

# 基于近红外高光谱成像技术的小麦不完善粒检测方法研究

刘欢<sup>1,2</sup>, 王雅倩<sup>1</sup>, 王晓明<sup>3</sup>, 安冬<sup>1\*</sup>, 位耀光<sup>1\*</sup>, 罗来鑫<sup>4</sup>, 陈星<sup>4</sup>, 严衍禄<sup>1</sup>

1. 中国农业大学信息与电气工程学院, 北京 100083
2. 山东科技大学电气与自动化工程学院, 山东 青岛 266590
3. 山东科技大学电气信息系, 山东 济南 250031
4. 中国农业大学植物保护学院, 北京 100094

**摘要** 小麦作为主要的粮食作物在我国农业生产、运输、食品加工等方面占有重要地位。不完善籽粒严重影响了小麦质量与粮食安全。不完善籽粒主要在生产、存储、包装等过程中产生, 目前我国小麦质量检测多以人工分选为主, 但存在人主观性较强, 肉眼易疲劳, 且费时费力等问题, 因此, 如何快速准确鉴别小麦不完善粒是现阶段提高生产率和保证粮食安全的重要问题。运用高光谱成像技术和特征波段选取方法提出一种快速有效的小麦不完善粒鉴别方法。利用近红外高光谱成像系统获得1 000粒小麦样本在862.9~1 704.2 nm共256个波段的高光谱反射图像, 其中包括健康粒、生芽粒、霉变粒和赤霉粒各250粒, 提取每个样本感兴趣区域的平均反射率光谱作为分类特征。本文首先对提取的全波段光谱信息进行窗口平滑、一阶导数差分、矢量归一化等数据预处理, 将原始光谱数据的隐藏信号放大并消除随机误差; 在预处理的基础上运用伪偏最小二乘(DPLS)和正交化线性判别分析(OLDA)对光谱进行特征提取, 降低数据的冗余度; 最后采用仿生模式识别(BPR)建立四类小麦的鉴别模型。实验结果表明, 采用全波段光谱信息建立的小麦不完善粒鉴别模型的平均识别精度达到97.8%, 分析结果可知, 利用近红外高光谱成像技术的全波段光谱信息对小麦不完善粒鉴别是可行的。尽管全波段光谱信息取得了较好的鉴别效果, 但高光谱成像设备较为昂贵, 获取高光谱全波段光谱信息数据量较大, 无法满足对现场设备运算速度的高要求, 因此, 采用连续投影算法(SPA)对全波段光谱数据进行特征波段的选择, 使波段数量由256维降低到10维, 从而提高系统的可行性和运算速度。采用选取的10个特征波段建立小麦不完善粒鉴别模型, 实验结果表明10个特征波段的平均识别精度仅为83.2%, 分析结果可知, 尽管采用10个特征波段提高了系统实时性, 但鉴别准确性较差。为达到与全波段特征基本相当的鉴别效果, 利用光谱特征与图像特征结合的方法建立小麦不完善粒鉴别模型, 将上述选取的10个特征波段的形态信息、纹理信息和光谱信息进行结合, 实验结果表明, 10个特征波段的光谱信息与图像信息结合使鉴别的平均识别精度达到94.2%, 此识别效果与利用全波段光谱数据的识别效果基本相当。利用高光谱成像系统探索了小麦不完善粒鉴别的可行性, 通过分析以上实验可知, 基于近红外高光谱成像技术对小麦不完善粒检测具有良好的效果, 在有效的提高运算速度的同时也保证了系统的鉴别精度, 为后期小麦不完善粒快速检测设备的开发提供了有效的研究方向。

**关键词** 小麦不完善粒; 高光谱成像技术; 连续投影算法; 光谱特征; 图像特征

**中图分类号:** S512.1 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2019)01-0223-07

## 引言

我国是世界上最大的小麦生产国之一, 小麦在我国粮食生产、流通、食品加工等方面占有重要地位。巨大的小麦产

量使小麦质量问题不断增加, 我国出台GB 1351—2008国家标准对小麦不完善粒给予明确规定, 小麦不完善粒由虫蚀粒、病斑粒(赤霉粒及黑胚粒)、破损粒、生芽粒、霉变粒组成, 这些不完善粒产生的质量问题不仅给粮食生产带来了巨大经济损失, 而且病变小麦的误食会危害人体健康, 甚至致

