

透射图像颜色特征在烟叶识别中应用的探索

刘华波¹, 贺立源^{1*}, 马文杰^{1,2}, 李翠英¹

(1. 华中农业大学资源与环境学院, 武汉 430070; 2. 莱阳农学院传播学院, 青岛 266109)

摘要: 在理论推导的基础上, 提出并应用烟叶反射和透射图像获取烟叶内在的质量信息, 并在实际分级试验中验证透射图像颜色特征的有效性。研究发现, 透射图像的三个颜色特征 H、S、I 和相对应的反射图像特征相关性小, 可以有效补充反射图像不能反映烟叶内在质量的不足。利用 7 个烟叶正组各 5 片烟叶, 采用费歇尔线性逐步判别法筛选, 获得反射和透射图像的色调作为识别参数, 建立判别模型。研究结果表明, 采用透射图像作为反射图像的补充可以提高烟叶分组识别的准确率, 具有应用前景。

关键词: 透射图像; 烟叶识别; 计算机视觉

中图分类号: TP391.41; S572

文献标识码: B

文章编号: 1002-6819(2007)9-0169-03

刘华波, 贺立源, 马文杰, 等. 透射图像颜色特征在烟叶识别中应用的探索[J]. 农业工程学报, 2007, 23(9): 169-171.

Liu Huabo, He Liyuan, Ma Wenjie, et al. Color features of transmittance images applied in the classification of tobacco leaves[J]. Transactions of the CSAE, 2007, 23(9): 169-171. (in Chinese with English abstract)

0 引言

烟叶收购前要进行外观品质分级, 按质定级, 按级定价。分级人员凭眼看手摸, 依靠经验进行判断, 不但费时费力, 准确度和客观公正性也难以保证。将计算机视觉技术应用于烟叶智能分级的研究始于 20 世纪 80 年代, 并在近十年取得了较快发展。张建平^[1,2]等运用计算机图像处理技术和色度学理论, 研究了烟叶外观品质特征的提取、描述与定量检测方法, 并建立了可训练的识别模型, 初步证实了计算机视觉在烟叶分级领域中的可行性。张惠民^[3]等讨论了基于图像特征的烤烟烟叶分级系统的建模及其实现, 给出了烟叶分级模型的数学描述。Tattersfield^[4]等提取烟叶的形状、颜色等特征, 对烤烟的生长部位和颜色进行识别。蔡健荣^[5], 张帆^[6]等在提取各类图像特征的基础上, 分别将人工神经网络和模糊数学应用于烟叶的计算机视觉识别。现有研究普遍采用烟叶的反射图像来提取烟叶颜色、形状和纹理特征, 并利用各种算法和训练方法获得识别模型。这些研究在源头就只限于单一的反射图像, 对内在质量没有可靠的数据支持, 影响了相关研究进展。比较突出的问题体现在两个方面: 一是烤烟从田间采收至定级出售, 要经过烘烤、调制、贮藏等过程, 外界因素容易使烟叶出现破损、皱折等情况, 造成叶尖角、脉相和叶面纹理特征极不规律甚至无法提取; 二是国家标准中规定的叶片结构、身份、油分等质量分级指标, 在人工分级中是以触觉为主进行判断^[7], 而在反射图像特征中没有与之相关性好的对应参数, 难以模拟人工分级的实际情况。

烟叶具有半透明特性, 烟叶组织除了存在疏密差异和厚薄不同外, 烟叶组织中还含有选择性吸收不同波长的各种有机化合物, 如氨基酸、脂肪、蛋白质、有机酸等, 尽管在叶片组织中, 它们的特征吸收难以分离, 但是它们的复合吸收特征则可能反映烟叶品质。光线从背面照射烟叶时, 烟叶的厚薄、叶肉组织的疏密和叶片细胞内含物的差异会引起透射光强度和光谱特性的变化。相对入射光而言, 烟叶透过光形成的透射图像有可能反映烟叶的这些内在质量特征。基于此理论推测, 本研究尝试采用反射和透射二种形式获得烟叶图像, 探讨透射图像特征在烟叶计算机视觉分级中的有效性。

