

高光谱成像的马铃薯叶片含水率分布可视化

孙红¹, 刘宁¹, 吴莉¹, 郑涛¹, 李民赞^{1*}, 吴静珠²

1. 中国农业大学现代精细农业系统集成研究教育部重点实验室, 北京 100083

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

摘要 为了快速检测马铃薯叶片的水分含量, 并探究受到干旱胁迫时叶片含水率变化情况, 利用高光谱成像对马铃薯叶片含水率进行检测和可视化研究。采集 71 个叶片, 用烘干法对叶片水分梯度进行控制, 共得到 355 个样本。使用高光谱分选仪器采集叶片 862.9~1 704.2 nm(256 个波长)的光谱成像数据, 采用称重法测量含水率。利用 Sample set partitioning based on joint X-Y distance(SPX Y)算法将总样本按照 2:1 的比例划分为建模集(240 个样本)和验证集(115 个样本)。对采集的数据进行光谱特征分析, 本文分别用 CA 和 RF 两种算法, 各筛选得到 15 个特征波长。基于 CA 筛选出相关系数高于 0.96 的 15 个波长分别为 1 406.82, 1 410.12, 1 403.62, 1 413.32, 1 416.62, 1 419.82, 1 400.32, 1 423.12, 1 426.32, 1 429.62, 1 432.82, 1 436.12, 1 439.32, 1 442.52 和 1 445.8 nm。基于 RF 算法筛选被选概率高于 0.3 的 15 个特征波长, 按照被选择概率值从大到小排列, 分别为 1 071.62, 1 041.12, 1 222.52, 1 465.22, 1 397.02, 1 449.02, 1 034.32, 1 523.22, 976.42, 1 172.52, 979.82, 1 165.82, 1 037.72, 1 426.32 和 869.8 nm。用 CA 和 RF 算法筛选到的特征波长建立 PLSR 模型, 分别记为 CA-PLSR 模型和 RF-PLSR 模型。利用高精度模型检测结果, 对马铃薯叶片含水率进行可视化分析, 首先计算马铃薯叶片图像每个像素点的含水率, 得到灰度图像, 然后对灰度图像进行伪彩色变换, 绘制出叶片含水率可视化彩色图像。为了体现马铃薯叶片烘干处理中含水率变化进程, 用 HSV 彩色模型对样本叶片的伪彩色图像进行分割, 获得分割图像结果, 显示出在某含水率区间的叶片面积比例。结果显示, CA 算法选取的 15 个波长均在 1 400.3~1 450.0 nm 范围内, CA-PLSR 模型的建模精度(R^2_c)为 0.975 5、建模集均方根误差(RMSEC)为 2.81%, 验证集精度(R^2_v)为 0.933 2、验证集均方根误差(RMSEV)为 2.31%。RF 算法选取的特征波长分布范围较 CA 法选取范围广, 具有局部“峰谷”特性, 且 RF-PLSR 模型的建模集精度(R^2_c)为 0.983 2、RMSEC 为 2.32%, 验证集精度(R^2_v)为 0.947 1、RMSEV 为 2.15%。选取 RF-PLS 模型计算马铃薯每个像素点的含水率, 得到伪彩色变换图像, 观察可知随着烘干时间的增加含水率逐渐下降; 并能够从叶片结构角度看到, 随着水分胁迫的加强, 叶片从边缘开始失水, 逐渐向叶片中间蔓延, 其中叶茎和叶脉的含水率较其他部位高。计算得到叶片伪彩色图像中含水率大于 90%, 80%和 70%的像素点占整个叶片图像的比例。利用高光谱成像技术可以实现马铃薯叶片的含水率检测与分布可视化表达, 为监测马铃薯生长状况以及叶片含水率分析提供新的理论根据。

关键词 含水率; 马铃薯叶片; 高光谱成像; 图像分割; 可视化

中图分类号: O433.4 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2019)03-0910-07

引言

水分是影响马铃薯作物生长发育的主要因素之一, 水分的缺失将会影响马铃薯植株和叶片的生理与形态结构, 进而

影响生长、产量与质量。传统的作物水分测量方法有烘干法、蒸馏法等, 精度高, 但过程繁琐, 测量周期长, 具有破坏性^[1-3]。近红外光谱给出分子中含氢基团(O—H, N—H, C—H)振动的合频与各级倍频的吸收信息, 通过样品的近红外光谱, 可以得到样品中含氢基团的特征信息^[3-6], 因而光