收稿日期: 2017-12-26, 修订日期: 2018-05-11

基金项目: 国家自然科学基金项目(31172260), 北京市科技创新基地培育与发展工程子专项(Z161100005016012)资助

作者简介: 刘欢, 1990年生, 山东科技大学电气与自动化工程学院硕士研究生 e-mail: Huan\_Liu\_33@yahoo.com

\* 通讯联系人 e-mail: andong@cau.edu.cn; wyg@cau.edu.cn

癌, 危及生命<sup>[1-3]</sup>。因此, 小麦质量鉴定显得尤为重要。

现阶段小麦不完善粒检测方法主要是人工检测, 此方法以其直观方便的优势被广泛应用, 然而, 人工肉眼挑选主观经验性太强, 肉眼易疲劳, 费时费力, 难以满足大批量的小麦质量检测, 由此基于机器视觉及近红外光谱分析的小麦不完善粒检测的方法被提出。机器视觉的方法通过采集小麦籽粒图像, 利用图像处理的方法得到小麦外观特征并对小麦的性质进行识别与统计, 然而由于机器视觉仅能采集样品外部特征而不能表达样品的内部差异, 因此对于外部差异较小的样本识别效果并不理想<sup>[4]</sup>。近红外光谱分析技术因能通过近红外波段得到样品内部有机分子含氢基团特征信息而被广泛的应用在食品质量检测和分级中<sup>[5-6]</sup>。然而近红外光谱分析技术易受环境和样品均匀性的影响, 对于小麦这种非均匀、小颗粒物质的应用并不理想, 因此近红外高光谱成像技术应运而生<sup>[5]</sup>。

高光谱成像技术集图像信息与光谱信息于一身, 在获得样品外部特征信息的同时, 也可以提供反映样品内部物理结构和化学成分的光谱信息, 从而实现图谱合一的特性<sup>[6]</sup>, 因此在形态相近和物质含量相差较大的分类问题上广泛应用。高光谱成像技术在食品、农产品质量与安全检测方面具有独特的优势。其中 Serranti 等<sup>[1]</sup>使用高光谱成像技术有效地检测了多种小麦损伤粒; Shahin 等<sup>[7]</sup>利用高光谱成像技术检测加拿大小麦样品中的赤霉粒, 其准确率达到了 90%; Singh 等<sup>[8]</sup>运用高光谱的 700~1 100 nm 波段信息对健康与真菌损伤小麦粒分类, 能正确识别 99%~100% 的真菌损伤小麦粒。上述研究表明, 高光谱成像技术已成功应用于单一指标或者多个指标的小麦品质检测, 根据实地(五得利面粉集团有限公司)调查可知小麦不完善粒中赤霉粒, 发芽粒、霉变粒出现的情况较多, 而目前对于赤霉粒, 发芽粒、霉变粒的多指标检测却鲜有报道, 因此本文结合近红外高光谱成像技术对赤霉粒, 发芽粒、霉变粒三类不完善粒进行分类检测。上述研究并没有考虑在实际生产中的可行性, 缺少应用价值, 本文提出的研究方法为小麦的快速质量检测提供了一定的研究方向。

使用高光谱成像技术对小麦不完善粒进行快速准确鉴别, 具体实现方案如下。

(1) 采用全波段的光谱特征信息探索健康粒、赤霉粒、发芽粒及霉变粒四类小麦籽粒的可区分性;

(2) 利用 SPA 算法选择光谱特征波段对小麦不完善粒进行检测, 实现小麦不完善粒鉴别的实时性;

(3) 在快速鉴别的基础上, 结合特征波段的光谱信息、纹理信息及形态信息对小麦不完善粒进行检测, 提高小麦不完善粒鉴别的准确性。

## 1 实验部分

### 1.1 图像采集与校正

研究采用图 1 所示的 GaiaSorter 高光谱成像系统(北京卓立汉光仪器有限公司)对样本进行采集, 高光谱设备主要由四个部分组成, 包括光谱相机、光源、样品置物台和计算

机。其中的光源是由 2 套 50W 钨灯光源组成, 两组光源被安置在平台两侧, 光源高度为 25 cm, 安装角度为 30°。光谱相机包括芬兰 SPECIM 公司的 Inspector 系列成像光谱仪和 InGaAs CCD 相机, 成像波段为 862.9~1 704.2 nm, 其光谱分辨率为 5 nm, 相机高度为 10 cm, 相机与计算机通过网口连接。相机的缝隙宽度为 30 μm。设备带有样本放置移动平台, 按照推扫式的方式对样本图像进行采集, 平台由步进电机进行控制, 电机与计算机通过串口进行通信。样品放在黑色背景板上, 选择纹理不明显、不反光和不粘灰尘的材质制作背景板, 减少图像采集时背景带来的误差。实验采用图像采集软件 Spectra View 控制高光谱成像系统采集小麦籽粒成像信息。在保证图像不失真的情况下设置平台的移动速度为 0.27 cm·s<sup>-1</sup>, 曝光时间为 35 ms。采集到的每组样本都包含有从 862.9~1 704.2 nm 近红外波段的 256 张图像。

