

玉米单倍体籽粒特征提取及识别*

宋 鹏¹ 吴科斌¹ 张俊雄¹ 陈绍江² 刘 金² 李 伟¹

(1. 中国农业大学工学院, 北京 100083; 2. 中国农业大学国家玉米改良中心, 北京 100193)

【摘要】 研究了玉米单倍体育种过程经导入基因标记后杂交诱导产生的单倍体与杂合体籽粒特征提取及识别方法,实现了单倍体籽粒的准确识别。以导入基因标记后杂交诱导产生的玉米籽粒为研究对象,根据颜色特征将玉米分为紫色标记、黄色胚乳及白色胚部3部分,通过分析此3部分像素在RGB及HSV颜色空间内的分布特点,提出了 $(G-B)R/B$ 模型,实现了紫色标记区域的有效提取;根据玉米籽粒形态特点及紫色标记部分的分布特点,提出了一种简单快速的单倍体籽粒识别方法。在单倍体分拣试验台上的试验结果表明,所提出的方法对单倍体的识别准确率达98.07%,且能满足动态检测要求。

关键词: 玉米 单倍体 识别 图像处理 颜色特征 遗传标记

中图分类号: TP242.6; TP391.4 文献标识码: A 文章编号: 1000-1298(2012)03-0168-05

Corn Haploid Grain Feature Extraction and Recognition

Song Peng¹ Wu Kebin¹ Zhang Junxiong¹ Chen Shaojiang² Liu Jin² Li Wei¹

(1. College of Engineering, China Agricultural University, Beijing 100083, China

2. National Maize Improvement Center, China Agricultural University, Beijing 100193, China)

Abstract

A feature extraction and identification method of corn haploid and hybrid kernels was studied. The corn kernels were produced by hybrid induction after importing gene markers to their male parent in corn haploid breeding process. The proposed method made the accurate identification of corn haploid kernels come true. Corn kernels produced by hybrid induction was researched. Firstly, the corn kernels were divided into three parts according to the color distribution, including purple marks, yellow endosperm and white embryo. Then a $(G-B)R/B$ model was proposed by analyzing color features of pixel samples that got from the three parts. This model could extract the purple marks area effectively. A simple and quick identification method of corn haploid kernels was proposed after researching on morphological characteristics of corn kernels and distribution characteristics of purple marks. Experiments finished in the corn haploid sorting platform, they showed the accurate rate of corn haploid identification was above 98.07% by the proposed method, and it met the requirements of dynamic test.

Key words Corn, Haploid, Recognition, Image processing, Color feature, Genetic marker

引言

单倍体是指只具有配子染色体的细胞或个体。利用单倍体技术获得纯系进而选育自交系,可加快选育进程、提高选育效率,是现代植物育种中快速、

高效的方法之一^[1]。由于自然产生单倍体的概率极低(0.05%~0.1%),即使人工诱导也不到10%^[2],因此研究如何从诱导产生的大量籽粒中快速准确地识别出单倍体籽粒对玉米单倍体育种非常必要。

收稿日期: 2011-05-26 修回日期: 2011-06-07

* 国家高技术研究发展计划(863计划)资助项目(2010AA101401)、国家自然科学基金资助项目(31071320)和中国农业大学基本科研业务费研究生科研创新专项(KYCX2010088)

作者简介: 宋鹏,博士生,主要从事计算机视觉、农业机器人技术研究,E-mail: speng0408@gmail.com

通讯作者: 李伟,教授,博士生导师,主要从事机器视觉及农业机器人技术研究,E-mail: liww@cau.edu.cn

目前对玉米单倍体的鉴定主要有形态学鉴定、细胞学和解剖学鉴定、分子标记鉴定及遗传标记鉴定等方法^[3-5],其中采用遗传标记法是鉴定单倍体最有效可靠的方法,通过导入 Navajo 标记后,诱导产生的单倍体与非单倍体籽粒具有不同的颜色标记特征^[6],目前主要采用人工方式进行分拣。虽然近年国内外学者对玉米品种识别、品质检测等进行了大量研究^[7-10],但在单倍体检测方面鲜有报道。本文以经遗传标记后诱导产生的玉米籽粒为研究对象,提出一种简单快速的玉米单倍体籽粒图像分割模型及识别方法,为玉米单倍体自动分选系统的实现奠定基础。

