

# 基于近红外高光谱图像分析的麦粒硬度分类研究

张红涛,田媛,孙志勇,母建茹,阮朋举,侯栋宸  
(华北水利水电大学 电力学院,河南 郑州 450011)

**摘要:**为对小麦硬度进行自动检测,采集不同硬度小麦品种的近红外高光谱图像,将光谱数据经过求导处理后,提取 950~1 645 nm 有效光谱区间数据,然后经过多元散射校正,建立偏最小二乘判别分析(PLS-DA)模型。采用 120 粒小麦对模型进行训练,剩余的 90 粒进行检验,总体上模型分类准确率为 99.63%。表明,采用近红外高光谱成像技术对单籽粒小麦硬度进行分类是可行的。  
**关键词:**近红外高光谱图像;光谱分析;偏最小二乘判别分析;小麦硬度;分类  
**中图分类号:**TP391.41;S512.1 **文献标志码:**A **文章编号:**1004-3268(2015)04-0181-04

## Hardness Classification of Wheat Kernel Based on Near-infrared Hyperspectral Imaging Technology

ZHANG Hongtao, TIAN Yuan, SUN Zhiyong, MU Jianru, RUAN Pengju, HOU Dongchen  
(Institute of Electric Power, North China University of Water Resources and Electric Power, Zhengzhou 450011, China)

**Abstract:** As the hardness of wheat may greatly influence the milling process, it is necessary to activate automatic detection of wheat hardness. To prepare for the research, the near-infrared hyperspectral images of wheat with different hardness were collected. The data were processed by derivation, and those in spectral range between 950—1 645 nm effective were extracted, after multiplicative scatter correction, with which a partial least squares discriminant analysis model(PLS-DA) was built. During the experiment, 120 wheat kernels were used to train the model, and the remaining 90 kernels were used to predict. Consequently, the accuracy rate of the model was 99.63%. The results showed that it was feasible to classify the hardness of wheat kernel based on near-infrared hyperspectral imaging technology.  
**Key words:** near-infrared hyperspectral image; spectral analysis; partial least squares discriminant analysis; wheat hardness; classification

小麦是我国重要的粮食作物,其种植面积仅次于水稻<sup>[1]</sup>。小麦籽粒硬度通过影响出粉率、破损淀粉粒数量、润麦加水量以及面粉颗粒度大小,最终决定小麦磨粉及其他加工品质,是衡量小麦品质的一项重要指标,对出入库质量控制具有非常重要的指导意义<sup>[2]</sup>。

小麦籽粒硬度的各种测定方法大多是基于压碎、研磨、磨耗或压痕的原理,主要包括角质率法、研磨时间法、损伤淀粉法、颗粒度指数法、单籽粒谷物

特性测定仪法和近红外法等<sup>[3-5]</sup>。随着我国粮食检测技术的逐渐提高,对小麦硬度检测提出了更高的要求。高光谱成像技术具有“图谱合一”的特性,不但包含了被检测物体空间结构的二维信息,还可得到任一感知点所对应像素的第三维信息,即光谱信息,目前已被广泛地应用于粮粒虫害、农产品外部品质、肉类鲜嫩度等检测中<sup>[6-8]</sup>。在粮粒硬度分类方面,Gerida 于 2009 年运用高光谱成像技术对小麦硬度分类进行了初步研究,根据粮粒中的每个像素对

收稿日期:2014-12-02  
基金项目:国家自然科学基金项目(31101085);河南省基础与前沿技术研究计划项目(122300410145);河南省高等学校青年骨干教师资助计划项目(2011GGJS-094);华北水利水电大学教学名师培育项目(2014108);华北水利水电大学 2014 年大学生创新创业计划项目(HSCX2004143)  
作者简介:张红涛(1977-),男,河南邓州人,副教授,博士,主要从事农业信息化研究。E-mail:zht1977@ncwu.edu.cn

杜伦麦、硬麦和软麦的硬度进行分类判别,识别率在 93% 以上<sup>[9]</sup>。

本研究采用近红外高光谱成像系统获取不同硬度小麦样本的高光谱图像,对获取到的高光谱图像计算平均光谱,经过多元散射校正(multiple scattering correction, MSC)预处理后,用偏最小二乘判别分析(partial least squares discriminant analysis, PLS-DA)法<sup>[10-12]</sup>对麦粒硬度进行分类判别,为单籽粒小麦硬度的检测提供一种快速无损的检测方法,这对小麦的分类储藏、精细加工等具有非常重要的意义。