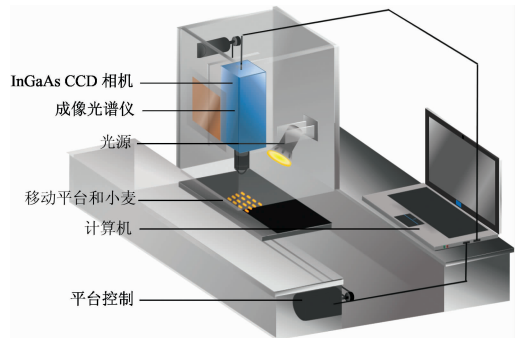


图 1 高光谱图像采集系统

Fig. 1 Hyperspectral imaging system

为减小研究中光源波动和相机存在暗电流等因素带来的测量误差, 对采集的高光谱图像进行黑白校正。黑白校正之后的数据可以更好的表现出某一物质在某一波段下的性质特征。我们将采集的原始图像信息( $I_{\text{sample}}$ )、黑色背景信息( $I_{\text{dark}}$ )、白色背景信息( $I_{\text{white}}$ )根据式(1)进行黑白校正,  $R$  为黑白校正后所得图像。

$$R = \frac{I_{\text{sample}} - I_{\text{dark}}}{I_{\text{white}} - I_{\text{dark}}} \quad (1)$$

其中黑背景数据采集时用设备镜头盖将相机封闭。白板背景采集将白板放置在与样本平齐的位置进行采集, 并且保证与样品采集时状态相同。

### 1.2 样本

本研究所用小麦样本由五得利面粉集团有限公司提供, 小麦品种选为现阶段北京市主推品种农大 211, 由经验丰富的专业检测人员通过人工筛选得到健康、赤霉病、发芽和霉变四类小麦籽粒, 小麦样本放在 5 °C 的干燥环境中储存, 在四类小麦样本中分别取出 250 粒作为本研究的实验样本。本研究共获得 25 幅图像, 为减小仪器漂移带来的影响, 每幅图像同时采集健康小麦、赤霉病小麦、发芽小麦、霉变小麦各 10 粒。

### 1.3 光谱与图像数据特征提取

通过高光谱成像系统, 可以采集到小麦的三维图像, 由于实验样本量较大, 采用 ENVI 手动标定感兴趣区域时间花费较多, 因此本文采用感兴趣区域自适应阈值分割和掩膜结

合的方法对种子感兴趣区域反射率光谱进行提取<sup>[9]</sup>。光谱提取步骤如图 2 所示，样本处理均在 MATLAB r2016a 上进行。

图 2 中，(a) 高光谱图像的获取(每幅图像包含 40 粒种子，每类各占 10 粒)；(b) 对高光谱图像进行图像增强、滤

波等预处理；(c) 动态选择阈值对样本二值化处理并标定感兴趣区域(ROIs)边界；(d) 通过边界坐标生成掩模，对 256 个波段的图像进行掩模，得到每粒种子 256 个波段的图像；(e) 去除背景信息，得到每粒种子 256 个波段的平均灰度值，从而得到光谱反射率特征曲线。