1 材料与方法

1.1 材料

试验烟叶为 2005 年山东临产 K326, 初烤后隔热隔光保存。全部由专家人工定组, 各组内不再分级。共选择正组内的七个组进行分组识别。分别为上部红棕(BR), 上部桔黄(BF), 上部柠檬黄(BL), 中部桔黄(CF), 中部柠檬黄(CL), 下部桔黄(XF), 下部柠檬黄(XL)。7 组烟叶每组各选择 5 片共 35 片烟叶进行建模训练, 另随机抽取 158 片烟叶进行模型的应用验证。

1.2 方法

本研究自行设计、制作了拍摄灯箱^[8], 灯箱为铝合金框架结构, 箱体大小为 100 cm × 75 cm × 120 cm, 可以容纳各种大小的烟叶。箱体顶部固定数码相机, 中间偏下放置厚度 4 mm 的双面磨砂玻璃作为载样台。灯箱上下部各并排固定 30 W 日光灯管(型号 T8D, 色温 6500K), 上部二支、下部 8 支。箱体的顶部、底部及四个侧面均用双层黑色棉布遮挡, 防止外界光线和灯箱内光的漫反射干扰。灯箱设置 3 位开关, 分别控制反射、透射和电源通断。采用数码相机进行图像采集, 并由存储卡转入计算机作进一步处理。图像像素大小为 640 × 480, 物距 75 cm, 拍摄前将舒展好的烟叶自然平铺在载样台上。对每一片烟叶, 分别开启顶光源或底光源, 拍摄反射和透射图像各一张, 使用 VC6.0 自编程序进行图像处理和像素统计。

基于实验目的和烟叶分级的特点, 采用 HSI(色调、饱和度和

收稿日期: 2006-11-30 修订日期: 2007-07-15

基金项目: 湖北省星火富民工程农村信息化促进专项“农村信息化体系建设与示范”资助

作者简介: 刘华波(1973-), 男, 湖北武汉人, 博士研究生, 主要进行农产品品质自动检测研究。武汉 华中农业大学资源环境学院, 430070。Email: liuhuabo@126.com

*通讯作者: 贺立源, 教授, 博士生导师, 主要从事农业信息工程研究。武汉 华中农业大学资源环境学院, 430070。Email: heliyuan@mail.hzau.edu.cn

强度)颜色模型提取烟叶特征。HSI 颜色模型将相关性很强的 RGB 三原色非线性地转化为光的波长特征(色调 H、饱和度 S)和强度特征(I),有利于判别率的提高和对结果的解释^[2,10]。

由于图像特征反映的是烟叶组织的综合性差异,质量指标与光学特征的关系复杂,实际在进行分组判别时不一定每个特征都有效,本研究采用费歇尔线性逐步判别法进行建模及验证,依次选择对最有效的特征参数进入判断模型,对判别作用不大的参数予以排除。先假定一组线性判别函数,然后根据已知信息对判别函数进行训练,使组间均值差异最大而组内离均差最小。确定了各函数式中的系数后,就可以根据函数式对待测数据进行判别。

2 结果与分析

供试 7 个组各 5 片烟叶的透射和反射图像的 HSI 特征参数的均值见表 1。

表 1 烟叶反射和透射图像的 HSI 颜色特征均值

Table 1 Mean characteristic values of each color feature of reflected and transmitted images of tobacco leaves based on HSI model

特征 参数		组别代码						
		BR	BF	BL	CF	CL	XF	XL
反射 图像	H	26.372	33.434	34.483	34.287	37.611	35.869	40.786
	S	0.399	0.458	0.460	0.447	0.434	0.473	0.467
	I	0.562	0.616	0.635	0.641	0.661	0.654	0.672
透射 图像	H'	42.856	44.61	46.113	40.462	41.849	43.745	43.198
	S'	0.725	0.704	0.676	0.755	0.756	0.676	0.656
	I'	0.312	0.253	0.366	0.285	0.292	0.362	0.377