## 1 材料和方法

### 1.1 试验仪器

采用近红外高光谱成像系统采集小麦籽粒的样本图像,成像系统主要由光谱成像单元、照明单元、位移单元、计算机和光箱五部分组成。光谱成像单元又称光谱相机,主要包括近红外相机、成像光谱仪、F/2.0 镜头;照明单元包括直流可调光源和玻璃光纤线性灯;位移单元由控制器和位移台组成<sup>[13]</sup>。近红外相机的光谱范围为 900 ~ 1 700 nm,成像光谱仪分辨率为 5 nm,狭缝宽 30  $\mu\text{m}$ 。光箱密封性能较好,以避免在采集过程中受到外界光线的影响。

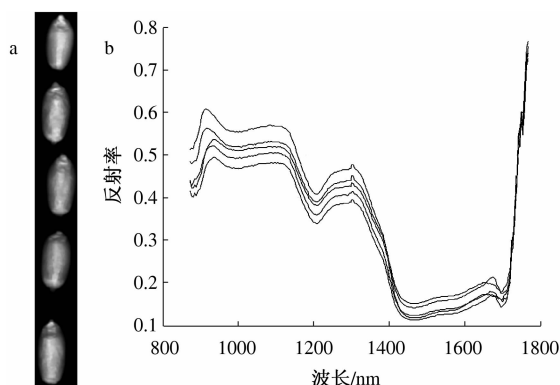
### 1.2 方法

1.2.1 样品处理 试验用小麦样本品种西农 979、豫麦 58 和众麦 1 号,均购于河南新乡平原新区农商新天地。经过小麦硬度指数测定仪测量,它们的硬度指数(HI)分别为 71、56、45,分别属于硬麦、混合麦和软麦。3 种硬度的小麦样品各取 70 粒,置于 25  $^{\circ}\text{C}$  干燥的环境中,使其温度均保持 25  $^{\circ}\text{C}$ ,水分在 9% ~ 15%,以供高光谱图像的采集。

1.2.2 高光谱图像的采集 在高光谱图像数据采集前,首先确定近红外相机的曝光时间以保证拍摄到的图像清晰,确定位移台的速度大小以避免图像尺寸和空间分辨率失真。图像采集时,采用曝光时间 3 000  $\mu\text{s}$ ,位移台的移动速度为 0.59 mm/s。采集前需对所获得的高光谱图像进行黑白标定,以去除部分在各波长下由于光源强度的分布不均造成的图像噪声。

近红外高光谱采集软件为 Spectral Cube,实际采集到的光谱范围是 871.6 ~ 1 766.3 nm,采样间隔为 3.5 nm,垂直于位移台运动方向的空间分辨率为 62.5  $\mu\text{m}$ 。将每 5 个小麦样品分为 1 组,把每组样品分别置于有黏性的黑色塑料板上,使样品处于 1 条直线上且互不接触。图像采集时,随着位移台的移动,实际近红外相机采集到的图像中麦粒是从下向

上排列的。图 1a 所示为 1 组小麦样品的高光谱灰度图像。



a. 高光谱图像; b. 平均光谱曲线

图 1 采集的 1 组小麦的高光谱图像及平均光谱曲线

1.2.3 高光谱图像的分析 对获取的麦粒高光谱图像中的每个波长图像进行自动滤波处理,去除图像中的噪声,分割出仅含单个麦粒的高光谱图像。提取出麦粒内每个像素的光谱曲线,将麦粒内所有像素的光谱曲线进行平均,作为该麦粒的平均光谱曲线。图 1b 所示为 1 组小麦样品中每个麦粒的平均光谱曲线。对于平均光谱曲线,在建立模式分类模型前,需消除信号的背景,确定有效波长区间的范围,并克服测量光的波长及样本颗粒的不均匀性和样本的折射指数等使光谱散射而导致的差异,需要对原始数据光谱进行一阶求导、二阶求导和多元散射校正处理。