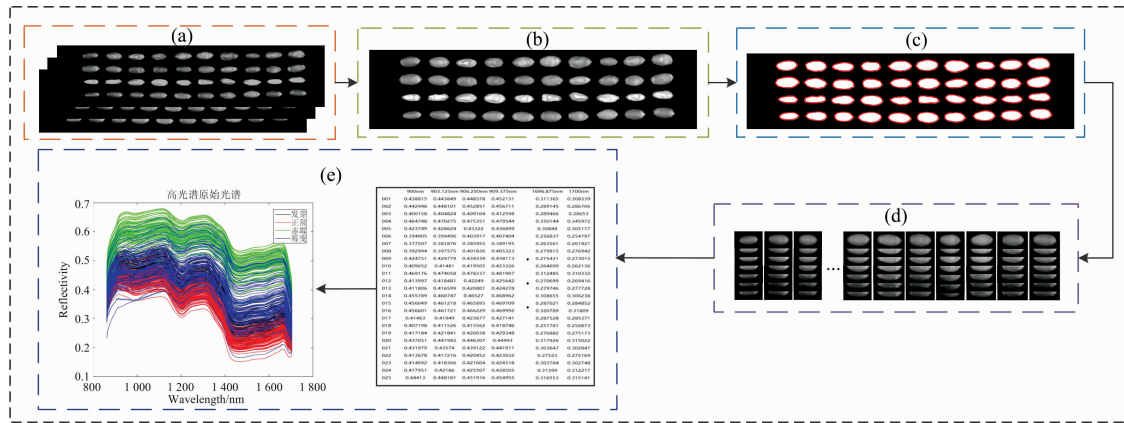


图 2 图像分割与光谱提取步骤

Fig. 2 The steps of image segmentation and spectral extraction

高光谱成像技术具有图谱合一的特点，其不仅可以获得样本的光谱信息，同时也可以得到样本的图像信息。本研究从纹理和形态两个方面进行图像特征的提取。研究采用的纹理统计量包括灰度分布统计量(均值、标准差、平滑度、一致性、熵、归一化的三阶矩)和灰度共生矩阵统计量(对比度、相关性、能量、局部均匀性，灰度共生矩阵熵)共 11 个统计量。对样本形态特征提取前需对图像进行二值化处理，将样本从背景中分离出来，对样本 10 个形态特征(面积、周长、长、宽、长宽比、矩形度、圆形成、等面积圆直径，紧致度，延展度)进行提取。

1.4 光谱特征波段选择

高光谱全波段的光谱维相邻波段间存在信息冗余，这些冗余信息会影响鉴别模型的精确性和稳定性，且高光谱的全波段的光谱信息处理速度较慢，不能满足实际不完善粒检测的要求。研究采用连续投影算法<sup>[10]</sup>(SPA)选择特征波段点，SPA 是一种使矢量空间共线性最小化的前向变量选择算法，通过提取全波段中的特征波段点消除原始光谱中冗余的信息，加快了样本处理速度。

1.5 数据处理与模型建立

光谱预处理是建立一个准确稳定模型的基础。光谱预处理使用平均窗口平滑(宽度为 9)，一阶导数差分(宽度为 9)和矢量归一化对光谱进行预处理，消除随机产生的误差并提高光谱的分辨率。经过预处理后依然无法充分提取样本特征信息，因此，在预处理的基础上进一步采用伪偏最小二乘(DPLS)和正交化线性判别分析(OLDA)对光谱进行特征提取和降维，使降维后的数据在不同维度之间相互正交。最后结合仿生模式识别<sup>[11]</sup>(BPR)建立四类小麦的仿生识别模型，其中采用二权值神经元作为基本覆盖单元，利用最小生成树连接基本单元。

2 结果与讨论

2.1 原始光谱分析

从高光谱图像中提取感兴趣区域的光谱信息，光谱范围是 862.9~1 704.2 nm 共 256 个波段点，由于前后段噪声较大，因此取 928.4~1 603.1 nm 共 205 个波段进行分析，四种类型的小麦籽粒在 928.4~1 603.1 nm 的平均原始光谱曲线如图 3 所示。从光谱整体特征上可以看出，四类小麦籽粒的呈现相似的光谱特性，其中赤霉粒由于水分的缺失呈现发白干瘪的特征，近红外光对水的反应较为敏感使赤霉粒反射率明显高于健康、发芽、霉变三种小麦籽粒。而健康粒、发芽粒和霉变粒三种类型的小麦籽粒差异较小。霉变粒因真菌侵蚀使籽粒表面物质发生改变，因此与健康籽粒有一定差

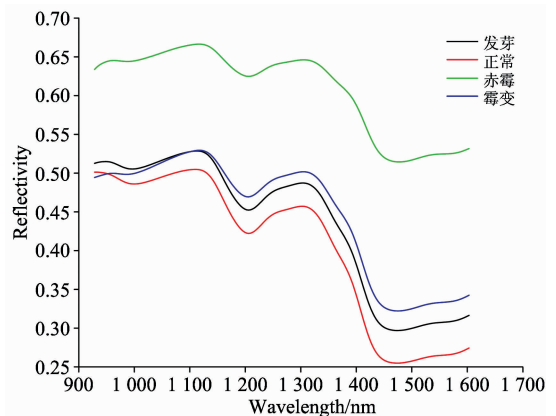


图 3 小麦籽粒的平均光谱反射率曲线

Fig. 3 The average spectral reflectance curve of wheat kernels

异, 霉变粒其霉变侵蚀程度不尽相同, 相比较赤霉粒, 霉变粒光谱曲线与健康籽粒光谱曲线距离更近。由于发芽粒对种子能量的消耗使得发芽粒与健康粒子有所区别, 但发芽粒物质成分变化相对较小, 所以相比较其他类型籽粒, 发芽粒与健康粒差异较小。

