主要从事农业机械化与自动化、农业信息化研究, E-mail: zhaocj@nercita.org.cn

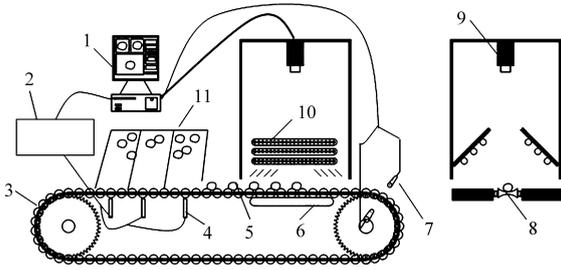


图1 自动检测分选机示意图

Fig.1 Diagram of general scheme

1. 计算机 2. 控制器 3. 传送链条 4. 卸料装置 5. 苹果
6. 摩擦带 7. 光电开关 8. 传输滚子 9. CCD 相机 10. 照明装置 11. 盛料箱

见光 LED 光源及一对中心波长为 820 nm 的近红外 LED 光源组成,LED 光源对称线状纵向分布于传输链条侧上方。

1.3 图像处理原理与方法

1.3.1 图像获取与预处理

利用可见-近红外双 CCD 相机采集苹果的 RGB 图像和 NIR 图像。研究发现,缺陷区域与正常表面在 R 分量上的亮度差异较大,因此提取 R 分量作为亮度校正与缺陷提取的图像。利用 NIR 图像对颜色不敏感和苹果与背景亮度反差较大的特点,单阈值分割 NIR 图像获得掩模,再用掩模去除苹果 R 分量图像的背景。

1.3.2 朗伯现象与亮度校正

苹果可以近似看作一个朗伯体^[9],由朗伯反射定律知:球面上任意点的亮度 I_D 等于入射光强度 I_L 乘以该点法向量与该点与光源连线间的夹角 θ 的余弦值,即

$$I_D = I_L \cos\theta \quad (1)$$

在检测时,由于相机与水果的距离远大于水果的尺寸,因此引起水果表面亮度不均的主要原因是夹角 θ 。朗伯反射模型如图 2a 所示。由图可知水果上表面各点的夹角 θ_1 小于侧面各点的夹角 θ_2 ,故水果表面的亮度呈现中间亮边缘暗的分布。