## 2 结果与分析

### 2.1 数据的预处理

对获取近红外高光谱图像的每个籽粒分别求出平均光谱,即为未经处理的原始数据,则 210 粒小麦的平均光谱曲线如图 2a 所示,3 种硬度小麦籽粒的平均光谱曲线如图 2b 所示。对原始数据光谱的一阶求导、二阶求导结果如图 2c、d 所示。由图 2b 可见,在波长 1 657 nm 之后 3 种硬度小麦的光谱曲线出现重合。由图 2c、d 可见,当波长小于 950 nm 和大于 1 645 nm 时,小麦籽粒的光谱波动较大。故在建立 PLS-DA 模型的时候,应去除小于 950 nm 和大于 1 645 nm 波段的数据,选取 950 ~ 1 645 nm 的波段进行处理分析。由近红外高光谱成像系统采集的 256 个信道波长数据可知,应采用第 20 ~ 224 个共 205 个波长下的小麦籽粒数据来建立 PLS-DA 模型。经过 MSC 校正后的第 20 ~ 224 个信道波长下的光谱数据如图 3 所示,可见,MSC 处理较好地消除了测量小麦籽粒时因颗粒不均、散射和各种噪声及物理扰动等使光谱产生的差异。

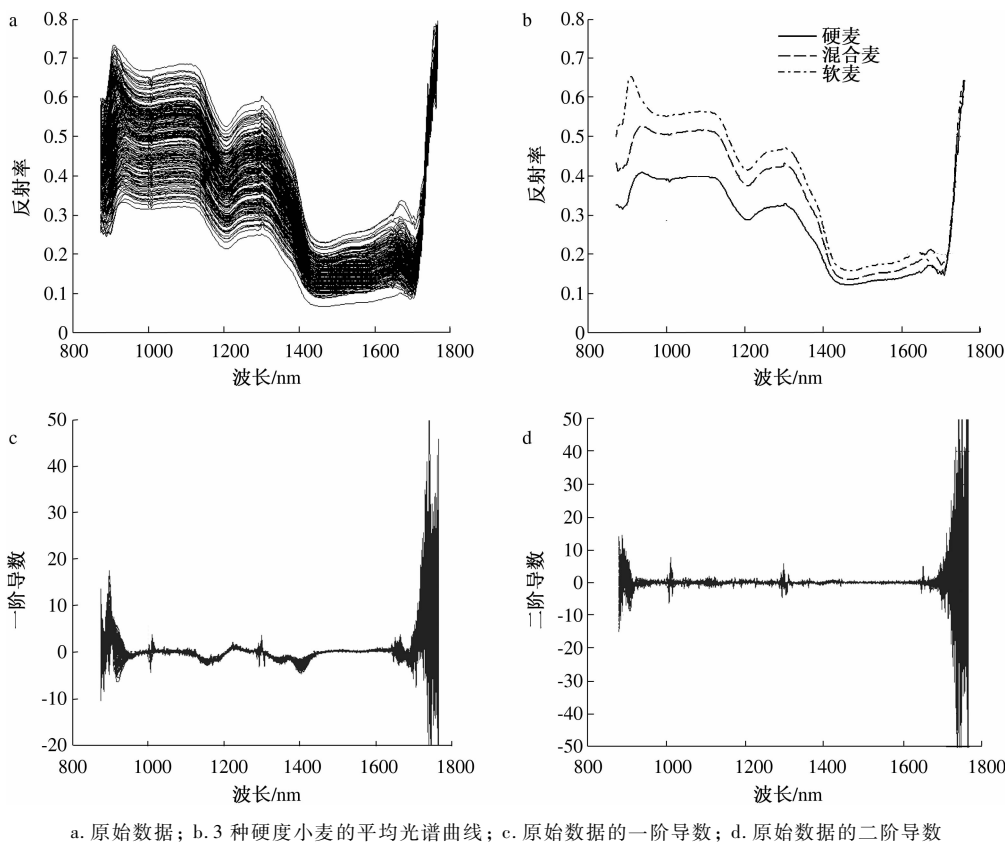


图 2 小麦籽粒光谱曲线及其预处理结果

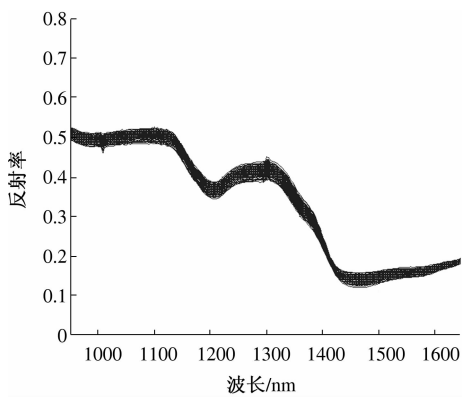


图 3 选取 950 ~ 1 645 nm 波段经 MSC 后的麦粒光谱

2.2 PLS - DA 模型的建立

在经过 MSC 处理的 210 个小麦籽粒数据中,每种硬度随机抽取 40 个麦粒,共 120 个麦粒作为训练集,剩余的 90 个麦粒作为检验集,以建立 PLS - DA 模型。该模型的输入为 120 个小麦籽粒在 950 ~ 1 645 nm 的平均光谱,输出分类变量  $Y$  中,硬麦、混合麦、软麦分别为(1 0 0)、(0 1 0)、(0 0 1),通过训练后的模型实际上是代表 3 种麦粒硬度分类的回归表达式,由于建模使用的是第 20 ~ 224 个波长下的光谱数据,故每个回归方程有 206 个系数。

