

基于光谱及成像技术的种子品质无损速测研究进展

王冬^{1,3}, 王坤², 吴静珠^{2*}, 韩平^{1,3*}

1. 北京农业质量标准与检测技术研究中心, 北京 100097
2. 北京工商大学食品安全大数据技术北京市重点实验室, 北京 100048
3. 农业农村部农产品质量安全风险评估实验室(北京), 北京 100097

摘要 种子是农业生产过程的重要生产资料。种子质量评价、活力与老化检测、纯度与真伪鉴别、分类与溯源研究是种子品质检测中的常见问题。种子质量主要包含种子含水率、蛋白含量、脂肪酸含量、淀粉含量等,是种子品质分级的重要指标,并且关系到种子存储过程的安全问题。种子活力是种子发芽和出苗率、幼苗生长的潜势、植株抗逆能力和生产潜力的总和;高活力种子具有明显的生长优势和生产潜力。种子老化是指种子活力的自然衰退,表现为种子变色、发芽率低、生长势差、作物减产。种子的纯度与真伪则会影响作物产量和农产品品质;而种子分类与溯源则是保证种子纯度与鉴别种子真伪的重要方法,进而为作物产量与产品品质提供保障。对于种子品质分析,传统方法通常需要对样品做不可逆的破坏性分析,且分析时间长、过程复杂,难以适应现代农业对种子生产环节的需要。因此,开展种子品质无损快速检测技术研究成为当前亟待解决的问题。近年来,随着化学计量学的发展和计算机技术的进步,近红外光谱法以其快速、无损、高效等优势,在农产品、食品、农业投入品等的无损快速分析方面得以广泛的应用。进一步地,将光谱技术与成像技术相结合,高光谱成像技术近年来日益兴起,相比较于传统的光谱技术,高光谱成像技术在获得待测样品的光谱信息的同时,还可以获取样品的空间分布信息以及图像特征。基于近红外光谱及高光谱成像等无损快速检测技术,从种子质量评价、活力与老化检测、纯度与真伪鉴别、分类与溯源研究四方面对近年来关于种子品质无损快速检测文献进行综述。在分析不同检测技术特点的基础上,分别就上述种子品质检测方面的问题加以整理。进而对种子品质无损快速检测的技术特点进行了总结与展望。

关键词 种子;近红外光谱;高光谱成像;无损检测

中图分类号: O657.33 **文献标识码:** R **DOI:** 10.3964/j.issn.1000-0593(2021)01-0052-08

引言

种子作为农业投入品,是农业生产过程的重要生产资料,其品质对农业经济有着至关重要的影响。种子质量是种子品质分级的重要指标,关系到种子存储过程的安全问题,并且可对种子等级划分以及优劣判定进行定量描述。种子活力直接影响种子的发芽速度、出苗整齐度以及抵抗逆境环境生长的能力,与植株的生长优势和生产潜力密切相关;而老化种子则是指种子活力的自然衰退,表现为种子变色、发芽率低、生长势差、作物减产。种子的纯度与真伪会直接影响作

物产量和农产品品质,与农业生产实际利益息息相关。种子分类与溯源是保证种子纯度与鉴别种子真伪的重要方法。

对种子品质的传统检测方法大多需要对种子进行破坏性分析,且检测时间长、过程复杂,难以满足现代农业对种子无损、快速生产的要求。随着化学计量学的发展和计算机技术的进步,近年来近红外光谱、高光谱成像等新型分析技术在无损检测领域实现了诸多成功的应用。近红外光谱技术具有无损、快速、高效等特点,可在不破坏样品的情况下,快速完成光谱数据采集,真正做到无损、快速分析;在数据模型的支持下,通过一次光谱数据扫描即可获得多个指标的分析结果,分析效率较传统方法明显提高。高光谱成像技术是

收稿日期: 2019-12-10, **修订日期:** 2020-04-12

基金项目: 北京市农林科学院科技创新能力建设专项储备性研究课题(KJCX20180409),北京工商大学食品安全大数据技术北京市重点实验室开放课题(BUBD-2017KF-11),科技部国家科技支撑计划项目(2014BAD04B05),科技部国家重点研发计划项目(2017YFD0201607)资助

作者简介: 王冬,1982年生,北京农业质量标准与检测技术研究中心副研究员 e-mail: wangd@brcast.org.cn

* 通讯作者 e-mail: hanp@brcast.org.cn; pubwu@163.com

一种将光谱技术和成像技术相结合的新型分析技术,在提供光谱信息的同时还可以提供样品的空间分布和图像特征信息,在对样品进行光谱信息采集的同时还可以实现样品的图像信息采集,从而为种子品质分析提供更多数据与信息。

本文围绕近红外光谱、高光谱成像等光学检测技术,针对种子品质无损快速检测,从种子质量评价、活力与老化检测、纯度与真伪鉴别、分类与溯源研究四个方面对近年文献进行综述,在此基础上,对种子品质无损快速检测技术进行总结与展望。

1 种子质量评价

种子质量关系着种子的健康状况,是种子等级划分以及衡量种子优劣的定量指标。种子质量主要包含种子含水率、蛋白含量、脂肪酸含量、淀粉含量等。其中,水分含量关系着种子储存过程的安全性。水分含量过高常会导致种子生虫、发霉,造成严重损失。近年来,很多学者针对种子水分含量的无损快速检测做了很多尝试^[1],并取得了一定的进展。芦兵^[2]等采用高光谱技术对水稻种子含水量采用模拟退火算法—支持向量回归建立了定量检测模型,预测集测定系数为 0.928 6。Zhang 等^[3]基于可见/近红外光谱和近红外高光谱成像技术实现对玉米种子水分含量的无损检测,采用无信息变量消除算法提取特征波长,建立偏最小二乘回归预测模型;模型校正集、预测集相关系数分别为 0.961 和 0.898,校正均方根误差、预测均方根误差分别为 0.491% 和 1.315%。康月琼等^[4]采用近红外光谱法对玉米、水稻种子的水分含量分别建立了校正模型,结果表明,玉米、水稻种子水分模型的多元相关系数分别为 0.837 1 和 0.832 1。