收稿日期: 2018-05-15, 修订日期: 2018-10-12

基金项目: 国家自然科学基金项目(31501219), 国家“十三五”重点研发计划课题(2016YFD0300600-2016YFD0300606, 2016YFD0300600-2016YFD0300610), 重点实验室课题(BKBD-2017KF03, 2018TC020)和青产学研 2017-12 资助

作者简介: 孙红, 女, 1980 年生, 中国农业大学信息与电气工程学院副教授 e-mail: sunhong@cau.edu.cn

* 通讯联系人 e-mail: limz@cau.edu.cn

谱分析法可支持马铃薯作物水分无损检测。

基于光谱吸收、透射等方法,一些学者利用水分敏感波长组合、植被指数等方法进行植物的水分含量检测^[7-8]。针对玉米叶片含水率检测,陈香等^[9]提取 800, 1 058, 1 323 和 1 423 nm 为水分敏感波长,用差值植被指数 DVI(1 423, 800 nm)和透射光谱 $T_{1\ 323}$ 和 $T_{1\ 058}$ 建立了水分多元线性回归模型。Das 等^[10]提取 1 391 和 1 830 nm 为水分敏感波长,构建了比值植被指数 RVI($R_{1\ 391}$, $R_{1\ 830}$)和归一化差异光谱指数 NDSI($R_{1\ 391}$, $R_{1\ 830}$),对小麦的相对含水量进行检测。上述研究者主要基于采样点一维光谱数据展开分析,对二维叶面上各位置的水分状态与分布情况没有精细化分析。

高光谱成像结合了光谱与图像,能够同时获得被检测目标的光谱和图像,可以精确地检测图像中每一个像素点的光谱数据以及图像信息,能够为可视化与检测提供技术依据。近年来国内外一些学者应用高光谱成像与可视化技术对作物病虫害以及叶绿素含量展开了研究^[11-13],在马铃薯作物叶片含水率定量检测方面还没有深入的讨论。

为了检测马铃薯叶片水分,作水分烘干实验,采集高光谱图像。首先提取水分吸收特征波长、建立马铃薯植株叶片含水率的检测模型,然后应用伪彩色处理技术绘制马铃薯叶片含水率分布可视化图像,最后基于图像分割技术提取叶片含水率梯度变化的面积百分比,以期为监测马铃薯生长状况以及叶片含水率分析提供新的途径。

1 实验部分

1.1 马铃薯叶片样本及样本含水率测定

2017 年 9 月,在中国农业大学信息与电气工程学院实验温室,培育马铃薯植株,选取水分管理关键时期(发棵期),从马铃薯植株上采集 71 个叶片。

将样本带回实验室,对每个叶片称重,质量记为 M_{f_0} ,扫描高光谱图像。然后将叶片放入 40 °C 烘箱,40 min 后取出,称其质量记为 M_{f_1} ,进行高光谱扫描。重复测量 3 次,每次烘干时间均为 40 min,分别得到质量 M_{f_2} , M_{f_3} 和 M_{f_4} ,最后将叶片放入 80 °C 的烘箱中,烘干至恒重,质量记为 M_d 。叶片含水率计算公式如式(1)

$$LWC_n = \frac{M_{f_n} - M_d}{M_{f_n}} \times 100\% \quad (1)$$

式中, LWC_n 为测量第 n 次的叶片含水率; M_{f_n} 为第 n 次称重质量; n 取值范围 0~4; M_d 为叶片干重。对鲜叶共进行 4 次烘干处理,得到含有鲜叶和烘干叶在内的 5 批×71 个叶片,共 355 个样本。

1.2 马铃薯叶片高光谱数据采集

采用 Gaia 高光谱成像系统采集马铃薯叶片光谱成像数据,结构如图 1 所示,系统主要由 OL23 型镜头、二维 CCD 面阵列、LT365 型侦测器、V10E 型光谱仪、均匀光源、可控移动载物台、计算机及相应配套控制软件组成。光谱范围为 862.9~1 704.2 nm,光谱分辨率为 2.8 nm,采样间隔为 0.65 nm。测量前预热 30 min 以消除基线漂移误差,设置系统曝光时间为 15 ms,载物台的移动速度为 5.5 mm·s⁻¹。

