

基于多示例学习的超市农产品图像识别

罗承成, 李书琴*, 唐晶磊

(西北农林科技大学 信息工程学院, 陕西 杨凌 712100)

(*通信作者电子邮箱 lsq_cic@nwsuaf.edu.cn)

摘要:为解决超市农产品价格需依靠人工记忆的问题,实现农产品的智能识别,提出了基于多示例学习的农产品图像识别方法。采用改进的单色块及其邻域算法(SBN)特征提取算法将训练样本组织成多示例包,利用多样性密度算法对正包和反包进行多示例学习,根据多样性密度最大化模型对测试样本进行识别。分别在自采集的多类别果蔬图像集以及 Amsterdam 图像库中的单类别果蔬图像上进行测试。结果表明该方法能够识别不同光照、存在干扰物的环境背景下,以任意方式摆放的多类别混合果蔬图像,识别率最高达到 94.21%,且对于单类别果蔬图像的识别优于全局方法。因此利用基于多示例学习的图像识别方法对超市农产品的自动售卖提供辅助具有可行性。

关键词:超市农产品;图像处理;模式识别;多示例学习;特征提取

中图分类号: TP391.413 **文献标志码:** A

Image recognition of agricultural products in supermarket based on multi-instance learning

LUO Cheng-cheng, LI Shu-qin*, TANG Jing-lei

(College of Information Engineering, Northwest A&F University, Yangling Shaanxi 712100, China)

Abstract: The pricing of agricultural products in supermarket needs to rely on artificial memory. In order to realize intelligent recognition of agricultural products, an image recognition method of agricultural products based on the multi-instance learning was proposed. An improved Single Blob with Neighbors (SBN) method was adopted to organize bags and meanwhile extract features of an image. The target concept was learned by maximizing Diverse Density (DD) and applied to images recognition. Experiments were performed on both multi-class produce image dataset by self-collection and single-class agricultural product images selected from Amsterdam Library of Object Images (ALOI). The experimental results show that the method is able to recognize multi-class agricultural product images captured under various illumination conditions and interference environment, and the recognition rate can achieve 94.21 percent. Additionally, the method performs better than global method when recognizing single-class agricultural product images.

Key words: agricultural products in supermarket; image processing; pattern recognition; multi-instance learning; feature extraction

0 引言

近年来,随着计算机视觉技术的发展,基于计算机视觉的农产品品质检测技术在单一品种农产品分级上得到了广泛的应用^[1-3],但是应用到超市售卖的多类农产品识别的研究则较少。

目前,超市售卖普遍采用条形码技术获取商品价格,但类似水果、蔬菜等农产品一方面需要称重,另一方面需要供顾客挑选,因而不便使用固定条形码,价格确认主要依靠人工记忆。这不仅增加了超市培训人员的经济、时间成本,同时,频繁出现售货员需要前往果蔬摆放处确认价格的情况也给购买者带来排长队等候的不便。因此,更加合理、快速的解决途径亟待提出。Veggie Vision 果蔬识别系统^[4]尝试了此方面的研究,采用独特设计的扫描系统,对果蔬进行图像采集,对果蔬的摆放方式有严格要求,通过图像的颜色、纹理信息对果蔬进行识别,主要面临硬件成本和准确性的问题。Rocha 等^[5]在其研究中则建立了单类别果蔬的图像库,并着重研究了特征和分类器的融合技术以便新品种的添加和识别。但以往的这些研究都只关注了单类别果蔬图像的识别,本文采用基于多

示例学习的图像处理技术,提出适合果蔬图像的包生成方法,旨在对混合果蔬图像进行识别,以适应超市果蔬售卖更进一步的需求。同时在实验中,对比了本文方法和以往研究所采用的全局方法在任意摆放方式下对单类别果蔬图像的识别效果。

1 材料与方法

1.1 图像样本的获取

实验样本包括 15 种果蔬,共 5189 幅图像,全部来源于超市售卖品种,包括洋葱、结球甘蓝、香瓜、柚子、橘子、猕猴桃、芋头、茄子、水萝卜、胡萝卜、荸荠、梨、苹果、青椒、巴西柑等。根据实验需求按以下原则进行图像采集:1)对每个单一品种采集 150 张以上图片;2)对每一品种与任意其他品种的混合方式采集 120 张以上图片;3)每张图片果蔬数量与摆放方式任意,果蔬之间允许粘连或分离;4)在一定数量的图片上设置类似红色胶带、塑料袋等干扰物以模拟超市环境;5)采用 3 种以上光照条件进行采集。