为了优化模型,需剔除模型中的奇异点,确定最

佳成分个数。在确定的光谱预处理方式和选择波段范围内,采用交互验证预测误差平方和(PRESS)来确定最佳主成分数,PRESS 值越小对应主成分个数越好,说明模型的预测能力越强。PLS - DA 模型中交互验证 PRESS 值随主成分数的变化如图 4 所示。同时在验证中,交叉的有效性计算如式(1)所示。

$$Q^2_h = 1 - \text{PRESS}(h)/\text{SS}(h - 1) \tag{1}$$

式中,PRESS( $h$ )代表交叉验证时抽取  $h$  个成分回归后的预测误差平方和;SS( $h$ )代表此时采用所有样本点抽取  $h$  个成分回归后的预测误差平方和。

在建模的每一步计算结束之前,均进行交叉有效性验证。如果第  $h$  步  $Q^2_h < 1 - 0.95^2 = 0.0975$ ,则

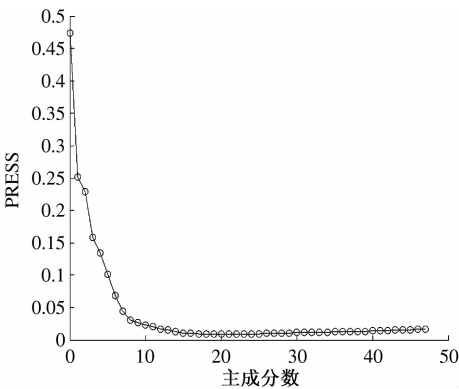
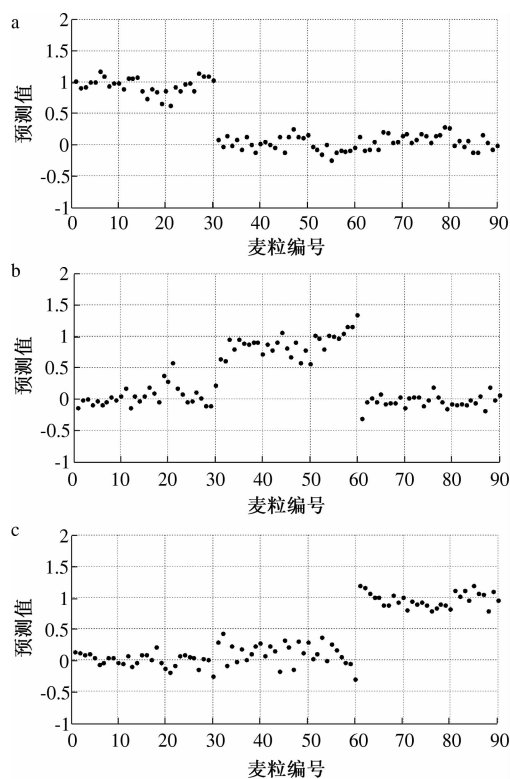


图 4 交互验证 PRESS 随主成分数的变化曲线

模型达到精度要求,可停止提取主成分;若  $Q_h^2 \geq 0.0975$ ,表示第  $h$  步提取的成分边际贡献显著,应该继续第  $h+1$  步的计算。当  $h=8$  时,  $Q_h^2 = 0.0757 < 0.0975$ ,停止交叉有效性验证,此时  $PRESS = 0.0452$ ,提取的主成分数为 8。