除水分含量外,很多学者对种子其他的质量指标研究了无损检测方法(请参阅本刊 35 卷第 2 期 346 页)。刘婷等^[5]采用近红外光谱技术建立了花生自然风干种子中的维生素 E 含量定量模型, $R^2=0.883 4$ 。唐月异等^[6]采用近红外光谱技术建立了花生自然风干种子中的蔗糖含量定量模型, $R^2=0.815 9$ 。张欣等^[7]采用近红外光谱技术建立了花生自然风干种子中的芥酸含量定量模型, $R^2=0.800 8$ 。刘盼等^[8]采用近红外光谱技术建立芝麻核心种质材料含油量和蛋白质含量的定量校正模型,含油量、蛋白质校正集相关系数(r)分别为 0.970 2 和 0.975 6。王纯阳等^[9]采用近红外光谱透射方式建立单粒稻种、糙米的蛋白质模型,稻种、糙米模型的测定系数(R^2)分别为 0.96 和 0.97。惠文凯等^[10]采用近红外光谱仪采集麻风树种子近红外光谱数据,建立麻风树种子含油量定量模型,校正集相关系数(r)为 0.955 6。陈俊锐等^[11]用近红外光谱仪对油菜种子进行质量分析,扫描波长范围 1 100~2 500 nm,结果表明,近红外光谱法测定油菜种子中芥酸、硫苷、含油量、油酸、水分测定值与化学法测定值的差异均不显著($p>0.05$),5 个指标相对误差分别为 2.03%,2.51%,0.89%,1.06% 和 2.50%,上述 5 个指标用近红外光谱法和化学法所测结果的相关系数分别是 0.999 9,0.999 6,0.997 7,0.992 2 和 0.966 4。Leonardo Velasco 等^[12]应用近红外反射光谱法估计油菜籽种子中脂肪酸组成,

通过气相色谱法对选定的样品进行分析,建立各个脂肪酸的校正模型,油酸,亚油酸,亚麻酸和芥酸的交互验证测定系数(R^2)值为 0.95~0.98;然而,棕榈酸和硬脂酸模型的交互验证 R^2 值始终低于 0.80;二十碳烯酸模型的交互验证 R^2 仅为 0.69,考虑到芥酸和二十碳烯酸之间的关系,改进后模型的 R^2 为 0.78~0.91。Wang 等^[13]采用 800~2 500 nm 波段的近红外光谱研究分析蚕豆种子营养成分定量测定,结合偏最小二乘回归算法建立了蛋白质、淀粉和总多酚的校正模型,蛋白质,淀粉和总多酚的测定系数(R^2)分别为 0.97,0.93 和 0.89。Zhang 等^[14]用近红外光谱建立了山茶籽的油脂含量模型,分别使用偏最小二乘回归和主成分回归建立了波长范围 4 000~10 000 cm^{-1} 的校正模型。结果表明,偏最小二乘回归的预测性能优于主成分回归,偏最小二乘回归的测定系数(R^2)、预测均方根误差分别为 0.94 和 0.495%。吴静珠等^[15]采用 876~1 729 nm 波段的近红外高光谱成像对小麦种子的含水量、蛋白含量和湿面筋含量建立可定量校正模型,并与传统近红外光谱数据所建模型进行比较。基于近红外高光谱数据,小麦种子含水量、蛋白含量、湿面筋含量模型的 R^2 分别为 0.931 2,0.941 5 和 0.970 3,RPD 值分别为 3.86,5.24 和 5.80;基于传统近红外光谱数据,小麦种子含水量、蛋白含量、湿面筋含量模型的 R^2 分别为 0.814 5,0.849 7 和 0.821 4,RPD 值分别为 2.35,2.67 和 2.53;结果表明,高光谱模型的准确性更高,预测性能更好。Francisco J Rodriguez-Pulido 等^[16]使用 950~1 650 nm 波段的近红外高光谱成像技术确定红葡萄和白葡萄种子中的黄烷醇含量,并使用对-二甲氨基基肉桂醛(DMACA)方法测定黄烷醇含量参考值。采用偏最小二乘回归算法建立校正模型,总黄烷醇含量的测定系数(R^2)为 0.73。Zhang 等^[17]使用 380~1 030 nm 波段的可见-近红外高光谱成像技术,通过采集油菜叶片的高光谱图像估计种子产量,采用偏最小二乘回归将平均光谱数据与相应的实际产量结合,测定系数(R^2)和预测均方根误差分别为 0.71 和 23.96。Wang 等^[18]采用 400~1 000 nm 波段的高光谱成像建立了不同保存时间的玉米种子硬度和弹性质地模型。采用偏最小二乘回归结合正交信号校正预处理算法建立校正模型,硬度、弹性、回弹性测定系数(R^2)分别为 0.901 2,0.874 4 和 0.847 7,预测均方根误差分别为 0.050 2,0.003 8 和 0.005 7;引入连续投影算法选择关键变量,建立正交信号校正-连续投影算法-偏最小二乘回归模型,硬度、弹性、回弹性的 R^2 分别为 0.836 5,0.821 7 和 0.793 0,预测均方根误差分别为 0.208 5,0.053 0 和 0.059 5。

在种子质量评价方面,近红外光谱技术可以给出较好的解决方案;而现有的关于高光谱成像技术在种子质量评价方面的报道多采用高光谱成像仪器采集种子样品的高光谱成像数据,再从成像数据中提取光谱信息建立校正模型,不仅可以实现对单粒种子的质量评价,而且可以得到准确度较高的结果。