用于图片采集的设备如下:图像采用 CCD 彩色数码相机(型号:Olympus FE280)获取,分辨率采用 2048×1536。相机

收稿日期:2011-11-10;修回日期:2012-01-17。 基金项目:陕西省烟草重大科技专项(K332021101)。

作者简介:罗承成(1986-),女,江苏南京人,硕士研究生,主要研究方向:智能信息系统;李书琴(1965-),女,陕西澄城人,教授,主要研究方向:智能信息系统;唐晶磊(1974-),女,河北邢台人,讲师,硕士,主要研究方向:虚拟现实、生物图像识别。

由三脚架固定在离采集台 50 cm 高处,并在镜头右后方(以使得反射光出离画面)设置 60 W 钨丝灯光源,顶部设置日光灯光源,采集时,分别采用两种单一光源、混合光源以及自然光拍摄。图 1 所示为部分图像样本。

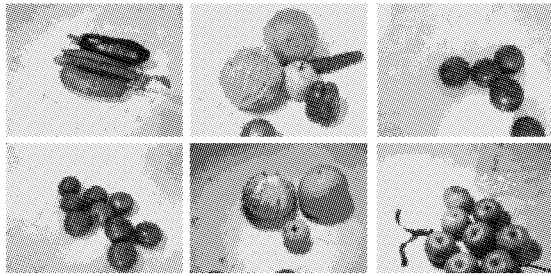


图1 自采集超市农产品图像样本

1.2 基于多示例学习的图像识别方法

多示例学习最初由 Dietterich 等^[6]提出,用以解决药物活性预测问题。在多示例学习中,训练集由有标记的包组成,每个包包含若干没有标记的示例。若一个包中至少有一个正例,则包被标记为正包;若包中所有示例均为负例,则该包被标记为负包。多示例学习不同于监督学习和无监督学习,被界定为半监督学习,解决了以往机器学习研究的盲区。

1.2.1 包生成方法

多示例包生成算法是近年来多示例学习应用于图像分类问题的研究热点,算法好坏直接影响最终的分类效果。目前已有的包生成方法主要分为两大类。一类是固定分割的方法,即对于任何图片使用同样相对位置的子图像产生示例。Maron 等^[7]提出单色块及其邻域算法(Single Blob with Neighbors, SBN),使用十字窗覆盖整幅采样图像以获取示例,提取 RGB 颜色分布信息作为特征,在自然场景图像分类上取得了良好的效果。而 Yang^[8]则认为目标区域仅在图像的有限区域出现,因此,在图像中指定了 20 个尺度不同的固定矩形区域及其镜像区域作为示例,并提取其灰度信息作为特征,在场景图像及物体识别图像上进行了实验。另一类是自适应分割的方法,即根据不同图像自动划分出该图像中的子区域以产生示例,主要以图像分割为基础^[9-10],将分割获得的各个区域作为示例,再对其提取相应的特征,Zhou 等^[11]在其研究中所采用的自组织特征映射(Self-Organizing feature Map, SOM)神经网络聚类包生成方法同 Maron 及 Yang 的方法进行了对比,其效果优于 Yang 方法,而未及 Maron 的 SBN 方法。

首先,农产品图像属于物体识别图像,区别于自然场景图像,有较明显的前景背景,因此,不适宜用示例覆盖整幅图像;其次,针对超市农产品售卖的情况,识别目标有可能出现在图像内任何位置,因此,对所有图像都采用同样的固定区域不一定命中目标;另外,从前人实验的比较结果来看,Yang 以灰度值为特征相比较以颜色、纹理为特征的方法丢失了图像部分信息,难以达到较好的效果。依据以上分析,本文在 Maron 的 SBN 方法基础上提出了一种自适应分割的包生成方法,具体流程如下:

1) 色彩空间转换及预处理。为满足超市售卖处理实时性,首先利用双线性插值算法将 2048×1536 的分辨率的图像缩放成 256×256 的图像。实验比较显示,所采用的分割算法对一幅原分辨率图像的处理时间平均为 0.1117 s,而对一幅采样后的图像处理时间平均为 0.0156 s。考虑光照变化对图像的影响,将图像从 RGB 色彩空间转换到 HSV(色调、饱和度和亮度)空间。