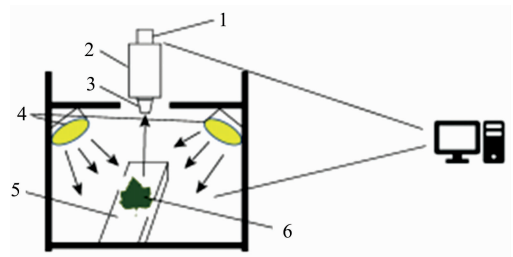


图 1 高光谱成像系统原理图

1: CCD; 2: 光栅光谱仪; 3: 成像镜头;
4: 光源; 5: 一维电移台; 6: 样品

Fig. 1 Schematic diagram of hyperspectral imaging system

1: Charge coupled device; 2: Grating spectrometer; 3: Imaging lens;
4: Light source; 5: Mobile station; 6: Sample

为了消除因光照不均匀、侦测器内部暗电流空间衍射效率分布差异及镜头不同位置透过率的差异等因素对采集数据的影响,在提取光谱成像数据前需要进行黑板校正,公式如式(2)

$$R = \frac{I - B}{W - B} \quad (2)$$

式中, I 为原始的马铃薯叶片数据, B 为黑板数据, W 为白板数据, R 为校正后的铃薯叶片数据。

在软件 ENVI5.1 中打开校正过的马铃薯叶片高光谱图像数据,计算整个叶片的反射光谱平均值作为此样本的光谱数据,最终得到 355 个样本×256 个波长的光谱数据矩阵,以用于数据分析、建模。

1.3 基于 SPXY 的样本集划分算法

采用 SPXY(sample set partitioning based on joint X-Y distance)算法划分模型的建模集和验证集。它是基于统计学角度的一种样本集划分方法,综合考虑光谱和化学性质的差异来选择建模集,先按照式(3)计算所有样本光谱数据两两之间的欧氏距离,选择距离最大的两个作为前两个建模集样本。

$$d_x(p, q) = \sqrt{\sum_i [x_p(i) - x_q(i)]^2}; p, q \in [1, N] \quad (3)$$

式中, $x_p(i)$ 和 $x_q(i)$ 是 p 和 q 样本在 i 波长处的光谱参数, I 是光谱的波长个数, N 是样本个数。

然后分别计算剩余的样本与已选择的两个样本之间的距离,对于每个剩余样本而言,其与已选样品之间的最短距离被选择,然后选择这些最短距离中相对最长的距离所对应的样本,作为第三个样品。重复以上步骤,直至所选的样品的个数等于事先确定的数目为止。

在式(3)的基础上考虑了性质空间因素 $d_y(p, q)$ 即

$$d_y(p, q) = \sqrt{(y_p - y_q)^2}; p, q \in [1, N] \quad (4)$$

式(4)中, y_p 和 y_q 是 p 和 q 样本的性质参数。

为了确保样本在光谱空间和性质空间由相同的权重,将 $d_x(p, q)$ 和 $d_y(p, q)$ 分别除以它们在数据集中的最大值,标准化的 xy 距离公式为

$$d_{xy}(p, q) = \frac{d_x(p, q)}{\max_{p, q \in [1, N]} d_x(p, q)} +$$

$$\frac{d_y(p, q)}{\max_{p, q \in [1, N]} d_y(p, q)} \quad (5)$$

1.4 特征波长选取方法

为了精简模型、提高模型精度,需要进行特征变量的筛选。对比了相关性分析(correlation analysis, CA)和随机蛙跳算法(random frog, RF)进行特征波长的筛选。

RF 算法是一种特征变量选取的方法。该算法类似于可逆跳转马尔可夫链蒙特卡洛(reversible jump Markov Chain Monte Carlo, RJMCMC),通过模拟一条服从稳态分布的马尔可夫链来计算每个变量被选择的概率,从而进行重要变量的筛选^[14]。该算法能够利用较少的变量迭代建模,是一种有效的高位数据变量选取方法。RF 算法主要的运算思想包括以下三步:

(1)输入一个初始变量子集 V_0 , 初始化时包含 N 个变量;