2 种子活力与老化检测

种子活力(Vigor)是种子发芽和出苗率、幼苗生长的潜

势、植株抗逆能力和生产潜力的总和,是种子品质的重要指标。高活力种子具有明显的生长优势和生产潜力,是农产品丰收、增产的重要保障。在种子活力无损快速检测领域,很多学者做了有意义的尝试^[19],并取得了一定的研究进展。范雪婷等^[20]采用透射方式采集单粒水稻种子的近红外吸收光谱,以近红外光谱数据为自变量,以种子活力(发芽率)为因变量建立定量校正模型,“日本晴”种子活力模型校正集测定系数 $R^2 = 0.944 4$,“9311”种子活力模型校正集测定系数 $R^2 = 0.986 1$ 。Wu 等^[21]采用近红外光谱建立了小麦种子发芽率定量模型,校正集和预测集相关系数(r)分别为 0.902 和 0.967。Chen 等^[22]采用 850~1 700 nm 波段范围的近红外高光谱成像系统收集高光谱图像,检测小麦的发芽程度。实验分别在小麦发芽 0, 12, 24 和 48 h 进行。提取胚胎和胚乳的原始光强度,然后将其更改为反射率作分析。比较了不同部位,不同品种,不同发芽程度的小麦的图像和光谱信息。结果表明,同一种子在萌芽 12 h 后,胚的反射率低于胚乳的反射率;在所测波段范围内,同一品种的小麦种子在 12, 24 和 48 h 萌发时,反射率随发芽时间的增加而增加,与其体内脂肪含量的变化有关。彭彦昆等^[23]采用 500~900 nm 波段的高光谱图谱融合技术提出了一种番茄种子图像采集并辨识种子特征进而将种子分级的算法,通过标准发芽试验得到种子活力结果,基于连续投影算法求得反映番茄种子活力的特征波长为:535, 577, 595, 654, 684, 713, 744, 768, 809 和 840 nm;在 713 nm 波长下的图像特征对活力结果判断分级正确率最高,校正集、验证集的正确率分别为 93.75% 和 90.48%。张婷婷等^[24]采用 400~1 000 nm 波段的高光谱成像对单粒小麦种子生活力进行特征波段筛选,并建立判别模型。结果表明,采用无信息变量消除-竞争自适应重加权采样-连续投影算法从全波段光谱的 688 个变量筛选出了 8 个关键变量:473, 492, 811, 875, 880, 947 和 969 nm,基于上述 8 个关键变量所建偏最小二乘判别分析模型效果最优,校正集和预测集的小麦种子生活力整体鉴别正确率分别为 86.7% 和 85.1%,较全波段偏最小二乘判别分析模型分别提高了 4.2% 和 2.1%;校正集和预测集的活种子鉴别正确率分别为 93.8% 和 84.4%。对模型进行优化筛选,种子发芽率的最终鉴别正确率达到 93.1%。Maor Matzrafi 等^[25]基于 423.6~878.9 nm 波段的高光谱技术和数据分析,预测棕榈果种子的发芽率,发芽种子和非发芽种子高光谱分类准确度分别为 81.9% 和 76.4%。Ashabahebwa Ambrose 等^[26]采用傅里叶变换近红外光谱和拉曼光谱评估玉米种子活力,使用傅里叶变换近红外光谱仪在 1 000~2 500 nm 波长范围内采集样品近红外光谱数据,并采集 170~3 200 cm^{-1} 范围的拉曼光谱数据。结果表明,两种光谱在识别有活力种子和无活力种子方面的准确度均可接近 100%,其中近红外光谱建模结果优于拉曼光谱。

另一方面,种子活力的自然衰退过程称为种子老化,这一过程在高温、高湿条件下往往会加快,具体表现为种子变色、发芽率低、生长势差、作物减产,不利于农业生产。在种子老化无损检测方面,许多学者基于光谱及成像技术展开了探索。吴小芬等^[27]采用 874~1 734 nm 的近红外高光谱成像

提取了两种常见水稻种子未老化、老化 48 h、老化 72 h 的光谱反射率,基于全波段光谱建立了支持向量机判别分析模型。结果表明,未老化种子与老化种子可以准确识别,而老化 48 h 和老化 72 h 种子之间无法准确识别,与基于种子活力参数的测量结果相符,不同水稻品质对老化的反映存在差异。李美凌等^[28]对不同老化程度水稻种子,采用 400~1 000 nm 波段的高光谱成像结合主成分分析-支持向量机算法,研究比较了不同活力水平的水稻种子活力差异。通过主成分分析算法获得主成分图像并确定特征波段,用支持向量机算法建立水稻种子活力鉴别模型,预测判别率可达 100%。Christian Nansen 等^[29]基于 423.6~878.9 nm 波段的高光谱成像数据研究了确定三种澳大利亚原生树种 *Acacia cowleana* Tate, *Banksia prionotes* L. F. 和 *Corymbia calophylla* 的发芽能力的方法。在每个时间点对种子进行高光谱成像,以获取单个种子的反射谱。采用正向线性判别分析算法选择特征变量。在大约 10~30 d 的老化实验中,发芽率从 90% 以上下降至 20% 以下。每个物种的 P_{50} 值(50% 萌发)为 19.3 (*A. cowleana*), 7.0 (*B. prionotes*) 和 22.9 (*C. calophylla*)。杨小玲等(请参阅本刊 36 卷 12 期 4028 页)采用 400~1 000 nm 波段的高光谱成像研究成熟与未成熟玉米种子的鉴别方法,通过图像处理对种子进行分类。研究表明,采用主成分分析算法对高光谱图像数据进行分解,其中第二主成分图像差异最明显;采用 640 nm/525 nm 的波段比成像可减轻玉米种子冠部浅色部分误识别为种子成熟度较低的不利影响,平均正确识别率 93.9%。

综上所述,采用近红外光谱及高光谱成像技术对种子活力进行检测,可为种子质量筛查提供有效的技术手段,结合化学计量学算法,可建立准确度较高的校正模型,有效提高了对种子活力与老化的识别正确率,从而可以有效提高种子质量,进而对促进农业生产起到积极作用。

3 种子纯度与真伪鉴别

种子纯度与真伪鉴别是品种选育、品种检验以及农业生产的重要环节,种子的真伪直接影响到种子的储藏、销售、育种、生产等各个方面,直接影响作物产量和农产品品质。种子纯度鉴定是提高作物产量和质量的重要措施。种子纯度鉴定技术的研究是全国种子检验工作的重点和难点之一,也是种子管理过程中必须要及时解决的问题^[30]。种子的纯度与真伪鉴别已成为国内外研究的热点。随着光谱技术和计算机图像处理技术的发展,近年来,机器视觉技术以及近红外光谱技术在种子纯度与真伪鉴别方面得到了较为广泛和深入的研究和应用^[31],也取得了较好的成果。