2) 目标矩形区域的自动标注。利用阈值分割算法实现背景区域和前景区域的分离,并用矩形窗口标注出目标区域。本文对传统的 Otsu 阈值分割算法进行改进,避开图像灰度值,利用图像的色彩饱和度 S 通道实现阈值分割。图像二值化之后,利用 10×10 的窗口进行中值滤波,以消除不属于前景的噪点。然后,通过计算前景区域的最小外接矩标注出目标区域,以便有效忽略分割错误的反射光斑。标注过程如图 2 所示。

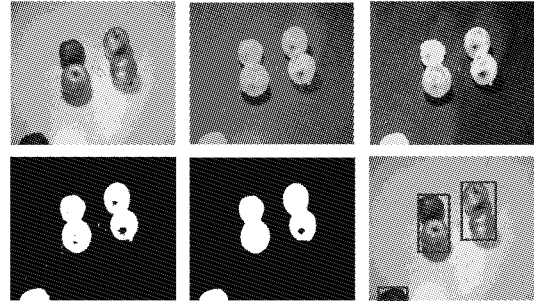


图2 目标区域自动标注过程

3) 颜色分布特征的提取。在 Maron 所提出的单色块及其邻域(SBN)算法中,主要利用了图像的 RGB 颜色分布,以色差作为特征同时反映了图像的颜色及纹理特性。本文将自动划分出的目标区域进一步划分成 $H \times H$ 的区块,以十字窗口覆盖整幅图像,并计算各十字窗口内部的 RGB 色差,共 15 维特征,每个十字窗口作为一个示例,每个包共包含 $(H-5)^2$ 个示例。以划分为 8×8 的区块为例,具体计算方法见图 3。

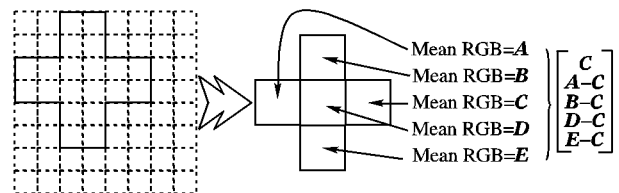


图3 单色块及邻域算法示意图

Maron 在其应用中统一将整幅图像划分成 8×8 的区块以提取特征。在本研究中,对目标区域采用了 $H = 8, 10, 12, 14, 16$ 的划分尺度,一个多示例包分别对应 9, 25, 49, 81, 121 个示例进行实验,实验结果反应了不同尺度的包生成方法对识别效果的影响。

1.2.2 图像识别

在基于多示例学习的图像处理中,通过生成的多示例训练包来寻找目标概念,判断新的图像中是否包含目标概念来进行图像识别。本文采用多样性密度(Diverse Density, DD)算法^[12]来学习目标概念。多样性密度算法的基本思想与多示例学习的概念相契合:在特征空间中,存在某点符合用户感兴趣的概念。每个正包都有至少一个示例接近该点,并且所有负包的示例都远离该点,也即寻找特征空间中的某点,使得该点接近尽可能多的正包,而远离所有负包。这里接近正包是指与正包中的某个或某几个示例接近,远离负包是指远离负包中任意示例。图 4 描述了在二维特征空间多样性密度最大点,其数学定义描述如下。

正包和负包分别表示为 B_1^+, \dots, B_n^+ 和 B_1^-, \dots, B_m^- , B_j^+ 表示 B_i^+ 包中的第 j 个示例,同样的, B_{jk}^+ 表示该示例的第 k 个特征值。要找到多样性密度最大点 t ,也即要求 t 满足式(1):

$$\arg \max_t \Pr(t | B_1^+, \dots, B_n^+, B_1^-, \dots, B_m^-) \quad (1)$$

假设多示例包满足独立同分布,那么根据贝叶斯定理可

以将式(1)转换为乘积项:

$$\arg \max_i \prod_i \Pr(t | B_i^+) \prod_i \Pr(t | B_i^-) \quad (2)$$

采用 noisy-or 模型对上式进行简化:

$$\Pr(t | B_i^+) = 1 - \prod_j (1 - \Pr(B_{ij}^+ = t)) \quad (3)$$

$$\Pr(t | B_i^-) = \prod_j (1 - \Pr(B_{ij}^- = t)) \quad (4)$$

并假设示例满足高斯分布:

$$\Pr(B_{ij} = t) = \exp(-\|B_{ij} - t\|^2) \quad (5)$$

其中 $\|B_{ij} - t\|^2 = \sum_k w_k^2 (B_{ijk} - t_k)^2$ 表示示例与目标之间的欧氏距离。获取多样性密度最大点采用梯度下降法,为降低梯度下降陷入局部极小值而未能命中目标的概率,以一个正包中的每一个示例为梯度下降的起点。