### 2.3 PLS-DA 模型的检验

利用经过 MSC 处理后数据建立的 PLS-DA 模型对检验集的小麦籽粒数据进行分析,结果如图 5 所示。对于第  $i$  个检测样本  $X_i$ ,模型预测输出值为  $Y_i$ ,若  $Y_{ik} \in [0.5, 1.5]$  ( $i=1, 2, \dots, 90; k=1, 2, 3$ ),则判定样本属于第  $k$  类;若  $Y_{ik} \in [-0.5, 0.5]$ ,则判定样本不属于第  $k$  类。由图 5 可见,经过 MSC 处理后的数据,对硬麦、混合麦、软麦特征的预测正确率分别为 100%、98.89%、100%,总体预测正确率为 99.63%,达到了较好的效果。



a. 采用硬麦模型; b. 采用混合麦模型; c. 采用软麦模型。90 个麦粒中,第 1~30、31~60、61~90 个分别为硬麦、混合麦、软麦

图 5 PLS-DA 模型预测结果

## 3 结论与讨论

采集了不同硬度小麦籽粒的近红外高光谱图像,经过多元散射校正处理后,建立 PLS-DA 模型对不同硬度的小麦进行分类,预测正确率为 99.63%。结果表明,采用近红外高光谱图像技术能有效地对小麦籽粒硬度进行分类。本研究应用 205 个

波长的数据对小麦硬度进行分析,数据量比较大,下一步需提取出更能表征麦粒硬度的特征波长区间进行 PLS-DA 建模,以提高模型判别的精度和运行的速度。由于我国小麦的品种比较多,小麦的硬度范围变化比较大,以后需进一步对单籽粒小麦硬度值的自动测定进行研究,以更好地满足粮食加工行业的需求。

### 参考文献:

- [1] 李雪,穆月英,付文革.近十年我国小麦生产影响因素的实证分析——侧重于 20 个小麦生产省[J].农业经济,2009,29(10):32-34.
- [2] 田素梅.两种小麦籽粒硬度测定方法比较试验[J].粮油仓储科技通讯,2012,28(4):38-39.
- [3] Moghaddam T M, Seyed M A R, Taghizadeh M. Applications of hyperspectral imaging in grains and nuts quality and safety assessment: A review[J]. Journal of Food Measurement and Characterization, 2013, 7(3):129-140.
- [4] Milica M P, Jasna S M. Near infrared spectroscopy-advanced analytical tool in wheat breeding, trade, and processing[J]. Food and Bioprocess Technology, 2013, 6(2):330-352.
- [5] 吴存荣,孙辉,周展明. GB 1351—2008《小麦》标准解读[J].粮食加工,2008,33(3):14-17,44.
- [6] 张红涛,任孟梅,胡玉霞,等.基于近红外超光谱成像的谷蠹活虫检测研究[J].河南农业科学,2014,43(4):152-155.
- [7] 李江波,饶秀勤,应义斌.农产品外部品质无损检测中高光谱成像技术的应用研究进展[J].光谱学与光谱分析,2011,31(8):2021-2026.
- [8] 陈全胜,张燕华,万新民,等.基于高光谱成像技术的猪肉嫩度检测研究[J].光学学报,2010,30(9):2602-2607.
- [9] Gerida D T. Near infrared hyperspectral imaging and chemometrics for exploration and classification of whole wheat kernels[D]. South Africa: Stellenbosch University, 2009.
- [10] 罗香,刘波平,张小林,等.偏最小二乘近红外光谱法测定瘦肉脂肪酸组成的研究[J].分析试验室,2007,26(10):25-29.
- [11] 张敏,杨进,付海燕,等.近红外光谱结合 PLS-DA 鉴别不同产地的川续断[J].三峡大学学报:自然科学版,2011,33(4):81-84.
- [12] 杨忠,任海清,江泽慧. PLS-DA 法判别分析木材生物腐朽的研究[J].光谱学与光谱分析,2008,28(4):793-796.
- [13] 胡玉霞,张红涛,谢俊明,等.粮虫近红外高光谱图像的采集系统设计与实现[J].河南农业科学,2012,41(6):112-114.