王丽萍等^[32]采用近红外光谱技术对油菜杂交种子建立纯度定量预测模型,向油菜杂交种子中加入非杂交的母本或父本种子构成校正集样品,建立定量校正模型,校正集测定系数 $R^2 = 0.980 0$ 。徐涿频等^[33]采用近红外漫反射光谱结合判别式偏最小二乘分类筛选法,对单粒水稻种子“新两优 6 号”与其父本、母本和其他假种子进行了区分,所建模型的灵敏度 Sn、命中率 Pr 和马修斯相关系数 Mcc 分别为 97.92%,

97.58%和 95.51%。冉航等^[36]用 4 个短波近红外波段(820, 910, 1 000 和 1 090 nm)的透射光谱和 4 个中波近红外波段(1 150, 1 250, 1 350 和 1 450 nm)的反射光谱获取种子光谱图像并提取纹理特征来鉴定玉米杂交种纯度。结果表明,透射和反射模型对 5 个玉米品种平均正确鉴别率均在 85%以上。Timothy Wilkes 等^[35]通过 400~1 000 nm 波段的多光谱成像和高光谱成像快速区分硬质小麦和掺假普通小麦品种,并以较低的误差和良好的可重复性估算出掺假百分比。3%掺假样本的分析显示出较小的正偏差(样本估计为 3.65%, 95%±1.41%),相关的变异系数(CV)为 15.50%。

从以上内容可见,采用近红外光谱及高光谱成像等技术对种子纯度与真伪进行识别可以建立准确度较高的校正模型;在现有文献报道中,可以使用较少的波段而非全波段光谱数据实现上述功能,这意味着可以大幅减少计算量,从而可为降低仪器成本、提高仪器工作效率提供理论基础与技术支持。

4 种子分类与溯源研究

种子是农业的根本,被誉为“农业的芯片”。对种子进行分类与溯源研究是种子监管过程的重要措施,是保障种子安全的有效技术手段。种子分类与溯源则是保证种子纯度与鉴别种子真伪的重要方法。在种子分类与溯源研究方面,很多学者做了有意义的尝试。

郑田甜(请参阅本刊 35 卷 3 期 622 页)采用 600~1 100 nm 波段的可见-近红外反射光谱对 3 种代表性花生进行分类识别,结果表明,马氏距离判别分析模型对花生种子预测集判别准确率达到 95%。进而建立了花生种子脂肪含量分析模型,小波分析结合主成分回归模型的预测偏差小于 0.25,相对误差范围 0.03%~1.03%。钱丽丽等^[36]采用 4 000~12 000 cm^{-1} 波段范围的近红外光谱对 2013 年至 2015 年来自建三江地区及五常地区的 291 份大米进行产地溯源研究,研究发现,不同地区来源的样品在波段 5 136~5 501 cm^{-1} 处均有显著差异,说明不同地区样品的近红外光谱存在显著性差异;采用因子化法建立的定性分析模型及聚类分析模型对建三江大米及五常大米的正确判别率均高于 97.00%;利用偏最小二乘法建立的定量分析模型对两地区大米的正确判别率分别为 95.83%和 94.00%。进一步地,钱丽丽等^[37]采用 5 000~5 500 和 7 000~7 500 cm^{-1} 波段范围的近红外光谱结合聚类分析和偏最小二乘法对黑龙江省 3 个水稻主产区的地理标志大米进行产地溯源研究,结果表明,用聚类分析建立的模型对建三江、五常地域预测正确率为 100%,响水地域预测正确率为 95.83%;五常、响水地域判别正确率为 100%,建三江地域判别正确率为 95.83%。采用偏最小二乘法建立定量分析模型对建三江、五常、响水三个地域的预测正确率分别为 95.83%,100%和 95.83%。宋雪健等^[38]应用近红外漫反射光谱结合化学计量学对不同状态下的小米进行产地溯源研究,在 12 000~4 000 cm^{-1} 范围内采用因子化法建立的定性分析模型和在特征波段(9 400.9~5 447.7 和 4 600.6~4 249.8 cm^{-1}) 范围内采用偏最小二乘法(PLS)建

立的定量分析模型,对肇源、肇州两个小米主产区的小米籽粒和小米粉末的正确鉴别率均在 90%以上,其中小米粉模型正确预测率要高于小米籽粒模型。此外,宋雪健等^[39]采用近红外漫反射光谱技术结合偏最小二乘法,对齐齐哈尔、佳木斯、五常地区的水稻样品建立产地判别模型,在 7 501.3~5 447.7 和 4 600.6~4 249.8 cm^{-1} 波段范围内结合矢量归一化预处理方式建立稻谷样品判别模型,对 3 个地区样品的正确判别率分别为 82.35%,88.23%和 70.58%;在波段 7 501.3~4 597.8 cm^{-1} 范围内结合最小-最大归一化预处理方式建立大米粉样品判别模型,对三个地区样品的正确判别率分别为 88.23%,94.12%和 88.23%。周子立等^[40]采用 400~1 000 nm 的近红外光谱结合 BP 神经网络建立大米品种鉴别模型,对 150 个建模样本的拟合残差为 9.863×10^{-6} ,识别率达到 100%。Kong 等^[41]采用 1 039~1 612 nm 波段范围的近红外光谱结合多变量数据分析技术研究了水稻种子品种鉴定方法,比较了偏最小二乘判别分析、簇类独立软模式、K 最近邻算法、支持向量机、随机森林五种机器学习算法,结果表明,偏最小二乘判别分析和 K 最近邻算法模型的分类精度均超过 80%,簇类独立软模式、支持向量机、随机森林算法模型的分类精度均达到 100%;通过 PLS-DA 模型的加权回归系数选择了 12 个优化波长,基于 12 个优化波长建立的 K 最近邻算法、支持向量机、随机森林模型的正确率均超过 80%。Francisco J Rodriguez-Pulido 等^[42]采用 914~1 715 nm 波段范围的近红外高光谱成像对两种土壤中的两个红葡萄品种(Tempranillo 和 Syrah)和一个白色品种(Zalema)进行分类,使用偏最小二乘回归预测模型和主成分分析和一般判别分析,其测定系数高于 0.95。张航等^[43]采集了 7 个品种小麦种子高光谱图像及 900~1 700 nm 范围的光谱信息,建立了主成分分析-支持向量机分类模型。结果显示,3 个品种间种子分类准确率平均达到 95%以上;4 个品种间种子分类准确率平均达到 80%左右;6 个小麦品种种子的分类准确率仅 66%左右。刘小丹等^[44]采用近红外高光谱成像技术,在波段 874~1 734 nm,采集杂交水稻种子高光谱图像,采用 PCA 初步探究了 3 类样本的可分性。采用连续投影算法提取 7 个特征波长 985.08, 1 106.00, 1 203.55, 1 399.04, 1 463.19, 1 601.81 和 1 645.82 nm,基于特征波长建立偏最小二乘判别分析和支持向量机模型,正确识别率达到 90%以上,其中支持向量机模型效果优于偏最小二乘判别分析模型,而全谱判别模型结果优于特征波长判别模型。Gao 等^[45]采用 1 400~1 600 nm 波段的高光谱成像技术研究了小麦种子在贮藏过程中的特征变化。收集从 2007 年到 2012 年的小麦籽粒的高光谱成像数据,采用主成分分析分析了包括 6 年在内的小麦籽粒的光谱数据,用簇类独立软模式算法对不同年份的粮食进行分类,结果表明,相邻年份之间二分法的分类精度达到 97.05%,六年混合分类的精度达到 82.50%。张初等^[46]采用 874~1 734 nm 波长范围的近红外高光谱图像,通过提取西瓜种子的光谱反射率,结合平滑算法,经验模态分解算法和小波分析对提取出的光谱数据进行去除噪声处理,采用连续投影算法和遗传偏最小二乘法选择特征波段范围为 1 042~1 646 nm;基于全波段光谱建立了偏