1 材料与方法

1.1 样本选择与图像获取

试验样本由中国农业大学国家玉米改良中心提供。对单倍体诱导系杂交诱导产生的玉米穗,经干燥、脱粒后进行粗筛,去除其中的杂质及不规则籽粒后获得所需试验样本。选用母本为 LC0995,诱导系父本为 CAU,导入 Navajo 标记后,杂交诱导产生的玉米籽粒为研究对象,其中包括单倍体籽粒、杂合体籽粒及极少量外来籽粒(二倍体籽粒)。

如图 1 所示,导入基因标记后诱导产生的玉米籽粒分为 3 类:①杂合体籽粒,胚部和胚乳顶部均具有紫色标记,占大多数,约为 90%。②单倍体籽粒,胚部无紫色标记,但胚乳顶部带紫色标记,约小于 10%。③外来籽粒,胚部和胚乳顶部均无紫色标记,

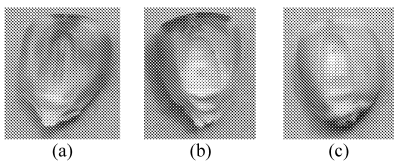


图 1 诱导产生的玉米籽粒

Fig. 1 Induced maize kernel

(a) 杂合体籽粒 (b) 单倍体籽粒 (c) 外来籽粒

含量极少,通常小于 1%。

采用文献[11]中的玉米单倍体分选系统获取玉米籽粒图像。该系统选用北京嘉恒中自图像技术有限公司生产的 OK_AC1300 型外触发面阵彩色摄像机和 Pentax 公司生产的焦距为 12.5 mm 的镜头采集图像,图像通过采集卡(OK_RGB10B,北京嘉恒中自公司)输入计算机。摄像机固定于系统传送带正上方 650 mm 处,玉米籽粒置于传送带上小孔内,小孔间距 40 mm,视场范围为 224 mm × 128 mm,所获图像分辨率为 896 × 512 像素。

1.2 样本图像分析

本文以从单倍体分拣试验台获取的动态图像为处理对象,待识别玉米籽粒置于分拣试验台传送带上的小孔内,如图 2a 所示。为加快图像处理速度,减少数据处理量,只对包含有玉米籽粒的区域进行处理,以小孔位置为参照,设定每个待处理区域的大小为 80 × 80 像素。经分析发现传送带、孔穴形成的背景与玉米籽粒在 R 通道差异较大,经图像平滑和 R 通道固定阈值处理后,可有效去除背景获取种子图像信息,背景去除后的图像如图 2b 所示。

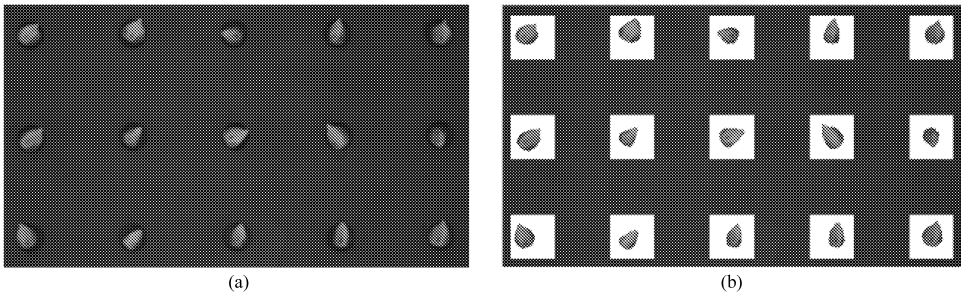


图 2 待处理种子图片

Fig. 2 To be processed seeds

(a) 原始图片 (b) 背景去除后图片

2 特征提取

诱导产生的 3 类玉米籽粒(单倍体、杂合体及外来籽粒)主要区别在于其颜色标记不同,因此可通过分析玉米籽粒的颜色特征,研究一种能有效提取紫色标记区域的颜色模型以实现单倍体与杂合体的准确识别。根据此 3 类玉米籽粒的不同颜色特征,将其分为 3 部分:紫色标记部分,该部分处于单