图 2b 为亮度校正算法的示意图。在图像处理时可以近似认为以水果上表面顶点为圆心,半径为 r 、宽度为 Δr 的圆环 A 内的所有像素的亮度是均匀的。

圆环 A 内像素的平均亮度为

$$I_M = \frac{1}{N} \sum_{i \in A} I_i \quad (i=0,1,\dots,N) \quad (2)$$

式中 I_M ——圆环区域 A 的平均亮度

I_i ——圆环 A 内像素点 i 的亮度

N ——圆环 A 内像素数目

圆环 A 内像素点 i 校正后的亮度为

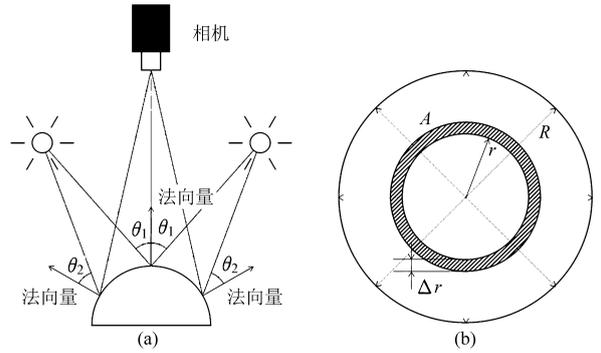


图2 朗伯模型示意图和校正原理图

Fig.2 Diagram of Lambertian reflection model and correction method

(a) 朗伯模型 (b) 校正原理示意图

$$I_{R_i} = 255 \frac{I_i}{I_M} \quad (i=0,1,\dots,N) \quad (3)$$

为提高校正精度,取 $\Delta r = 1$,即同心圆环退化为宽度只有 1 个像素的圆。在实现过程中本文重点改进了文献[6]的方法,本文校正过程为:①视觉检测系统在线采集苹果的 RGB 图像和 NIR 图像,并提取 R 分量作为亮度校正图像。②分割 NIR 图像获得苹果掩模,进而求取苹果最小外接圆的圆心 (x,y) 及半径 R 。③以 (x,y) 为圆心,构造起始半径 $r = 1$ 的圆形掩模,并提取 R 分量图像的一圈,利用式(2)和式(3)对该圈的图像亮度校正。④累加步骤③获得的单圈校正图像,并使圆形掩模半径 r 增加 1 个像素,重复步骤③,直至 $r = R$,此时便获得苹果的亮度校正图像。

1.3.3 缺陷候选区提取

经过亮度校正后,缺陷区域仍保持较低亮度,而正常表面的亮度则被提升至高亮,亮度校正提高了缺陷区域和正常表面的对比度,并且有效补偿了水果边缘区域的亮度,解决了过分分割问题,有利于缺陷尤其是边缘部位缺陷的提取。本文使用单阈值分割的方法提取缺陷候选区。

1.4 候选区特征选择与提取

选择候选区内像素的颜色、纹理和区域的方差信息作为候选区的特征。其中颜色特征包括 R 、 G 、 B 通道值, H 、 I 、 S 通道值,颜色特征可以直接从图像中获取。

纹理特征是基于 R 通道单色图像构造的局部灰度共生矩阵求取的,包括熵、能量、相关、惯性矩和局部平稳性特征。首先以目标像素为中心构造 $k \times k$ 的邻域区域;其次以该区域像素的 R 通道灰度构建灰度共生矩阵;最后基于局部灰度共生矩阵计算 5 个纹理特征:熵 E_i 、能量 E_n 、相关 C_r 、惯性矩 I_m 和局部平稳性 L_s 作为目标像素的纹理特征。纹理特

征的具体计算式为

$$\begin{cases} E_t = - \sum_{i=1}^k \sum_{j=1}^k G(i,j) \lg G(i,j) \\ E_n = \sum_{i=1}^k \sum_{j=1}^k (G(i,j))^2 \\ C_r = \sum_{i=1}^k \sum_{j=1}^k \frac{ijG(i,j) - u_i u_j}{S_i S_j} \\ I_m = \sum_{i=1}^k \sum_{j=1}^k (i-j)^2 G(i,j) \\ L_s = \sum_{i=1}^k \sum_{j=1}^k \frac{G(i,j)}{1 + (i-j)^2} \end{cases} \quad (4)$$

其中

$$u_i = \sum_{i=1}^k \sum_{j=1}^k iG(i,j)$$

$$u_j = \sum_{i=1}^k \sum_{j=1}^k jG(i,j)$$

$$S_i = \sqrt{\sum_{i=1}^k \sum_{j=1}^k G(i,j) (i - u_i)^2}$$

$$S_j = \sqrt{\sum_{i=1}^k \sum_{j=1}^k G(i,j) (j - u_j)^2}$$

式中 $G(i,j)$ ——灰度共生矩阵元素

k ——灰度共生矩阵的维数

值得注意的是,在求取纹理特征前,为了降低尖锐噪声的影响,需要对图像进行平滑处理,本文采用 3×3 圆盘形高斯模板对图像进行平滑处理。

除此之外,本文还提取了候选区域的方差信息,其计算式为

$$S^2 = \sum_{i=1}^N \sum_{j=1}^N \frac{(t_i - \bar{t})^2}{M} \quad (5)$$

式中 M ——缺陷候选区像素的数目

t_i ——像素灰度

\bar{t} ——像素灰度的平均值

1.5 AdaBoost 算法

AdaBoost 是一种迭代算法,基本思想是通过训练训练集数据得出几种简单的弱分类器,然后把把这些弱分类器联合起来,提升为最终的强分类器。

设有一组训练集数据为: $(x_1, y_1), (x_2, y_2), \dots,$

(x_m, y_m) ,其中 x_i 表示样本特征, y_i 表示与之对应的样本标号,其中 $y_i \in \{0,1\}$, $y_i = 0$ 时, x_i 表示负样本; $y_i = 1$ 时, x_i 表示正样本。算法的实现过程为:

(1) 样本权重初始化。

设正样本数目为 n_1 ,负样本数目为 n_2 ,则正样本的权值初始化为 $w_1(i) = 1/n_1$,负样本的权值初始化为 $w_1(i) = 1/n_2$ 。

(2) 训练 T 个弱分类器。

对于 $t = 1, 2, \dots, T$,执行以下步骤:

① 训练弱分类器 $f_t(x)$ 。

$$f_t(x) = L(w, w_t) \in \{0,1\} \quad (6)$$

② 计算分类器 $f_t(x)$ 的分类错误率 e_t 。

$$e_t = \frac{1}{2} \sum_{i=1}^m w_t(i) |f_t(x_i) - y_i| \quad (7)$$

③ 计算分类器 $f_t(x)$ 的权重 α_t 。

$$\alpha_t = \ln \frac{1 - e_t}{e_t} \quad (8)$$

④ 调整样本的权重,并归一化。

权重调整

$$w_{t+1}(i) = \begin{cases} w_t(i) & (f_t(x_i) = y_i) \\ w_t(i) e(\alpha_t) & (f_t(x_i) \neq y_i) \end{cases} \quad (9)$$

权重归一化

$$w_{t+1}(i) = \frac{w_{t+1}(i)}{\sum_{j=1}^m w_{t+1}(j)} \quad (i = 1, 2, \dots, m) \quad (10)$$

(3) 强分类器 $C(x)$ 为

$$C(x) = \text{sign} \left(\sum_{i=1}^T \alpha_i f_i(x) \right) \quad (11)$$

通过以上步骤便可以设计出高精度的 AdaBoost 强分类器,且分类器不需要特征筛选,也不会出现过拟合现象。本文使用 AdaBoost 分类器识别苹果的缺陷和果梗-花萼。具体算法步骤包括图像获取、亮度校正、缺陷候选区分割和识别分选,流程如图 3 所示。

2 结果与讨论

2.1 亮度校正与缺陷候选区提取

图 4 为亮度校正前、后缺陷候选区的提取结果。

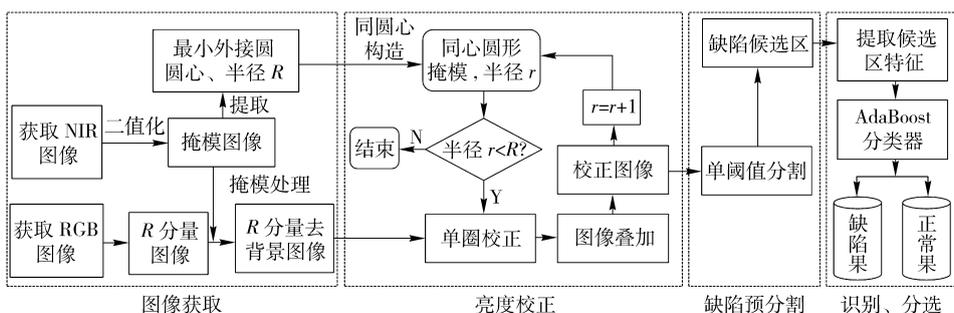


图 3 算法流程图

Fig. 3 Flowchart of algorithm

其中,图4a是未经亮度校正的图像,图4b为4a中灰线处像素点的亮度曲线。可以看出,在苹果的边缘亮度较低,甚至与缺陷区域亮度接近,因此使用 $T=33$ 的阈值进行缺陷提取时,存在过分割现象,分割结果如图4c所示。