2.1 数据分析

从特征参数的组间均值差异显著性分析(表 2)可以看出,除反射图像饱和度 S 外,其他 5 变量对应的组间均值差异都极显著,说明判别变量能够表明分类特征。

表 2 参数组间均值方差检验和相关系数

Table 2 ANOVA and correlation coefficients

		反射图像			透射图像		
		H	S	I	H'	S'	I'
反射 图像	H	1.000					
	S	0.294	1.000				
	I	0.843	-0.064	1.000			
透射 图像	H'	0.008	0.128	-0.049	1.000		
	S'	-0.245	-0.216	-0.175	-0.508	1.000	
	I'	0.302	0.224	0.228	0.455	-0.526	1.000
方差 检验	F	73.191	1.841	8.375	9.455	4.981	8.966

注:相关系数 $n = 35, r_{0.05} = 0.325, r_{0.01} = 0.418$;
方差检验 $F_{0.05}(6, 28) = 2.445, F_{0.01}(6, 28) = 3.528$ 。

比较各变量之间的相关系数可见,反射图像的 H 与 I 呈显著正相关,H 与 S、S 与 I 之间无显著相关;透射图像的 H、S、I 等 3 参数中任 2 参数之间均达到极显著相关,颜色特征的效应降低。从反射和透射三对颜色参数间的相关性看,相关系数最大值

为 0.302,发生在反射图像的色度和透射图像的亮度之间,其它参数两两间相关系数的绝对值都低于 0.3,均没有达到显著相关水平。此结果反映,反射特征和透射特征具有较好独立性,透射可以作为反射的有效补充应用于烟叶的质量评价。

2.2 建立判别模型

将七个组共 35 片烟叶的六个颜色特征变量和对应的专家分组数据输入 SPSS 软件中,使用逐步判别法^[11]进行分析。研究分别对反射图像 H、S、I、透射图像 H、S、I 和综合这 6 个参数进行判别,每次增加一个对判别影响最大的变量,引入判别方程,同时验证判别率。在判别正确率不再升高(或增加不显著)时,停止引入。对反射和透射图像特征的判别结果表明,两者分别都只有色调特征入选判别方程,而其他参数对本研究对象的判别贡献不大。在联合反射和透射特征参数的判别中,反射和透射的色调均能进入判别方程,其它参数也被淘汰。表 3 列出了判别分析的结果,可以看出二种图像的色调特征在分组识别中起主要作用,单独采用反射图像的色调比单独采用透射图像的色调识别率要高,但联合使用反射、透射图像的色调进行判别,可以显著提高识别率。根据贝叶斯概率公式,在假定各类别先验概率相等的条件下,得出表 4 所列线性分类函数。

表 3 不同变量参与判别结果对照

Table 3 Comparison of discrimination results by different variables sets

参与判别变量	入选判别变量	识别成功率/%
H, S, I	H	68.6
H', S', I'	H'	60.0
H, S, I, H', S', I'	H, H'	88.6

表 4 费歇尔线性判别函数系数表

Table 4 Coefficients of Fisher's linear discrimination functions

组别	H	H'	常量
BR	14.552	17.059	-559.364
BF	19.856	16.891	-710.636
BL	20.464	17.472	-757.611
CF	21.127	14.747	-662.490
CL	23.542	14.942	-757.321
XF	21.897	16.118	-747.184
XL	25.846	15.139	-856.026

2.3 识别模型验证

为了进一步检验本识别模型的效果,研究将未参与建模训练的一批样本进行了测试,总识别率达到 84%(表 5),表明使用线性判别模型,综合利用反射和透射图像的色调进行烟叶分组识别效果比较稳定,具有进一步研究的价值。

表 5 计算机识别与人工分级吻合率

Table 5 Tallying rates between computer vision recognition and manual grading

组别	BR	BF	BL	CF	CL	XF	XL	总计
样本数/片	12	16	25	20	25	30	30	158
吻合数/片	10	12	18	18	24	26	25	133
吻合率/%	83	75	72	90	96	87	83	84