倍体和杂合体胚乳顶端及杂合体胚部;黄色胚乳部分,为杂合体与单倍体中无紫色标记的胚乳部分;无颜色标记的白色胚部,为单倍体和外来籽粒的胚部及杂合体胚部中无紫色标记的区域。因此,紫色标记区域是有效区分单倍体与杂合体的重要依据。

为实现紫色标记区域的有效提取,从试验图片中截取不同玉米籽粒中分别处于以上 3 部分的图像块进行分析。截取 15 颗不同种子上不同颜色区域

共计 1 800 个像素点,每个颜色区域 600 个像素点,形成颜色块。采用随机函数随机选取 600 个像素点中的 50 个点,分别分析其处于 RGB 及 HSV 颜色空

间中的颜色分布属性,以期获得能有效分割紫色标记区域的颜色模型。3 个不同颜色块在 RGB、HSV 颜色空间各通道的分布如图 3 所示。

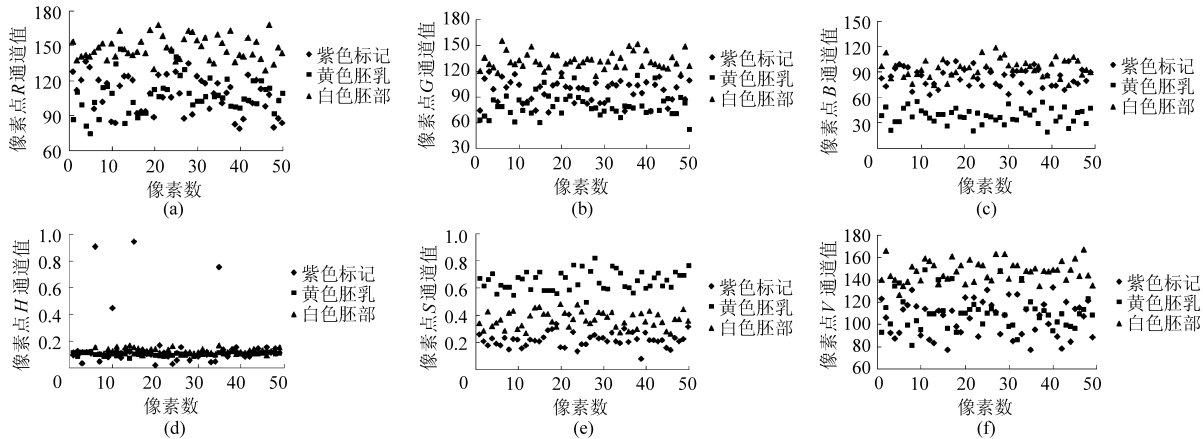


图 3 3 个颜色块在不同颜色通道分布图

Fig. 3 Three color blocks distribution in different color channels

(a) *R* 通道分布图 (b) *G* 通道分布图 (c) *B* 通道分布图 (d) *H* 通道分布图 (e) *S* 通道分布图 (f) *V* 通道分布图

由图 3 可以看出,在 RGB 空间内,*B* 通道可用于区分黄色胚乳部分,白色胚部区域与黄色胚乳部分 *R* 通道值与 *B* 通道值的差别较大,且 *R* 通道与 *G* 通道差值明显,而紫色标记区域的 *R*、*G*、*B* 通道值交错分布,相差较小;在 HSV 空间内,紫色标记部分和黄色胚乳部分在 *S* 通道的差别较大,而白色胚部和紫色标记部分在 *V* 通道差别较大。