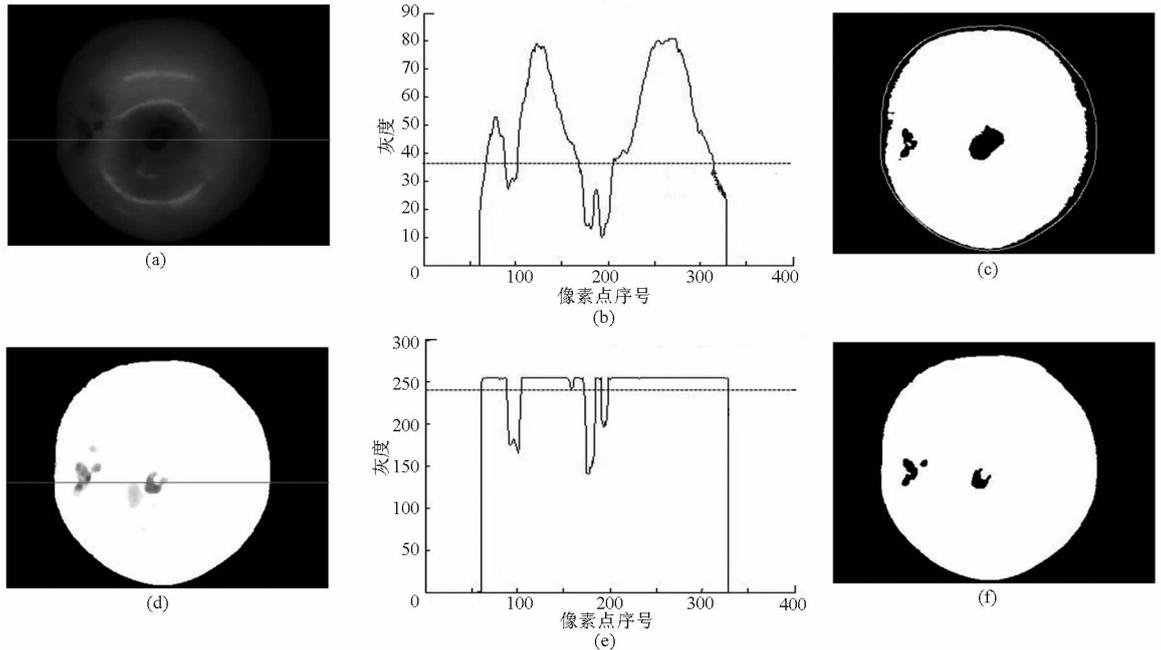


图4 亮度校正效果图

Fig. 4 Results of defect candidate regions extraction before and after lightness correction

(a) 去除背景的 R 分量图像 (b) 图4a中灰线处像素点的亮度曲线 (c) 图4a所提取的缺陷候选区分割结果
(d) 图4a亮度校正图像 (e) 图4d中灰线处像素点的亮度曲线 (f) 图4d所提取的缺陷候选区分割结果

2.2 缺陷与果梗-花萼的识别分类

为了测试算法的有效性,在线采集苹果样本的图像,离线进行了缺陷与果梗-花萼的识别分类。此阶段,以每个缺陷候选区为一个整体,提取候选区内所有像素的RGB空间和HIS空间的各分量的平均值作为区域的颜色特征,提取其内部5像素×5像素邻域内的熵、能量、相关、惯性矩以及局部平稳性的平均值作为区域的纹理特征,同时提取 R 分量图像中候选区的方差信息,然后以区域为单位,运用AdaBoost分类器进行分类。识别结果如图5所示。

对50个样本进行检测,正确识别率为96%。结果表明算法可以有效区分苹果的缺陷和果梗-花萼,但处理每一个苹果图像的时间为700~1400ms(所选择的测试图像的尺寸为640像素×480像素,存在缺陷候选区1~5个,每个候选区的面积为300~1500像素之间),算法的实时性较差,不利于苹果的快速在线检测。