(2)基于原始的变量子集 V_0 , 提出一个候选变量子集 V^* , 包含 N^* 个变量;选择 V^* 作为 V_1 代替原始的变量子集 V_0 。直到 M 次迭代终止, 计算完成;

(3)计算每个变量的选择概率, 以此作为筛选变量的标准, 概率越大说明这个变量越具有代表性。

1.5 建模方法

偏最小二乘回归(partial least squares regression, PLSR)是光谱分析中应用最广泛的一种建模方法, 基于主成分提取思想, 能够解决变量间的自相关和多重共线性问题。PLSR 同时对光谱反射率矩阵与叶片含水率矩阵进行主成分分解, 分解过程中将光谱矩阵和叶片含水率矩阵相关联, 建立二者之间的线型回归模型, 以用来检测马铃薯叶片含水率。采用留一交叉验证法(leave-one-out cross validation, LOOCV)进行内部交叉验证, 以交叉验证均方差(root mean square error of cross validation, RMSECV)为标准选取最优特征变量数(optLV)。使用建模集、验证集模型决定系数 R_c^2 , R_v^2 以及均方根误差 RMSEC 和 RMSEV 为评价指标。

1.6 马铃薯叶片含水率的可视化

高光谱图像数据上每一个像素点都有一条包括全波长的光谱反射率曲线, 将每个像素点上的光谱数据代入到建立的 PLSR 模型中计算出相应像素点的含水率, 得到灰度图像; 再利用 ENVI 软件中的工具包 ENVI Classic+IDL 将不同的含水率用不同的颜色表示, 绘制马铃薯叶片含水率的可视化分布图, 具体步骤如下:

(1)提取特征波长下马铃薯叶片高光谱图像。

(2)提取相应特征波长图像中每个像素的光谱反射率值。

(3)依据 PLSR 模型计算叶片每个像素点的含水率值, 形成灰度图像。

(4)对灰度图像进行伪彩色处理, 得到马铃薯叶片含水率可视化彩色分布图。

利用 HSV 彩色模型进行图像分割提取出叶片含水率梯度分布与面积百分比。其中 HSV 模型是根据颜色的直观特性而创建的一种颜色空间, 这个模型中颜色的参数分别是: 色调(hue, H), 饱和度(saturation, S), 明度(value, V)。根据 H 分量对伪彩色图像进行分割处理。H 分量的取值用角

度量, 取值范围为 $0^\circ \sim 360^\circ$, 从红色开始按逆时针方向计算, 红色为 0° , 绿色为 120° , 蓝色为 240° 。通过合理设定 H 分量阈值, 提取相应区间的图像颜色。

2 结果与讨论

2.1 马铃薯叶片高光谱特性曲线

用 ENVI 软件提取整个叶片的反射率平均值作为此样本的光谱数据, 分别提取每个样本在 862.9~1 704.2 nm 范围内的反射率光谱。对采样新鲜叶片进行烘干处理, 按照处理时间划分采集 5 批数据, 共 355 个样本, 每批样本的平均光谱如图 2 所示, 总体而言各样本在 900.9~1 300.3 nm 保持较高的反射率, 其中由于 C—H 第二倍频的吸收, 在 1 170 nm 附近出现小反射率波谷; 由于在 1 225 和 1 360 nm 处分别存在 C—H 键的二级倍频和一级倍频的吸收, 在 1 300 nm 附近出现波峰。由于水分子 O—H 键在 1 450 nm 附近的强吸收, 导致反射率在 1 300.3~1 452.3 nm 急剧下降, 在 1 450 nm 附近有明显的波谷存在。

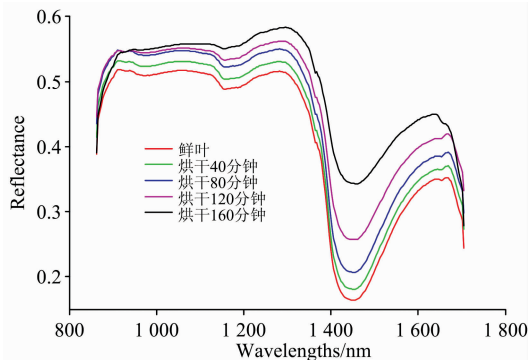


图 2 马铃薯叶片反射光谱曲线

Fig. 2 Reflectance spectra of potato leaves

2.2 马铃薯叶片含水率统计

总体 355 个样本, 含水率在 11.76%~92.66% 区间, 平均值为 74.49%, 标准差为 16.18%。根据 SPXY 算法对样本划分建模集和验证集, 240 个样本为建模集, 115 个样本为验证集, 划分结果如表 1 所示, 其中建模集的叶片含水率取值范围为 11.76%~92.66%, 覆盖了验证集的叶片含水率值, 说明用 SPXY 算法划分得到的建模集和验证集是合理的, 可以用于后续的建模。