## 2.2 基于光谱全波段的小麦不完善粒鉴别

为了探究全波段光谱数据对健康粒与三种不完善粒小麦间的可区分性, 将样本按照 1:1 的比例分为训练集和测试集, 分别对训练集和测试集进行预处理, 剔除训练集中存在的异常样本, 并利用交叉验证的方法来确定 DPLS 的最优主元数, 再用 OLDA 将数据降至 4 维, 最后结合 BPR 建立四类小麦的仿生识别模型。测试结果统计见表 1。

表 1 基于全波段光谱特征的模型识别结果  
Table 1 Identification results of model based on spectral characteristics of full wavebands

样本数	训练集平均识别率/%	测试集识别率/%				平均识别率
		健康	赤霉	发芽	霉变	
1 000	100	99.2	100	96	96	97.8

从图 4 可以看出健康粒、赤霉粒、发芽粒、霉变粒四类样本具有较为明显的可区分性, 由此可见, 采用近红外高光谱的全波段光谱信息可以实现小麦不完善粒的定性判别。从表 1 可得训练集的识别精度在 100%, 测试集的平均识别精度达到 97.8%, 其中健康粒识别精度是 99.2%, 赤霉粒识别精度达到 100%, 发芽粒与霉变粒识别精度均为 96%。以上研究结果表明, 采用近红外高光谱全波段光谱信息可以实现对小麦不完善粒的有效鉴别。

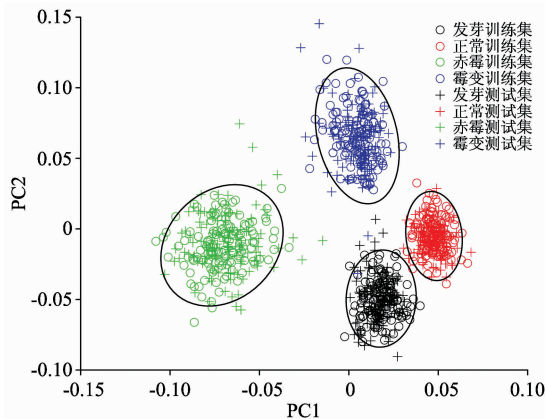


图 4 训练集和测试集样本的二维分布图

Fig. 4 Two-dimension distribution map of samples in training set and test set

## 2.3 基于特征波段的小麦不完善粒鉴别

由于全波段的高光谱数据在实际应用中存在鉴别实时性差, 高光谱仪器设备价格昂贵等问题, 使得采用全波段的光谱数据难以满足实际生产需要, 因此本研究采用 SPA 算法从未经过预处理的 205 个波段的原始样本集中提取特征波

段。由于样品的分类精度与最优波段点的数量有关<sup>[12]</sup>。因此分析 5~25 个不同数量的波段点对正确识别率的影响(表 2), 从表 2 可以看出随着特征波段的增加, 训练集和测试集的平均分类精度也在不断上升, 特征波段点在 10 个左右, 测试集的平均识别精度只能达到 83.2%, 当选取 25 个波段时, 测试集的平均识别率达到 96.2%, 基本接近于全波段光谱维测试集的平均识别率。然而 25 个特征波段并不利于便携式仪器的开发, 因此本文在选择 10 个光谱特征波段点的基础上, 增加 10 个特征波段点的图像维信息, 所选 10 个特征波段点如图 5 所示, 分别是 928.4, 945.6, 983.2, 1 095.2, 1 202.5, 1 311.9, 1 370.9, 1 406.8, 1 452.3 和 1 603.1 nm。其中在 900~1 000 nm 为 O—H 的三倍频, 与碳水化合物吸收有关, 在 1 100~1 200 nm 左右为 C—H 的三倍频, 与淀粉为主的碳水化合物吸收有关, 在 1 400 nm 左右为 O—H 的二倍频, 与酚类化合物的吸收有关<sup>[13]</sup>。

表 2 不同特征波段数的平均正确识别率  
Table 2 The verage correct recognition rate of different band numbers

波段数	特征波段	
	训练集平均识别率/%	测试集平均识别率/%
5	89.9	64.4
10	91.5	83.2
15	95.16	88.60
20	96.8	94
25	98.4	96.2