3 讨论与结论

目前的烟叶分级,在国内外均采用人工感官判断的方法。其中一个非常重要但又难于严格把握的品质因素就是烟叶的颜色^[12],很多研究围绕计算机辨色展开,以提高不同分级人员判定结果的一致性。但是相关研究将辨色目标多集中在烟叶反射图像上,忽略了烟叶自身特性对不同波长辐射的吸收特征。本研究从理论角度论证了烟叶透射图像可能蕴涵的烟叶的结构和身份特征,在专门设计的灯箱环境中,分别获取烟叶的反射和透射图像,采用计算机视觉技术提取各自的色度,并综合应用反射和透射图像的色度进行判别分析,初步解决了通常只能由人工触觉判断的烟叶内在质量的特征表征问题,为扩展计算机视觉的应用领域提供了一种有益的思路。

研究结果证明,将透射图像用于烟叶的分组识别是可行的,方法简单,弥补了单独使用反射图像的局限性,如果配合比较容易识别的叶片长度和残伤就可能实现对7个正组25级烟叶的智能识别,较为系统的识别方法尚待进一步研究与验证。

[参 考 文 献]

- [1] 张建平,吴守一,方如明,等. 烟叶自动分级模型的建立与训练[J]. 农业工程学报, 1997: 179– 183.
- [2] Zhang J P, Sokhansanj S, Wu S, et al. A transformation technique from RGB signals to the munsell system for color analysis of tobacco leaves [J]. Computers and Electronics in Agriculture, 1998, 19: 155– 166.
- [3] 张惠民,韩力群,段正刚. 基于图像特征的烟叶分级[J]. 武汉大学学报·信息科学版, 2003, 28(3): 359– 362.
- [4] George Tattersfield, Keith Forbes. Classification of tobacco leaves by color and plant position[C]. Proceedings of the Ninth Annual Symposium of the Pattern Recognition Association of South Africa, 1998.
- [5] 蔡健荣,方如明,张世庆,等. 利用计算机视觉技术的烟叶质量分选系统研究[J]. 农业工程学报, 2000, 16(3): 118– 122.
- [6] 张帆,张新红,张彤. 模糊数学在烟叶分级中的应用[J]. 中国烟草学报, 2002, 8(3): 44– 48.
- [7] 王卫康. 烤烟》国标中分级因素的概念及把握[J]. 烟草科技, 2004, (5): 44– 48.
- [8] 马文杰,贺立源,刘华波,等. 成像环境因素对烟叶图像采集结果的影响及校正研究[J]. 中国农业科学, 2006, 39(12): 2615– 2620.
- [9] 周冀衡,朱小平,王彦亭,等. 烟草生理与生物化学[M]. 合肥: 中国科学技术大学出版社, 1996.
- [10] Chen Y B, Ling P P. Evaluation of HSI colorimetric system for intensity invariant spectral feature extraction [C]. St. Joseph, Mich: ASAE, 2001.
- [11] 任志娟. SPSS 中判别分析方法的正确使用[J]. 统计与决策, 2006, (2): 157.
- [12] 张杰,巫迎红,周慧玲. 烟叶的计算机辨色系统[J]. 中国烟草学报, 2002, 8(2): 38– 40.

Color features of transmittance images applied in the classification of tobacco leaves

Liu Huabo¹, He Liyuan¹, Ma Wenjie^{1,2}, Li Cuiyin¹

(1. College of Resources and Environment, Huazhong Agricultural University, Wuhan 430070, China;

2. Department of Communication, Laiyang Agricultural College, Qingdao 266109, China)

Abstract: A novel image analysis method was proposed in the tobacco leaf grading system based on transmittance theory. Both kinds of images were captured under reflected and transmitted light. The test showed that three pairs of color features from reflectance and transmittance image had very low correlations. Six color features were applied to develop recognition model by using Fisher's linear discrimination method. The result shows that the combination of both kinds of images obtained a higher classification rate than either single image type did. It is applicable to introduce transmittance image into the tobacco leaf grading system.

Key words: transmittance image; classification of flue-cured tobacco leaves; computer vision