2.3 苹果缺陷的在线检测

为了提高算法的实时性,实现苹果缺陷的快速在线检测,本文对上述算法进行了改进。对于每一个候选区,随机提取其内部的7个像素点的颜色、

图4d为4a的亮度校正图像,图4e为4d中灰线处像素点的亮度曲线。可以看出,经过校正,缺陷仍保持较低亮度,而正常表皮则呈现高亮。使用 $T=235$ 的阈值进行缺陷提取时,边缘没有出现过分割现象,分割结果如图4f所示。

5×5邻域的纹理特征及像素的方差信息(7个像素的方差分别赋予每个像素),每个像素的信息分别代表所在候选区的特征,把这7组特征送入分类器进行分类、投票,并以得票较高的类别作为候选区的最终类别。不足7个像素的区域则选取所有像素进行分类、投票。经过改进,算法处理一个水果图像(测试样本图像同2.2节)的时间小于320ms,实时性得到较大提高。对测试样本进行检测,检测结果如表1所示。

由实验结果可知,正常果的平均检测正确率为97.5%,缺陷果的检测正确率为93.3%,总体检测正确率为95.7%。误检主要与果梗-花萼在图像中呈现的特征与缺陷太相似、分类算法训练不完善等原因有关。为了提高准确率,后续研究可以增加训练样本数量,或采用高光谱相机寻找对缺陷或者果梗-花萼更为敏感的波段图像,结合本文算法进行缺陷检测。