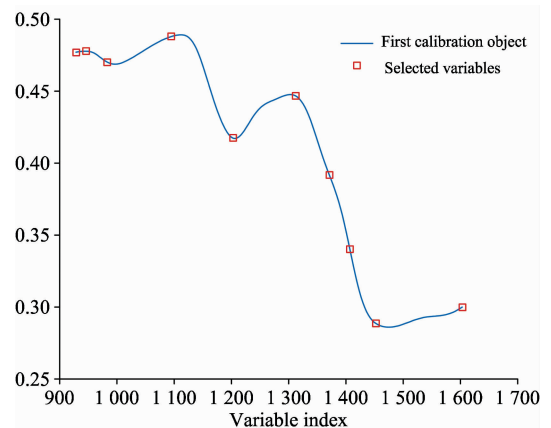


图 5 基于 SPA 算法的特征波段点选择

Fig. 5 Characteristic band points selected by SPA algorithm

## 2.4 基于图像特征和光谱特征的小麦不完善粒鉴别

从以上研究可以看出, 只使用小麦 10 个特征波段的光谱信息, 虽然在一定程度上提高了鉴别速度, 但分类精度无法达到小麦不完善粒检测的要求, 因此在提取特征波段光谱信息的基础上进一步提取小麦样本的图像特征。小麦的图像特征从小麦的纹理特征和形态特征两方面进行提取。

研究采用 10 个形态特征来表征小麦的形态, 分别计算健康粒、赤霉粒、发芽粒、霉变粒每类特征的类内标准差

std1 和类间标准差 std2, 以其标准差比值(std2/std1)来评价四类小麦在该特征上的差异大小, 其中标准差比值越大, 表明在这个特征下四类的差异越大, 10 个特征的标准差比值如表 3 所示。

表 3 统计四种类型小麦籽粒的形态特征标准差比值

Table 3 Statistics of standard deviation ratio of morphological characteristics of four wheat varieties

特征	标准差比值	特征	标准差比值
面积	1.154	矩形度	1.181
周长	1.063	圆形度	1.317
长	1.009	等面积圆直径	1.160
宽	1.263	紧致度	1.234
长宽比	1.296	延展度	1.035

从表 3 中可以看出, 四类种子的圆形度和长宽比差异较大, 标准差比值分别为 1.317 和 1.296。选择四类样本中前 20 粒种子, 分别得出四类小麦圆形度(图 6)和长宽比(图 7)的对比分布图, 从图中可以看出, 赤霉粒的圆形度要小于其他三种类型的小麦, 长宽比则高于其他三种类型的小麦, 而对比健康、发芽和霉变三种类型的小麦在圆形度和长宽比方面并无明显差异, 三种类型的圆形度和长宽比之间均存在较大的交叉。通过观察小麦样本并结合研究结果来看, 可以发现, 赤霉籽粒较为干瘪, 饱满度明显低于其他三类, 而其他三类外形特征却十分相似。通过对形态特征进行分类(表 4), 可以看出由于健康、发芽、霉变三类小麦形态差异不大, 使得采用形态特征分类的测试集的平均识别率仅为 33.6%, 四种类型的小麦在形态特征上没有表现出明显的可分性。

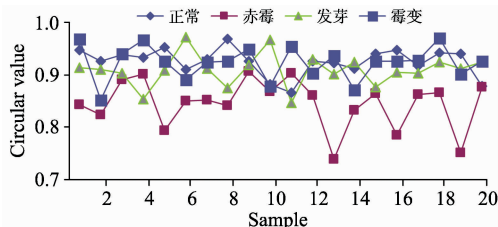


图 6 圆形度对比分布图

Fig. 6 Comparative distribution of circularity

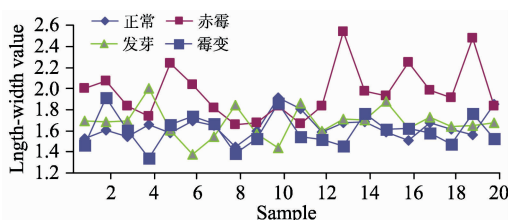


图 7 长度比对比分布图

Fig. 7 Comparative distribution of length-width ratio

近红外高光谱图像中对不同物质吸光程度不同导致图像信息不同, 化学变化不完善籽粒与物理损伤的不同之处在于其成分含量的不同而非形态差别。图像的纹理特征可以得

到图像的本质信息而非种子的形态信息, 从侧面反映了物质的变化。

由于只使用形态特征无法可靠的获得有效的分类信息, 因此研究采用 10 个特征波段的纹理特征, 建立鉴别模型, 研究统计结果如表 4 所示, 可以得出选取的 10 个特征波段的纹理信息, 训练集和测试集识别精度均在 90% 以上, 实验结果表明纹理信息可以获取小麦不完善籽粒鉴别的有效信息。

表 4 三种方式的识别结果

Table 4 Identification results of three methods

建模特征	建模集识别率/%	测试集识别率/%				
		健康	赤霉	发芽	霉变	平均
光谱特征	91.5	83.2	84	96.8	68.8	83.2
纹理特征	95.9	89.6	92	94.4	84.8	90.2
形态特征	66.9	41.6	30.4	35.2	27.2	33.6