计算出目标概念后,对于测试图像,采用同样的包生成方法,并按照式(5)计算图像为正包的概率以判断图像中是否包含用户感兴趣概念。在本文中,对于多标签的图像,分别训练得出每种标签的目标概念,测试时分别计算图像包含每种标签的概率,并为图像标注概念。

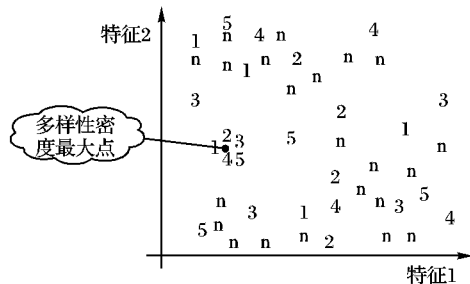


图4 多样性密度最大点在二维空间的表示

2 实验结果与讨论

2.1 多示例学习与全局方法处理果蔬图像的结果比较

目前,大多数农产品图像处理技术采用全局的方法,也即分割出图像前景之后,提取前景的全局颜色特征、纹理特征等。在超市果蔬识别上,Veggie Vision 系统和 Rocha 的研究中都采用了全局颜色直方图特征(Global Color Histogram, GCH)。因此,在实验中,采用 GCH 和支持向量机(Support Vector Machine, SVM)相结合的方法和本文多示例学习(Multiple-Instance Learning, MIL)方法进行了对比。由于全局方法不能处理混合果蔬,故未采用本研究采集的混合果蔬图像。供试图像为阿姆斯特丹目标识别图像库(Amsterdam Library of Object Images, ALOI)中的单类别农产品图像,包含 360°全方位的果蔬图像,如图 5 所示为一组来自 ALOI 的花椰菜图像。

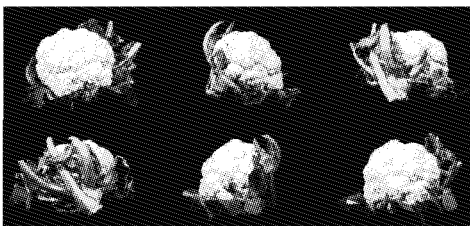


图5 ALOI 数据集中农产品图像

ALOI 图像库为每种果蔬提供了 111 幅图像。在实验中,对每一类果蔬模型的训练,随机选取 111 幅图像中的若干幅,例如 40 幅,作为正例样本,在其他种类的果蔬图像中选取 40 幅作为负例样本,剩余的 71 幅图像作为测试集,重复此操作

10 次,获得平均识别率。如图 6 所示的实验结果是采用全局方法和多示例方法,并在多示例方法中应用不同的尺度参数 H 的 5 种果蔬平均识别率。

分析统计结果可知:

1) 对于任意摆放方式的果蔬,由于其全局颜色分布特征可能产生较大变化,使得全局的方法难以达到更高的识别率,而本文方法能够通过调整尺度参数的调整,使示例细化的程度加大,对于包含多种躯干组成,或者摆放具有非刚性特性的果蔬,能够更有利于命中包中正例,使识别率优于全局方法。

2) 随着尺度参数 H 的增大,多示例学习的识别率有所提高。但是当 H 增大到一定程度(当 $H = 16$ 时),会造成识别准确率大幅下降,分析原因在于示例细化的同时增大了正例噪声,容易使多示例学习过程陷入不利局部极值。

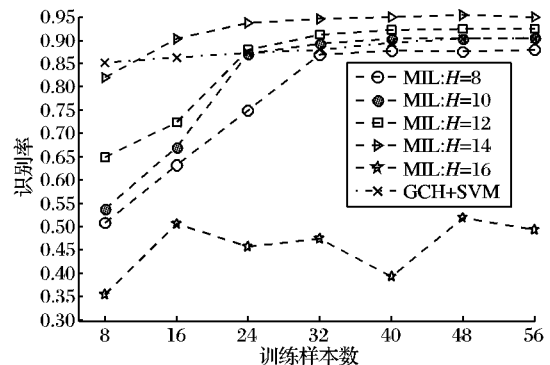


图6 单类别果蔬图像应用多示例学习和全局方法的识别率

2.2 多示例学习在混合果蔬图像上的识别

在混合果蔬的实验中,以采集的 5 189 幅图像作为样本集。在每种果蔬图像中选出 40 幅单类别图像和 60 幅多类别图像作为测试集,其余作为训练集,以 $H = 14$ 作为尺度参数,对每种果蔬分别训练识别模型,并在包含混合果蔬的测试集上进行识别。测试时,首先对每幅图像用本文的包生成方法进行处理,生成多示例包,再依次用训练得到的 15 个模型进行匹配,标识出识别获得的标签。与人工标注对比结果,各类别果蔬的识别率和错检率如图 7 所示。