3 结论

(1)为了解决因亮度不均造成的缺陷候选区的过分割问题,改进了文献[6]中的亮度校正方法。

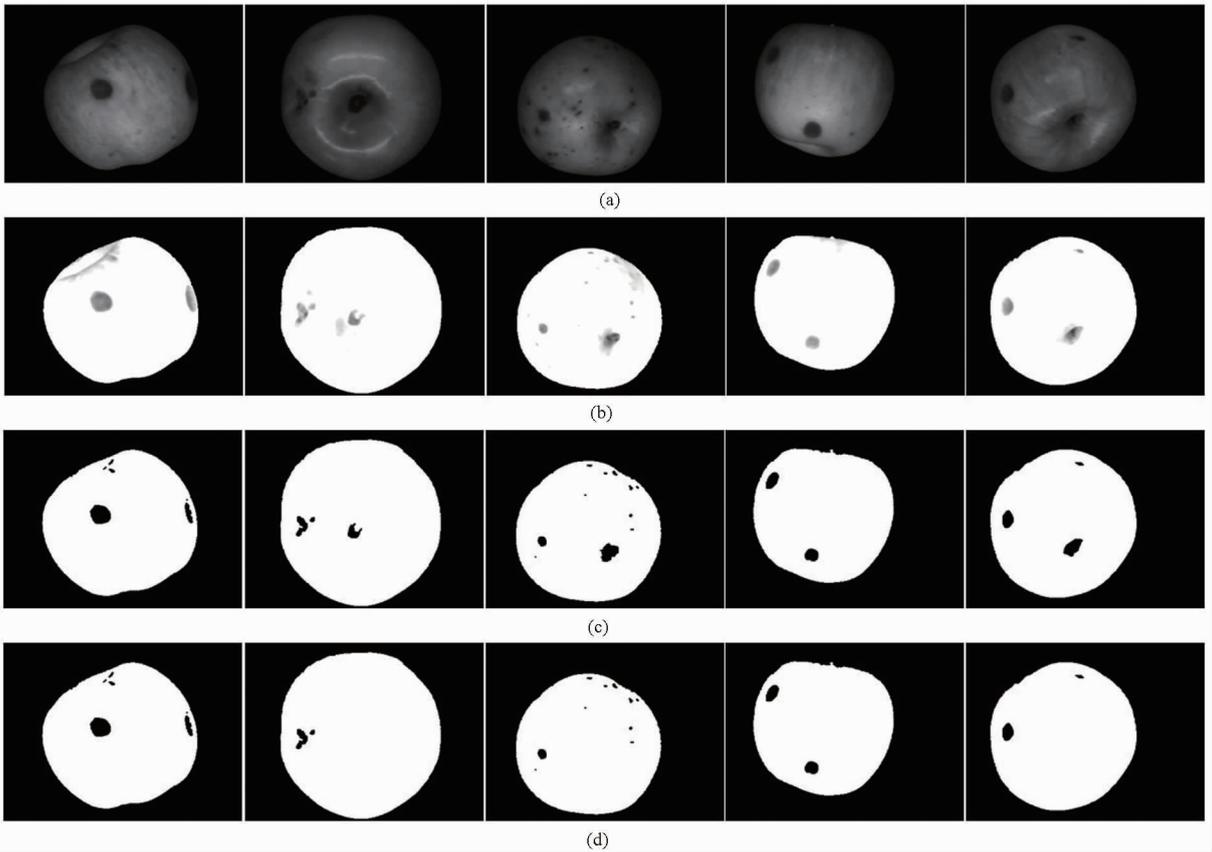


图 5 识别结果图

Fig. 5 Results of defects detection

(a) 去背景的 R 分量图像 (b) 亮度校正图像 (c) 缺陷候选区阈值分割图像 (d) 缺陷识别后的图像

表 1 检测结果

Tab. 1 Detection results

| 类型 | 样本类型 | 样本数目 | 检测结果/个 | | 检测正确率/% |
|-----|---------|------|--------|----|---------|
| | | | 正确 | 错误 | |
| 正常果 | 含有果梗-花萼 | 23 | 21 | 2 | 91.3 |
| | 不含果梗-花萼 | 57 | 57 | 0 | 100 |
| 缺陷果 | 含有果梗-花萼 | 18 | 16 | 2 | 88.9 |
| | 不含果梗-花萼 | 42 | 40 | 2 | 95.2 |
| 合计 | | 140 | 134 | 6 | 95.7 |

本文方法不受苹果大小、形状、位置的影响。亮度校正后使用单阈值便可简单有效地提取缺陷候选区。

(2)为了验证特征的有效性,以每个连通的候选区为一个整体,提取其内部所有像素的颜色、纹理

的平均值和方差信息作为候选区的特征,运用 AdaBoost 分类器进行分类,缺陷识别率达到 96%,所选特征有效地完成了苹果缺陷与果梗-花萼的识别分类。

(3)为了提高算法实时性,以每个连通的候选区为一个整体,随机提取候选区内的 7 个像素的信息分别代表候选区的特征,并送入 AdaBoost 分类器进行分类、投票,以投票结果作为每个候选区的最终类别,高效地完成苹果缺陷与果梗-花萼的在线识别。

(4)实验结果表明,本文算法处理每个苹果的时间在 320 ms 以内,检测的总体正确率为 95.7%,适合于苹果缺陷的快速在线检测。

参 考 文 献

- 1 Devrim Unay, Bernard Gosselin, Olivier Kleynen, et al. Automatic grading of Bi-color apples by multispectral machine vision [J]. Computers and Electronics in Agriculture, 2011, 75(1): 204 - 212.
- 2 Bennedsen B S, Peterson D L, Amy Tabb. Identifying defects in images of rotating apples [J]. Computers and Electronics in Agriculture, 2005, 48(2):92 - 102.
- 3 Zou Xiaobo, Zhao Jiewen, Li Yanxiao, et al. In-line detection of apple defects using three color cameras system [J]. Computers and Electronics in Agriculture, 2010, 70(1):129 - 134.
- 4 Mehl Patrick M, Chen Yud-Ren, Kim Moon S, et al. Development of hyperspectral imaging technique for the detection of apple surface defects and contaminations [J]. Journal of Food Engineering, 2004, 61(1):67 - 81.