在分类模型中, 样本包含的信息越丰富, 分类模型精度越高, 因此在光谱特征波段的基础上增加了图像信息(纹理信息和形态信息)。综合多种特征信息后, 对四种小麦不完善粒进行检测, 测试结果如表 5 所示。

表 5 光谱信息与图像信息结合的识别结果

Table 5 Identification results of spectral information combined with image information

建模特征	建模集识别率/%	测试集识别率/%				
		健康	赤霉	发芽	霉变	平均
光谱+纹理	99.2	97.6	93.6	96.0	85.6	93.2
光谱+纹理+形态	98.8	99.2	91.2	99.2	87.2	94.2

从表 5 中可以看出, 光谱信息与图像的纹理和形态信息相结合时平均识别率由 83.2% 升到 94.2%, 基本可以达到全波段光谱信息的鉴别效果, 因此采用 10 个特征波段的光谱信息与图像信息相结合的方法可以实现小麦不完善粒的快速准确检测。

### 3 结 论

利用 862.9~1 704.2 nm 的近红外高光谱成像技术分别从光谱信息和图像信息对小麦健康粒和不完善粒进行识别检测, 通过研究结果可得以下结论:

通过建立的全波段光谱模型, 对健康粒和不完善粒小麦的总体鉴别精度达到 97.8%, 在一定程度上说明采用近红外高光谱成像技术可以实现对小麦不完善粒的有效检测; 为满足实际应用, 增强实时性, 采用 SPA 算法进行波段选取, 通过研究结果可知, 选择 10 个特征波段建立模型的测试集平均识别率仅在 83.2% 左右, 因此在提取特征波段光谱信息的基础上, 提取特征波段的图像信息。

为提高特征波段的鉴别精度, 分别采用纹理信息和形态信息建立鉴别模型, 研究表明, 由于健康粒、发芽粒、霉变粒形态相似, 无法通过形态信息对这三类小麦进行有效

鉴别, 而选择 10 个波段的纹理信息进行鉴别, 测试集识别率达到 90.2%, 识别精度稍高于仅采用特征波段的光谱信息。

由于分别采用特征波段的光谱信息和图像信息的鉴别精度均低于全波段光谱数据, 因此将光谱信息与图像信息、形态信息相结合进行小麦不完善籽粒的鉴别, 通过研究结果可

知, 将光谱信息与纹理信息结合所得总体分类精度最高, 达到 94.2%, 其识别精度基本可达到全波段光谱信息的鉴别精度, 与此同时特征波段的选取使模型鉴别速度显著高于全波段光谱鉴定方法。

致谢: 感谢五得利面粉集团有限公司提供小麦样本。

## References

- [ 1 ] Serranti S, Cesare D, Bonifazi G. *Biosystems Engineering*, 2013, 115(1): 20.
- [ 2 ] Delwiche S R, Kim M S. *Sensing and Instrumentation for Food Quality and Safety*, 2011, 5: 63.
- [ 3 ] Senthilkumar T, Jayas D S, White N D G. *Journal of Stored Products Research*, 2016, 65: 30.
- [ 4 ] Rahman Anisur, Cho Byoung-Kwan. *Seed Science Research*, 2016, 26: 285.
- [ 5 ] Dong Gao, Guo Jian, Wang Cheng. *International Journal of Agricultural and Biological Engineering*, 2017, 10(2): 251.
- [ 6 ] ZHAO Jia-wen, CHEN Quan-sheng, LIN Hao(赵杰文, 陈全胜, 林 颢). *Introduction Modern Imaging Technology and Its Application in the Detection of Food and Agricultural Products(现代成像技术及其在食品、农产品检测中的应用)*. Beijing: China Machine Press(北京: 机械工业出版社), 2011.
- [ 7 ] Shahin M A, Symons S J. *Journal of Food Measurement and Characterization*, 2012, 6: 3.
- [ 8 ] Singh C B, Jayas D S, Paliwal J, et al. *International Journal of Food Properties*, 2012, 15: 11.
- [ 9 ] Huang Min, Tang Jinya, Yang Bao. *Computers and Electronics in Agriculture*, 2016, 122: 139.
- [ 10 ] Tian Minglu, Ban Songtao, Chang Qingrui. *Transactions of the Chinese Society of Agricultural Engineering*, 2016, 32(21): 102.
- [ 11 ] WANG Shou-jue(王守觉). *Acta Electronica Sinica(电子学报)*, 2002, 30(10): 1417.
- [ 12 ] Huang Min, He Chujie, Zhu Qibing. *Applied Sciences*, 2016, 6: 183.
- [ 13 ] YAN Yan-lu, ZHAO Long-lian, HAN Dong-hai, et al(严衍禄, 赵龙莲, 韩东海, 等). *Foundation and Application of Near-Infrared Spectroscopy Analysis(近红外光谱分析基础与应用)*. Beijing: China Light Industry Press(北京: 中国轻工业出版社), 2005.