最小二乘判别分析, 基于特征波长建立了反向传播神经网络判别模型和极限学习机判别模型, 建模集和预测集识别正确率均达到 100%。Gao^[47]等采用 874~1 734 nm 波段的近红外高光谱成像技术研究来自中国四个不同地理来源(江苏省、四川省、海南省、台湾省)的 240 个麻风树种子样品, 然后分别通过光谱和图像处理对获得的数据集进行分析。连续投影算法用于选择有效波长; 通过主成分分析缩小关注区域图像尺寸; 建立了最小二乘支持向量机分类模型, 样本预测正确识别率为 93.75%。Qiu^[48]等使用高光谱成像识别水稻种子品种, 获得了两个不同光谱范围(380~1 030 和 874~1 734 nm)的 4 个水稻种子品种的高光谱图像。提取了 441~948 和 975~1 646 nm 范围内的光谱数据, 使用不同数量的训练样本构建了 K 邻近、支持向量机和卷积神经网络模型; 实验结果表明, 用较长波长的反射率值构建的模型比用较短波长的反射率值构建的模型效果更好。Zhao^[49]等将高光谱成像系统用于玉米种子的品种分类, 共评估了 12 900 粒玉米种子, 包括 3 个不同的品种, 提取了 975.01~1 645.82 nm 的光谱数据; 建立了使用径向基函数的神经网络判别模型, 通过加载主成分进行最佳波长选择, 校正、预测准确度分别为 93.85% 和 91.00%。Zhu^[50]等使用波段范围 975~1 650 nm 的近红外高光谱成像对七种棉籽进行分类, 通过像素主成分分析形成的得分图像显示, 不同棉籽品种之间存在差异; 根据主成分分析的载荷数据选择有效波长, 使用自设计卷积神经网络和残差网络提取深层特征, 建立偏最小二乘判别分析、逻辑回归和支持向量机模型, 验证和预测的分类准确度均超过 80%。Zhao 等^[51]采集了三种葡萄籽在 874~1 734 nm 波段范围内的高光谱图像, 通过小波变换逐像素对光谱进行预处理, 并提取每个葡萄种子的光谱; 对高光谱图像进行主成分分析, 使用六个主成分的图像定性识别不同品种; 采用支持向量机建立判别模型, 校正、预测的准确度分别为 94.3% 和 88.7%。Feng 等^[52]使用波段范围 874.41~1 733.91 nm 的近红外高光谱成像从两个品种(huaidao-1 和 nanjing46)中识别 CRISPR/Cas9 诱导的水稻突变体; 采用主成分分析对高光谱成像数据进行分解, 并结合支持向量机和极限学习机建立分类判别模型; 进一步地, 采用主成分分析结合连续投影算法筛选光谱波长; 结果表明, 对于 huaidao-1, 校准集、预测集的分类准确率分别为 93.00% 和 92.75%, 对于 nanjing46, 校正集、预测集的分类准确率分别为 91.25% 和 89.50%。此外, Feng 等^[53]采用 874.41~1 733.91 nm 波段的近红外高光谱成像结合化学计量学数据对包含 cry1Ab/cry2Aj-G10evo 蛋白及其非转基因(GM)亲本的 GM 玉米籽粒进行分类, 采用偏最小二乘判别分析建立判别模型, 校正和预测准确度均接近 100%。Yang 等^[54]采用 924~1 657 nm 波段的高光谱成像对玉米进行分类研究。对 14 个品种的 1120 个玉米种子的高光谱图像, 基于无监督关键偏度波长选择算法选择 19 个关键波长, 变量数占全波长的 8.68%, 基于最小二乘支持向量机建立多模型策略用于品种识别; 结果表明, 用于全波长数据的多模型对测试集的分类准确度达到 98.18%, 高于单模型的 96.36%; 当使用由无监督关键偏度波长选择算法所选的 19 个关键波长时, 分类