- 5 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(3):83–89.
- 6 黄文倩,李江波,张驰,等. 基于类球形亮度变换的水果表面缺陷提取[J]. *农业机械学报*,2012,43(12):187–191.
Huang Wenqian, Li Jiangbo, Zhang Chi, et al. Detection of surface defects on fruits using spherical intensity transformation [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2012, 43(12):187–191. (in Chinese)
- 7 Unay D, Gosselin B. Stem and calyx recognition on ‘Jonagold’ apples by pattern recognition [J]. *Journal of Food Engineering*, 2007, 78(2):597–605.
- 8 Throop J A, Aneshansley D J, Anger W C, et al. Quality evaluation of apples based on surface defects: development of an automated inspection system [J]. *Postharvest Biology and Technology*, 2005, 36(3):281–290.
- 9 Gomez-Sanchis J, Molto E, Camps-Valls G, et al. Automatic correction of the light source on spherical objects. An application to the analysis of hyperspectral images of citrus fruits [J]. *Journal of Food Engineering*, 2008, 85(2):191–200.
- 10 李江波,饶秀勤,应义斌,等. 基于掩模及边缘灰度补偿算法的脐橙背景及表面缺陷分割[J]. *农业工程学报*,2009, 25(12):133–137.
- 11 李江波,饶秀勤,应义斌. 水果表面亮度不均校正及单阈值缺陷提取研究[J]. *农业机械学报*,2011,42(8):159–163.
Li Jiangbo, Rao Xiuqin, Ying Yibin. Correction algorithm of illumination nonuniformity on fruit surface and defects extraction using single threshold value[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2011, 42(8):159–163. (in Chinese)
- 12 Diaz R, Faus G, Blasco M, et al. The application of a fast algorithm for the classification of olives by machine vision [J]. *Food Research International*, 2000, 33(3~4):305–309.
- 13 Diaz R, Gil L, Serrano C, et al. Comparison of three algorithms in the classification of table olives by means of computer vision [J]. *Journal of Food Engineering*, 2004, 61(1):101–107.
- 14 Tao Y, Wen Z. An adaptive spherical image transform for high-speed fruit defect detection [J]. *Transactions of the ASAE*, 1999, 42(1):241–246.
- 15 Patel K K, Kar A, Jha S N, et al. Machine vision system: a tool for quality inspection of food and agricultural products [J]. *Journal of Food Science and Technology*, 2012, 49(2):123–141.
- 16 Juliana Freitas Santos Gomes, Fabiana Rodrigues Leta. Application of computer vision techniques in the agriculture and food industry: a review [J]. *European Food Research and Technology*, 2012, 235:989–1000.

On-line Identification of Defect on Apples Using Lightness Correction and AdaBoost Methods

Zhang Baohua¹ Huang Wenqian² Li Jiangbo² Zhao Chunjiang² Liu Chengliang¹ Huang Danfeng¹
(1. *State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, Shanghai 200240, China*
2. *Beijing Research Center of Intelligent Equipment for Agriculture, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China*)

Abstract: An algorithm was proposed to on-line identify the defects and stem – calyx on apples based on lightness correction method and AdaBoost classifier. The ‘Fuji’ apples were selected as the experiment object. First, the RGB images and NIR images of apples were acquired, and NIR images were binarized to obtain the mask images. Second, the *R* component images were corrected by using proposed lightness correction algorithm and the defect candidate regions were obtained by binarizing the corrected images with a single threshold. Third, every candidate region was treated as a mask, and the information of random seven pixels in the candidate region were selected as the characteristics of the selected candidate region. Finally, an AdaBoost classifier was used to classify these candidate regions by voting way, and the category of candidate region was determined according to the final voting results. For the investigated 140 samples, the results with a 95.7% overall detection rate under acquisition speed of three apples per second indicated that the proposed algorithm was effective.

Key words: Apple defects Machine vision Lightness correction AdaBoost On-line identification Stem – calyx