## Study on Detection Method of Wheat Unsound Kernel Based on Near-Infrared Hyperspectral Imaging Technology

LIU Huan<sup>1, 2</sup>, WANG Ya-qian<sup>1</sup>, WANG Xiao-ming<sup>3</sup>, AN Dong<sup>1\*</sup>, WEI Yao-guang<sup>1\*</sup>, LUO Lai-xin<sup>4</sup>, CHEN Xing<sup>4</sup>, YAN Yan-lu<sup>1</sup>

1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China
2. College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao 266590, China
3. Department of Electrical Engineering and Information Technology, Shandong University of Science and Technology, Ji'nan 250031, China
4. College of Plant Protection, China Agricultural University, Beijing 100094, China

**Abstract** Wheat is a major food crop and occupies an important position in Chinese agricultural production, transportation, and food processing. Unsound kernel seriously affects wheat quality and food security. Wheat unsound kernel is mainly produced during production, storage, and packaging. At present, the manual sorting method is the main method for detecting wheat kernel quality in China. It is subjective, time-consuming, laborious, and costly. Therefore, the rapid and accurate identification method of the wheat unsound kernel will increase productivity and ensure food security. So the method for rapid and accurate detection of wheat unsound kernel was proposed by using hyperspectral image technology and the method of characteristic band selection. In this paper, near-infrared hyperspectral imaging system was used to collect hyperspectral reflection image of 1 000 wheat kernels (including healthy kernels, sprouted kernels, mildewed kernels, and kernels infected with *Fusarium head blight*, their respective amount are 250) at 862.9~1 704.2 nm (a total of 256 bands) and the average reflectivity of each sample were extracted from region of interests of hyperspectral images as classification characteristics. This paper conducts pre-processing for the extracted full-wave bands spectral information through window smoothing, first order derivative and vector normalization. It will also amplify hidden signals of the original spectral data and erase random errors. On the basis of pre-processing, feature extraction by applying discriminant partial least squares (DPLS) and orthogonal linear discriminant analysis (OLDA) to lower the redundancies of the data. Finally, it establishes identification model for 4 kinds of wheat through pattern recognition (BPR).

The experiment results showed that the average identification accuracy of the model for wheat unsound kernel established by using full-wave bands spectral information is 97.8%. The analysis also proves that it is feasible to detect wheat unsound kernel by using near-infrared hyperspectral imaging technology. Though full-wave bands spectral information achieved better detection effect, the high costs of hyperspectral imaging equipment and large amount of hyperspectral full-wave bands spectral information data fail to meet the high requirement of calculation for site equipment. Therefore, this paper uses successive projections algorithm (SPA) to select characteristic bands among full-wave bands data and lower the number of band from 256 dimensions to 10 dimensions to improve the operation and calculation speed of the system. So 10 characteristic bands were taken to establish identification model for wheat unsound kernel. The experiment results showed that the average identification accuracy of the 10 characteristic bands is only 83.2%, which means that though the 10 characteristic bands improve the real-time capability of the system, but they show worse identification accuracy. In order to achieve the identification effect that is basically equivalent to the characteristics of the whole band, this paper uses the combination of spectral features and image features to establish identification model of the wheat unsound kernel. All kinds of relevant information (morphological information, texture information, spectral information) of the kernel of the above 10 selected wave bands are integrated. The experimental results showed that the combination of spectral information and image information in 10 characteristic bands can achieve an average identification accuracy of 94.2%. Its identification effect is basically consistent with the use of full-wave bands spectral data. This paper uses hyperspectral imaging system to explore the feasibility of wheat unsound kernel detection. From the analysis of the above experiment, it can be seen that near-infrared hyperspectral imaging technology shows better results in the detection of wheat unsound kernel. It can guarantee the identification accuracy of the system while improving calculation speed so it offers an effective research orientation for later development of equipment that is able to detect wheat unsound kernel in a quick way.

**Keywords** Wheat unsound kernel; Hyperspectral imaging technology; Successive projections algorithm; Spectral features; Image features

(Received Dec. 26, 2017; accepted May 11, 2018)

\* Corresponding authors