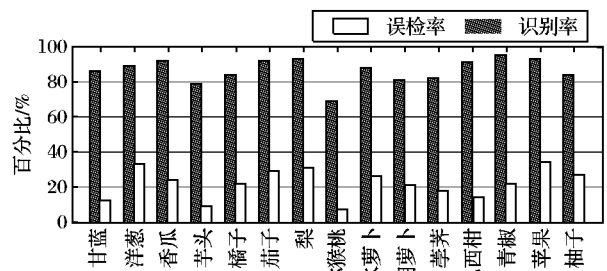


图7 各类别果蔬模型在混合图像上的匹配结果

由实验结果可以看出,大部分果蔬识别的识别率达到了 80% 以上,其中,青椒获得了最高识别率 94.21%。另外,对于一些颜色相近的果蔬,例如梨和香瓜,则会出现较高的误检率,在一些只有梨而不包含香瓜的图像中,误判图像与香瓜模型相匹配。分析其原因,尽管 SBN 方法包含了一定与纹理特性相关的颜色分布信息,但颜色均值信息在训练中获得的相关系数更大。

3 结语

本文研究了基于多示例学习的图像处理技术在超市果蔬

(下转第 1566 页)

同时对三种数码相机 Olympus FE360、Casio EX-Z2 和 Canon IXUS950 的识别率分别提高了 8.67%、4.67% 和 5.33%。

表3 $FAR = 10^{-3}$ 下各数码相机的判定阈值 t 、 FRR 和识别率

相机类型	Lukas 算法			$\sigma_1^2 = 8$ 和纹理聚类分析		
	t	FRR	识别率/%	t	FRR	识别率/%
Olympus	0.0098	2.1×10^{-3}	88.00	0.0115	3.6×10^{-4}	96.67
Casio	0.0067	8.6×10^{-2}	93.33	0.0072	2.4×10^{-2}	98.00
Canon	0.0076	4.4×10^{-2}	94.67	0.0096	5.0×10^{-3}	100.00

4 结语

本文针对传统的基于模式噪声图像来源鉴别方法的不足,提出了一种利用噪声方差和纹理复杂度分析的图像来源鉴别新方法:一方面,通过对插值像素点和非插值像素点的噪声方差进行分析,去除相机模式噪声中的插值像素点,以更好地反映相机的模式噪声特性;另一方面,通过对样本的纹理复杂度进行聚类分析,去除测试样本中的高纹理区域,减小背景噪声对相关系数计算造成的误差。实验表明本文算法具有更好的源相机辨识能力,与经典的基于模式噪声的相机源识别方法相比,在相同的 FAR 下具有更小的 FRR 和更高的识别率。一般情况下图像的纹理复杂度估计范围为 $[0, 60]$, 本文算法对该范围内的图像是适用的,但对于一些特殊的高纹理复杂度图像,由于进行纹理聚类分析之后,测试样本中参与相关性计算的像素点数变少,相关系数的误差将变大,算法有待进一步改进。

参考文献:

- [1] KHARRAZI M, SENCAR H T, MEMON N. Blind source camera identification[C]// Proceedings of IEEE International Conference on Image Processing. Piscataway: IEEE Press, 2004: 709–712.
- [2] SWAMINATHAN A, WU M, LIU K J R. Nonintrusive component forensics of visual sensors using output images[J]. IEEE Transactions on Information Forensics and Security, 2007, 2(1): 91–106.
- [3] CHOI K S, LAM E Y, WONG K K Y. Automatic source camera identification using the intrinsic lens radial distortion[J]. Optics Express, 2006, 14(24): 1551–1565.
- [4] JOHNSON M K, FARID H. Exposing digital forgeries through chromatic aberration[C]// Proceedings of the 8th Workshop on Multimedia and Security. Piscataway: IEEE Press, 2006: 48–55.
- [5] LUKAS J, FRIDRICH J, GOLJAN M. Digital camera identification from sensor pattern noise[J]. IEEE Transactions on Information Forensics and Security, 2006, 1(2): 205–214.
- [6] FRIDRICH J. Digital image forensics[J]. IEEE Signal Processing Magazine, 2009, 26(2): 26–37.
- [7] LI C T, LI Y. Digital camera identification using colour-decoupled photo response non-uniformity noise pattern[C]// Proceedings of IEEE International Symposium on Circuits and Systems. Piscataway: IEEE Press, 2010: 3052–3055.
- [8] YOICHI T, HITOSHI K. Digital camera identification based on the clustered pattern noise of image sensors[C]// Proceedings of IEEE International Conference on Multimedia and Expo. Piscataway: IEEE Press, 2011: 1–4.
- [9] LI C T. Source camera identification using enhanced sensor pattern noise[C]// Proceedings of IEEE International Conference on Image Processing. Piscataway: IEEE Press, 2009: 1509–1512.
- [10] 胡永健, 简超, 俞兵华. 利用彩色图像信息的源相机辨识策略[J]. 计算机应用, 2010, 30(10): 2694–2697.
- [11] JUN T, YASUYUKI M, TSUKASA O, et al. Estimating demosaicing algorithms using image noise variance[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2010: 279–286.
- [12] DIRIK A E, MEMON N. Image tamper detection based on demosaicing artifacts[C]// Proceedings of 16th IEEE International Conference on Image Processing. Piscataway: IEEE Press, 2009: 1497–1500.
- [13] KAZUYA M, HITOSHI K. An improved camera identification method based on the texture complexity and the image restoration[C]// Proceedings of the 2009 International Conference on Hybrid Information Technology. New York: ACM, 2010: 74–90.
- [6] DIETTERICH T G, LATHROP R H, PEREZ T L, et al. Solving the multiple instance problem with axis-parallel rectangles[J]. Artificial Intelligence, 1997, 89(1): 31–71.
- [7] MARON O, RATAN A L. Multiple-instance learning for natural scene classification[C]// Proceedings of the 15th International Conference on Machine Learning. San Francisco: Morgan Kaufmann Publishers, 1998: 341–349.
- [8] YANG C. Image database retrieval with multiple-instance learning techniques[C]// Proceedings of 16th International Conference on Data Engineering. Piscataway: IEEE Press, 2000: 81–82.
- [9] 戴宏斌, 张敏灵, 周志华, 等. 一种基于多示例学习的图像检索方法[J]. 模式识别与人工智能, 2006, 19(2): 179–185.
- [10] 王春燕, 袁津生. 一种结合多示例学习的图像检索方法[J]. 计算机系统应用, 2010, 19(6): 212–215.
- [11] ZHOU Z H, ZHANG M L, CHEN K J, et al. A novel bag generator for image database retrieval with multi-instance learning techniques[C]// Proceedings of the 15th IEEE International Conference on Tools with Artificial Intelligence. Washington, DC: IEEE Computer Society, 2003: 565–569.
- [12] MARON O. Learning from ambiguity[D]. Boston: Massachusetts Institute of Technology, 1998.

(上接第1562页)

图像上的应用。以自采集的方式构建了混合果蔬图像库,根据超市果蔬图像组成、分布、光照等因素的不确定性,对SBN包生成算法进行改进,并应用多样性密度算法训练果蔬模型,分别在单类别果蔬和多类别果蔬图像上进行了实验。结果表明,应用基于多示例学习的图像识别方法处理混合果蔬图像是可行的。

参考文献:

- [1] 展慧, 李小昱, 王为, 等. 基于机器视觉的板栗分级检测方法[J]. 农业工程学报, 2010, 26(4): 327–331.
- [2] 马东昱, 孙龙清. 基于图像特征的籽棉品级分级模型[J]. 计算机应用, 2010, 30(8): 2235–2238.
- [3] 张俊雄, 荀一, 李伟, 等. 山竹的计算机视觉分级方法[J]. 农业机械学报, 2009, 40(11): 176–179.
- [4] BOLLE R M, CONNELL J H, HAAS N, et al. Veggie vision: a produce recognition system[C]// 3rd IEEE Workshop on Applications of Computer Vision. Piscataway: IEEE Press, 1996: 244–251.
- [5] ROCHA A, HAUAGGE D C, WAINER J, et al. Automatic fruit and vegetable classification from images[J]. Computers and Electronics in Agriculture, 2010, 70(1): 96–104.