根据以上分析,发现在 RGB 和 HSV 颜色空间内均可实现紫色标记区域的分割,在 RGB 颜色空间内玉米籽粒紫色标记部分的 *R/B* 值偏小,且黄色胚乳部分与胚部无标记的白色部分 *G* 通道值与 *B* 通道值相差较大,可采用 $(G - B)R/B$ 模型进行分割。而在 HSV 颜色空间内,由于其进行空间转换需要进行大量的非线性运算,虽然也能实现紫色标记区域的分割但大大降低算法效率,难以满足动态检测实时性需求。采用 $(G - B)R/B$ 模型进行目标区域分割,该模型对样本的统计结果如图 4 所示。

统计结果表明紫色区域的 $(G - B)R/B$ 值集中分布于 44 以下,为避免少量白色胚部区域像素误

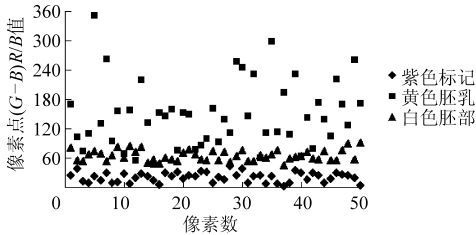


图 4 样本在分选模型下的分布图

Fig. 4 Sample distribution under the distinguishing model

判,采用 *R* 通道固定阈值进行限制。

$$T = \begin{cases} 0 & \left((G - B) \frac{R}{B} < 44 \text{ 且 } R < 135 \right) \\ 255 & \text{(其他)} \end{cases} \quad (1)$$

其处理效果如图 5 所示。

3 单倍体识别

诱导产生的 3 类玉米籽粒中,外来籽粒的比例小于 1%,其与单倍体籽粒的主要区别在于胚乳顶端是否具有紫色标记。由于动态分选过程的玉米籽粒置于传送带上小孔内,摄像机无法准确获取其胚

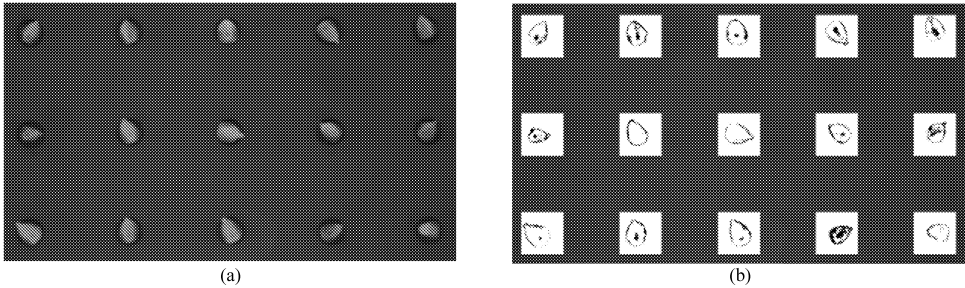


图 5 提取效果实例

Fig. 5 Result of processing

(a) 原始图 (b) 提取结果

乳顶端颜色信息,为提高算法效率及适用性,图像处理时将外来籽粒和单倍体籽粒归为一类,只区分单倍体籽粒及杂合体籽粒。

分选过程中,由于玉米籽粒置于传送带上小孔内,籽粒边缘在光源照射下的阴影特征与紫色标记部分相似,且杂合体和单倍体的胚乳顶端均具有紫色标记,使用上述颜色模型进行分割后,边缘及所有紫色标记区域均被有效提取,为了准确识别单倍体籽粒与杂合体籽粒,采用判断籽粒胚部是否具有紫色标记的方式进行识别,因此在提取出颜色标记后需对标记位置进行判断,若胚部区域具有颜色标记为杂合体籽,否则为单倍体籽粒。

由于玉米籽粒情况各异,如果采用软件算法提取胚部区域^[12]后再进行判断,不但算法复杂且会大大降低处理效率,难以满足动态检测的速度要求。在此,采用一种简单方式进行识别判断。随机取该

品种的 50 颗杂合体籽粒,人工测量通过籽粒尖端的最长轴,并统计胚部及胚部紫色标记区域与最长轴的位置关系。统计结果表明,将最长轴等分成 7 等份,从玉米籽粒尖端处开始,胚部区域包含最长轴约 6/7,且紫色标记区域处于胚部中心偏胚乳位置处。