精度为 96.57%。Yang 等^[55]基于从 430~980 nm 波段的可见和近红外高光谱图像中提取光谱、组合形态和外观特征, 建立了用于玉米种子品种分类的方法; 使用连续投影算法构建了光谱特征向量, 从每个玉米粒中提取形态特征——面积, 圆度, 纵横比, 实心度和纹理特征——能量, 对比度, 相关性, 熵及其标准偏差作为外观特征, 采用支持向量机和偏最小二乘判别分析算法建立分类模型; 与偏最小二乘判别分析模型相比, 支持向量机模型的识别准确率更高, 胚侧和背侧识别准确率分别为 98.2% 和 96.3%。Zhang 等^[56]用 380~1 030 nm 波段的可见光-短波近红外区域高光谱成像区分玉米种子的不同品种, 获取了六种玉米种子共 330 个样品的高光谱图像; 使用主成分分析和核主成分分析来探索光谱数据的内部结构, 选择三个关键波长——523, 579 和 863 nm, 根据最佳波长从每个单波长图像的灰度共生矩阵中提取四个纹理变量, 包括对比度、均匀性、能量和相关性, 最后通过最小二乘支持向量机和反向传播神经网络, 使用主成分、核主成分和质地特征相结合, 建立了玉米种子识别模型, 准确度为 98.89%。Huang 等^[57]基于 400~1 000 nm 波段的高光谱成像研究玉米种子分类方法。使用连续投影算法选择用于玉米种子品种分类的最佳波长, 选择特征区域, 然后引入主成分分析和多维缩放以变换/减少分类特征; 建立最小二乘支持向量机模型, 模型准确度达到 90% 以上, 优于使用原始光谱和图像特征模型的 83.68%, 而仅使用光谱特征模型的准确度仅为 76.18%。Wang 等^[58]将在 400~1 000 nm 波长范围内提取的光谱信息与机器学习结合, 对 10 个大豆品种进行分类。对原始数据进行预处理之后, 进行主成分分析, 从主成分分析选择的三个特征图像中提取出纹理特征参数; 对比了簇类独立软模式、偏最小二乘判别分析、遗传神经网络和 Takagi-Sugeno 模糊神经网络, 训练集平均准确度高于 96%, 预测集平均准确度高于 84%, 其中遗传神经网络模型的预测准确度最高, 达到 92%。Huang 等^[59]使用 400~1 000 nm 波段的高光谱成像建立玉米种子分类方法, 共评估了 2 000 粒种子, 包括不同年份的四个玉米种子品种; 使用最小二乘支持向量机建立了基于种子平均光谱特征的分类模型, 并使用增量支持向量数据描述, 以实现模型的在线更新, 分类准确率达到 94.4%, 比未更新模型高 10.3%。He 等^[60]采用 400~1 000 nm 波段的高光谱成像建立玉米种子分类模型, 基于最小二乘支持向量机分类器将三个不同年份种植的典型玉米品种的 2 000 个种子进行分类。模型总体正确率达到 98.3%。Wang 等^[61]采用 400~1 000 nm 波段的可见/近红外高光谱成像确定玉米种子的地理起源和年份。分别从包含两个玉米的高光谱图像的三个区域——胚芽, 胚乳和整个玉米籽粒提取光谱特征; 主成分分析显示不同玉米种子存在差异; 将偏最小二乘判别分析用于三个不同区域的光谱特征, 以识别玉米种子的起源和年份最高准确度达到 99.19%, 验证组准确度达到 98.44%。

在种子分类与溯源研究方面, 近红外光谱技术可在一定程度上提供解决方案, 而高光谱图像技术的介入则可提供更多的形态学数据信息, 并且可以针对单粒种子进行形态学分析, 因此可在一定程度上提高分类、溯源的准确度。然而,

光谱学和形态学数据的融合方法仍有待继续研究。

5 结 论

种子品质对于保障农业生产具有至关重要的作用。对种子品质进行无损快速检测不仅是现代农业生产的需要,而且是农产品丰收、增产的重要保障,同时对保证农产品数量安全具有重要意义。调研文献可知,近红外光谱技术在种子质量评价方面的应用很多。近红外光谱技术具有快速、无损、高效等优点,因此在种子内部化学物质含量的检测中可发挥强大的优势。然而,由于近红外光的穿透性有限,对厚度较大的种子的穿透性仍存在一定的局限;当种子内部品质变化引起种子表层组织化学性质改变时,采用近红外漫反射光谱对种子品质进行检测可取得较好的效果。拉曼光谱技术近年来也在种子品质检测方面有所应用,然而实验过程中的荧光

干扰仍是困扰拉曼光谱技术应用的难题。近红外高光谱成像技术不仅可以提供样品的光谱信息,而且还可以提供样品的空间分布信息和图像特征。近红外高光谱成像技术按光谱波段大致可分为近红外高光谱成像技术和可见-短波近红外高光谱成像技术;其中,近红外高光谱成像技术所提供的光谱信息的可解释性较强,而可见-短波近红外高光谱成像所携带的光谱信息相对而言其可解释性较弱。由此可见,不同波段的高光谱成像侧重点不同,近红外高光谱成像更侧重于样品的近红外光谱信息,而可见-短波近红外高光谱成像则更侧重于样品的形态学信息。因此,在种子活力与老化、纯度与真伪、分类与溯源方面,高光谱成像技术则体现出更为明显的优势。然而,光谱数据和形态学数据的融合仍是本领域的一个难点,采用何种融合方式以及融合数据的筛选将成为今后的研究重点。

References

- [1] WEI Li-feng, JI Jian-wei(魏利峰, 纪建伟). Journal of Chinese Agricultural Mechanization(中国农机化学报), 2016, 37(7): 80.
- [2] LU Bing, SUN Jun, YANG Ning, et al(芦 兵, 孙 俊, 杨 宁, 等). Journal of Southern Agriculture(南方农业学报), 2018, 49(11): 2342.
- [3] Zhang Y M, Guo W C. International Journal of Food Science & Technology, 2020, 55: 631.
- [4] KANG Yue-qiong, HAO Feng(康月琼, 郝 风). Seed(种子), 2004, 23(7): 10, 16.
- [5] LIU Ting, WANG Chuan-tang, TANG Yue-yi, et al(刘 婷, 王传堂, 唐月异, 等). Shandong Agricultural Sciences(山东农业科学), 2018, 50(6): 163.
- [6] TANG Yue-yi, WANG Xiu-zhen, LIU Ting, et al(唐月异, 王秀贞, 刘 婷, 等). Shandong Agricultural Sciences(山东农业科学), 2018, 50(6): 159.
- [7] ZHANG Xin, TANG Yue-yi, HU Dong-qing, et al(张 欣, 唐月异, 胡东青, 等). Shandong Agricultural Sciences(山东农业科学), 2018, 50(10): 138.
- [8] LIU Pan, ZHANG Yan-xin, LI Dong-hua, et al(刘 盼, 张艳欣, 黎冬华, 等). Chinese Journal of Oil Crop Sciences(中国油料作物学报), 2016, 38(6): 722.
- [9] WANG Chun-yang, MA Yu-han, LIU Bin-mei, et al(王纯阳, 马玉涵, 刘斌美, 等). Journal of Nuclear Agricultural Sciences(核农学报), 2019, 33(10): 2003.
- [10] HUI Wen-kai, WANG Yi, CHEN Xiao-yang(惠文凯, 王 益, 陈晓阳). Journal of Beijing Forestry University(北京林业大学学报), 2018, 40(1): 1.
- [11] CHEN Jun-kun, HE Ping, XU Chun, et al(陈俊锟, 何 萍, 徐 春, 等). Anhui Agricultural Sciences(安徽农业科学), 2013, 41(3): 985.
- [12] Velasco L, Becker H C. Euphytica, 1998, 101: 221.
- [13] Wang J J, Liu H, Ren G X. The Crop Journal, 2014, 2(1): 28.
- [14] Zhang K, Tan Z L, Chen C C, et al. Energy & Fuels, 2017, 7: 1.
- [15] WU Jing-zhu, LIU Qian, CHEN Yan, et al(吴静珠, 刘 倩, 陈 岩, 等). Transducer and Microsystem Technologies(传感器与微系统), 2016, 35(7): 42.
- [16] Rodríguez-Pulido F J, Hernández-Hierro J M, Nogales-Bueno J, et al. Talanta, 2014, 122: 145.
- [17] Zhang X L, He Y. Industrial Crops and Products, 2013, 42: 416.
- [18] Wang L, Pu H B, Sun D W, et al. Food Analytical Methods, 2015, 8(6): 1535.
- [19] LI Xiao-fan, WANG Cheng, SONG Peng, et al(李孝凡, 王 成, 宋 鹏, 等). Seed(种子), 2019, 38(6): 61.
- [20] FAN Xue-ting, ZHU Ming-dong, YANG Chen-guang, et al(范雪婷, 朱明东, 杨晨光, 等). Hybrid Rice(杂交水稻), 2019, 34(4): 62.
- [21] Wu J Z, Dong W F, Liu Q, et al. Transactions of the Chinese Society of Agricultural Engineering, 2015, 31(Supp. 2): 272.
- [22] Chen J Y, Chen H H, Wang X D, et al. The Characteristic of Hyperspectral Image of Wheat Seeds During Sprouting. 7th International Conference on Computer and Computing Technologies in Agriculture (CCTA) 2013, pp 408-421 (10.1007/978-3-642-54344-9-47). (hai-01220944).
- [23] PENG Yan-kun, ZHAO Fang, BAI Jing, et al(彭彦昆, 赵 芳, 白 京, 等). Transactions of the Chinese Society for Agricultural Machinery(农业机械学报), 2018, 49(2): 327.

- [24] ZHANG Ting-ting, XIANG Ying-ying, YANG Li-ming, et al(张婷婷, 向莹莹, 杨丽明, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2019, 39(5): 1556.
- [25] Matzrafi M, Herrmann I, Nansen C, et al. *Frontiers in Plant Science*, 2017, 8: 474.
- [26] Ambrose A, Lohumi S, Lee W H, et al. *Sensors and Actuators B: Chemical*, 2016, 224: 500.
- [27] WU Xiao-fen, ZHAO Guang-wu, QI Heng-nian(吴小芬, 赵光武, 祁享年). *Journal of Anhui Agricultural Sciences(安徽农业科学)*, 2017, 45(29): 12.
- [28] LI Mei-ling, DENG Fei, LIU Ying, et al(李美凌, 邓 飞, 刘 颖, 等). *Acta Agriculturae Zhejiangensis(浙江农业学报)*, 2015, 27(1): 1.
- [29] Nansen C, Zhao G P, Dakin N, et al. *Journal of Photochemistry and Photobiology B: Biology*, 2015, 145: 19.
- [30] WU Xiu-ting(吴秀婷). *Agriculture Engineering Technology(农业工程技术)*, 2017, 10: 81.
- [31] WEI Li-feng, JI Jian-wei(魏利峰, 纪建伟). *Hubei Agricultural Sciences(湖北农业科学)*, 2016, 55(21): 5445, 5478.
- [32] WANG Li-ping, ZHAO Xing-zhong, CHEN Wen-jie, et al(王丽萍, 赵兴忠, 陈文杰, 等). *Chemical Analysis and Meterage(化学分析计量)*, 2017, 26(5): 43.
- [33] XU Zhuo-pin, FAN Shuang, CHENG Wei-min, et al(徐涿频, 范 爽, 程维民, 等). *Chinese Agricultural Science Bulletin(中国农学通报)*, 2017, 33(2): 142.
- [34] RAN Hang, CUI Yong-jin, JIN Zhao-xi, et al(冉 航, 崔永进, 靳召晰, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2017, 37(9): 2743.
- [35] Wilkes T, Nixon G, Bushell C, et al. *Food and Nutrition Sciences*, 2016, 7: 355.
- [36] QIAN Li-li, SONG Xue-jian, ZHANG Dong-jie, et al(钱丽丽, 宋雪健, 张东杰, 等). *Food Science(食品科学)*, 2018, 39(16): 321.
- [37] QIAN Li-li, SONG Xue-jian, ZHANG Dong-jie, et al(钱丽丽, 宋雪健, 张东杰, 等). *Journal of the Chinese Cereals and Oils Association(中国粮油学报)*, 2017, 32(10): 185, 196.
- [38] SONG Xue-jian, QIAN Li-li, ZHOU Yi, et al(宋雪健, 钱丽丽, 周 义, 等). *Food Research and Development(食品研究与开发)*, 2017, 38(11): 134.
- [39] SONG Xue-jian, QIAN Li-li, ZHOU Yi, et al(宋雪健, 钱丽丽, 周 义, 等). *Farm Products Processing(农产品加工)*, 2017, (5): 13.
- [40] ZHOU Zi-li, ZHANG Yu, HE Yong, et al(周子立, 张 瑜, 何 勇, 等). *Transactions of the Chinese Society of Agricultural Engineering(农业工程学报)*, 2009, 25(8): 131.
- [41] Kong W W, Zhang C, Liu F, et al. *Sensors*, 2013, 13: 8916.
- [42] Rodríguez-Pulido F J, Barbin D F, Sun D W, et al. *Postharvest Biology and Technology*, 2013, 76: 74.
- [43] ZHANG Hang, YAO Chuan-an, JIANG Meng-meng, et al(张 航, 姚传安, 蒋梦梦, 等). *Journal of Triticeae Crops(麦类作物学报)*, 2019, 39(1): 96.
- [44] LIU Xiao-dan, FENG Xu-ping, LIU Fei, et al(刘小丹, 冯旭萍, 刘 飞, 等). *Transactions of the Chinese Society of Agricultural Engineering(农业工程学报)*, 2017, 33(22): 189.
- [45] Dong G, Guo J, Wang C, et al. *International Journal of Agricultural and Biological Engineering*, 2017, 10(2): 251.
- [46] ZHANG Chu, LIU Fei, KONG Wen-wen, et al(张 初, 刘 飞, 孔汶汶, 等). *Transactions of the Chinese Society of Agricultural Engineering(农业工程学报)*, 2013, 29(20): 270.
- [47] Gao J F, Li X L, Zhu F L, et al. *Computers and Electronics in Agriculture*, 2013, 99: 186.
- [48] Qiu Z J, Chen J, Zhao Y Y, et al. *Applied Sciences*, 2018, 8: 212.
- [49] Zhao Y Y, Zhu S S, Zhang C, et al. *RSC Advances*, 2018, 8: 1337.
- [50] Zhu S S, Zhou L, Gao P, et al. *Molecules*, 2019, 24: 3268.
- [51] Zhao Y Y, Zhang C, Zhu S S, et al. *Molecules*, 2013, 23: 1352.
- [52] Feng X P, Peng C, Chen Y, et al. *Scientific Reports*, 2017, 7: 15934.
- [53] Feng X P, Zhao Y Y, Zhang C, et al. *Sensors*, 2017, 17: 1894.
- [54] Yang S, Zhu Q B, Huang M, et al. *Food Anal. Methods*, 2017, 10: 424.
- [55] Yang X L, Hong H M, You Z H, et al. *Sensors*, 2015, 15: 15578.
- [56] Zhang X L, Liu F, He Y, et al. *Sensors*, 2012, 12(12): 17234.
- [57] Huang M, He C J, Zhu Q B, et al. *Applied Sciences*, 2016, 6: 183.
- [58] Wang R T, Tan K Z, Li M Y, et al. *Journal of Computational Methods in Sciences and Engineering*, 2019, 19: 1001(doi: 10.3233/JCM-193562).
- [59] Huang M, Tang J Y, Yang B, et al. *Computers and Electronics in Agriculture*, 2016, 122: 139.
- [60] He C, Zhu Q, Huang M, et al. *American Society of Agricultural and Biological Engineers*, 2016, 59(6): 1529.
- [61] Wang Q G, Huang M, Zhu Q B. Characteristics of Maize Endosperm and Germ in the Geographical Origins and Years Identification Using Hyperspectral Imaging. *American Society of Agricultural and Biological Engineers Annual International Meeting, ASABE*, 2014, 6: 4420.