如图 6a 所示,采用最大曲率法求出玉米籽粒的尖端点位置。在提取种子轮廓后,分别计算尖端点与轮廓上各点间距离,确定与尖端点最长距离点的位置,进而根据该点与尖端点确定籽粒最长轴。依据上述统计结果,将最长轴等分为 6 份,以距尖端分别为 3 倍和 4 倍等分距离的两点为焦点绘制椭圆,使得椭圆其中一端点处于最长轴上与尖端点约 6 倍等分距离处,处于椭圆内的紫色点被认为是胚部标记点,据此进行单倍体籽粒判别。动态提取效果如图 6b 所示,统计椭圆内部黑色像素数,取阈值为 10,作为识别单倍体与杂合体的判断依据。

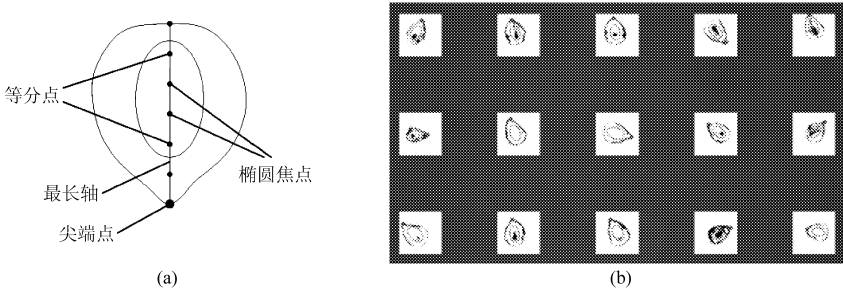


图 6 单倍体识别结果
Fig. 6 Figure of haploid identification
(a) 单倍体判断方法示意图 (b) 动态提取效果

4 试验结果分析

为验证该算法的检测精度和实时性,在由中国农业大学研制的玉米单倍体分选试验台^[11]上,采用由中国农业大学国家玉米改良中心提供的 LC0995 - CAU 玉米籽粒进行试验,从中随机选取杂合体和单倍体籽粒共 360 颗,其中杂合体籽粒 308 颗,单倍体籽粒 52 颗。试验时将所有玉米籽粒胚部朝上,放置于分选试验台传送带上的小孔内。根据本文所提出的分割模型和识别方法进行动态识别试验,将算法的识别结果与实际情况进行对比,结果如表 1 所示。

表 1 LC0995 - CAU 试验结果
Tab. 1 Result of test

实际所属 类别(数量)	判别结果		准确率/%
	杂合体	单倍体	
杂合体(308)	290	18	94. 20
单倍体(52)	1	51	98. 07

由于不同杂合体籽粒胚部颜色标记深浅各异,308 颗杂合体籽粒中有 18 颗被误判为单倍体籽粒,准确率为 94. 20%;而 52 颗单倍体中仅有 1 颗被误判为杂合体,准确率为 98. 07%,这主要由于胚部的褶皱在光照环境下形成阴影,其特征与紫色标记区域特征相似,从而造成误判。

5 结束语

分析了玉米籽粒的胚部、胚乳及紫色标记部分的颜色特征,提出了一种 RGB 组合算法。试验表明,该算法能实现玉米籽粒紫色标记区域的有效提取;通过提取玉米籽粒的尖端及最长轴,简单运算后,根据特定椭圆内的紫色标记区域像素点的数量进行单倍体识别。试验表明,所提出的方法对单倍体籽粒的准确识别率达 98. 07%,对杂合体的识别准确率达 94. 20%,且其算法简单满足实时分选要求,为单倍体自动分选系统的实现奠定良好的基础。