Progress in Research on Rapid and Non-Destructive Detection of Seed Quality Based on Spectroscopy and Imaging Technology

WANG Dong^{1,3}, WANG Kun², WU Jing-zhu^{2*}, HAN Ping^{1,3*}

1. Beijing Research Center for Agricultural Standards and Testing (BRCAS), Beijing Academy of Agricultural and Forestry Sciences, Beijing 100097, China
2. Beijing Key Laboratory of Big Data Technology for Food Safety, Beijing Technology and Business University, Beijing 100048, China
3. Laboratory of Quality & Safety Risk Assessment for Agro-products (Beijing), Ministry of Agriculture and Rural Affairs, Beijing 100097, China

Abstract Seed is an important means of production in the process of agricultural production. The quality evaluation, vigor and aging detection, purity and authenticity identification, classification and traceability are common problems in seed quality detection. Seed quality mainly includes the content of moisture, protein, fatty acid, starch, and so forth, which is the important indices of seed quality classification. Moreover, seed quality is related to the safety of storage. Seed vigor is the sum of seed germination and emergence rate, seedling growth potential, plant stress resistance and production potential. High vigor seeds are of obvious growth advantage and production potential. Seed aging refers to the natural decline of seed vigor, which is manifested by seed discoloration, low germination rate, poor growth potential and crop yield reduction. The purity and authenticity of seeds will affect crop yield and agricultural product quality. Seed classification and traceability is an important method to ensure the purity and identify the authenticity of seeds, by which, crop yield and product quality will be guaranteed. For seed quality analysis, it usually need to do irreversible destructive analysis on samples according to the traditional methods, which is time-consuming with complex procedures. It is obvious that traditional methods are difficult to meet the needs of modern agriculture for seed production. Therefore, it is an urgent problem to carry out the research on non-destructive and rapid detection technology of seed quality. In recent years, with the development of chemometrics and the progress of computer technology, near-infrared spectroscopy, with the advantages of fast, non-destructive and high efficiency, has been widely applied in the non-destructive and rapid analysis of agricultural products, food, agricultural inputs, and so on. In addition, combined with spectral and imaging technology, hyperspectral imaging technology is rising in recent years. Compared with the traditional spectral technology, hyperspectral imaging technology can acquire not only the spectral information of the sample but also the spatial distribution information and image characteristics of it. In this paper, based on the near-infrared spectroscopy and hyperspectral imaging technology, the literature of seed quality non-destructive detection from the aspects of seed quality evaluation, vigor and aging detection, purity and authenticity identification, classification and traceability research were reviewed. Based on the analysis of the characteristics of different detection technologies, the problems of seed quality detection are sorted out, respectively. Furthermore, the technical characteristics of non-destructive and rapid detection of seed quality are summarized and prospected.

Keywords Seed; Near-infrared spectroscopy; Hyperspectral imaging; Non-destructive detection

(Received Dec. 10, 2019; accepted Apr. 12, 2020)

* Corresponding authors