参 考 文 献

- 1 梁文科,戚廷香,徐尚忠. 单倍体技术在玉米育种中的应用和新进展[J]. 玉米科学,2004,12(3):13~15,18.
Liang Wenke, Qi Tingxiang, Xu Shangzhong. Progress and utilization of haploidy methods in maize breeding [J]. Journal of Maize Sciences, 2004, 12(3): 13~15,18. (in Chinese)
- 2 才卓,徐国良. 玉米单倍体育种研究进展[J]. 玉米科学,2008,16(1):1~5.
Cai Zhuo, Xu Guoliang. The advances in haploid breeding of maize [J]. Journal of Maize Science, 2008,16(1):1~5. (in Chinese)
- 3 韩赞平,王彬,张泽民. 玉米的单倍体育种[J]. 种子,2010,29(1):120~122.
Han Zanping, Wang Bin, Zhang Zemin. Haploid breeding of maize [J]. Seed, 2010,29(1):120~122. (in Chinese)
- 4 陈绍江,宋同明. 利用高油分的花粉直感效应鉴别玉米单倍体[J]. 作物学报,2003,29(4):587~590.
Chen Shaojiang, Song Tongming. Identification haploid with high oil xenia effect in maize [J]. Acta Agronomica Sinica, 2003, 29(4):587~590. (in Chinese)
- 5 汤飞宇,丁菲,王国英. 利用 SSR 标记检测来源于玉米孤雌生殖的双单倍体[J]. 江西农业大学学报,2004,26(6):859~862.
Tang Feiyu, Ding Fei, Wang Guoying. Identification of doubled haploid plants from maize gynogenesis by SSR molecular marker[J]. Acta Agriculturae Universitatis Jiangxiensis, 2004,26(6): 859~862. (in Chinese)
- 6 姜昱,王玉民,王中伟,等. 我国玉米单倍体育种技术研究进展[J]. 玉米科学,2008,16(6):48~51.
Jiang Yu, Wang Yumin, Wang Zhongwei, et al. The advances at haploid breeding of maize in China [J]. Journal of Maize Sciences, 2008,16(6): 48~51. (in Chinese)
- 7 张俊雄,荀一,李伟. 基于形态特征的玉米种子表面裂纹检测方法[J]. 光学精密工程,2007,15(6):951~956.
Zhang Junxiong, Xun Yi, Li Wei. Detection of surface cracks of corn kernel based on morphology [J]. Optics and Precision Engineering, 2007, 15(6): 951~956. (in Chinese)
- 8 程洪,史智兴,么炜,等. 基于支持向量机的玉米品种识别[J]. 农业机械学报,2009,40(3):180~183.
Cheng Hong, Shi Zhixing, Yao Wei, et al. Corn breed recognition based on support vector machine[J]. Transactions of the Chinese Society for Agricultural Machinery, 2009, 40(3): 180~183. (in Chinese)
- 9 王玉,刘贤喜,苏庆堂,等. 多对象特征提取和优化神经网络的玉米种子品种识别[J]. 农业工程学报,2010,26(6):199~204.
Wang Yu, Liu Xianxi, Su Qingtang, et al. Maize seeds varieties identification based on multi-object feature extraction and optimized neural network [J]. Transactions of the CSAE, 2010, 26(6): 199~204. (in Chinese)
- 10 权龙哲,祝荣欣,雷溥,等. 基于 K-L 变换与 LS-SVM 的玉米品种识别方法[J]. 农业机械学报,2010,41(4):168~172.
Quan Longzhe, Zhu Rongxin, Lei Pu, et al. Recognition method of maize cultivars based on K-L transform and LS-SVM [J]. Transactions of the Chinese Society for Agricultural Machinery, 2010, 41(4): 168~172. (in Chinese)
- 11 宋鹏,吴科斌,张俊雄,等. 基于计算机视觉的玉米单倍体自动分选系统[J]. 农业机械学报,2010,41(增刊):249~252.
Song Peng, Wu Kebin, Zhang Junxiong, et al. Sorting system of maize haploid kernels based on computer vision [J]. Transactions of the Chinese Society for Agricultural Machinery, 2010, 41(Supp.): 249~252. (in Chinese)
- 12 韩仲志,赵友刚,杨锦忠. 基于籽粒 RGB 图像独立分量的玉米胚部特征检测[J]. 农业工程学报,2010,26(3):222~226.
Han Zhongzhi, Zhao Yougang, Yang Jinzhong, Detection of embryo based on independent components for kernel RGB images in maize [J]. Transactions of the CSAE, 2010, 26(3): 222~226. (in Chinese)