表 1 建模集与验证集划分统计 (%)

Table 1 Statistical results of calibration set and validation set (%)

数据集	样本量	最大值	最小值	平均值	标准差
总样本	355	92.66	11.76	74.49	16.18
建模集	240	92.66	11.76	70.47	18.00
验证集	115	90.24	59.62	82.88	5.24

2.3 特征波长筛选

采集马铃薯叶片光谱范围为 862.9~1 704.2 nm, 应用

并比较了基于统计相关分析(CA)和基于仿生学的 RF 算法进行特征波长选取。

2.3.1 基于相关系数筛选特征波长

对马铃薯叶片含水率值与光谱反射率进行相关性分析,处理结果如图 3 所示,图中显示每个波长点处相关系数的大小。以相关系数的绝对值从大至小排列,筛选出相关系数高于 0.96 的 15 个波长作为特征波长,分别为 1 406.82, 1 410.12, 1 403.62, 1 413.32, 1 416.62, 1 419.82, 1 400.32, 1 423.12, 1 426.32, 1 429.62, 1 432.82, 1 436.12, 1 439.32, 1 442.52 和 1 445.8 nm, 提取得到的 15 个特征波长均在 1 400.3~1 450.0 nm 范围内,此波段内存在 O—H 键的特征吸收波长,表征了水分吸收峰处的光谱特性。

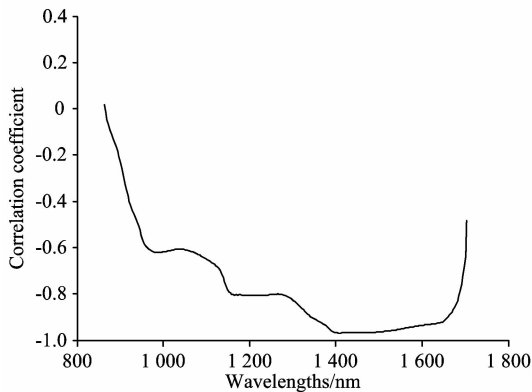


图 3 光谱反射率与叶片含水率相关系数曲线
Fig. 3 Correlation between spectral reflectance and leaf water content

2.3.2 基于 RF 算法筛选特征波长

Random Frog (RF) 算法与 PLSR 方法相结合,根据 PLSR 模型中每个变量的回归系数绝对值大小作为每次迭代过程中该变量是否被选择或者提出的依据。然后,基于不同的波长点具有不同的选择可能性进行特征波长的提取,运算后结果如图 4,横坐标是波长,纵坐标是被选择概率,概率是评价波长变量重要性的标准,概率越大的波长变量越重要。

设定 0.3 作为叶片含水率对应的特征波长选择的阈值,最终共筛选出 15 个特征波长,按照被选择概率值从大到小排列,分别为 1 071.62, 1 041.12, 1 222.52, 1 465.22, 1 397.02, 1 449.02, 1 034.32, 1 523.22, 976.42, 1 172.52, 979.82, 1 165.82, 1 037.72, 1 426.32 和 869.8 nm。对照图 4 与图 3,总体而言 RF 提取波长位于 860~1 550 nm 范围之内,包括 CA 筛选的 1 400.3~1 450.0 nm 范围内的 1 426.32 和 1 449.02 nm;此外,RF 算法筛选的波长还包括了其他范围的相关性局部“峰谷特征”,如 976.42, 979.82 nm 和 1 165.82, 1 172.52 nm。其中,976.42, 979.82 nm 位于 970 nm 附近,体现了 O—H 键吸收特征;1 165.82 和 1 172.52 nm 位于 1 150~1 180 nm 范围,是 O—H 键和 C—H 键吸收特征波长。综上分析可知,与 CA 法相比较,RF 算法能够更加全面地提取作物体现含氢基团的特征。

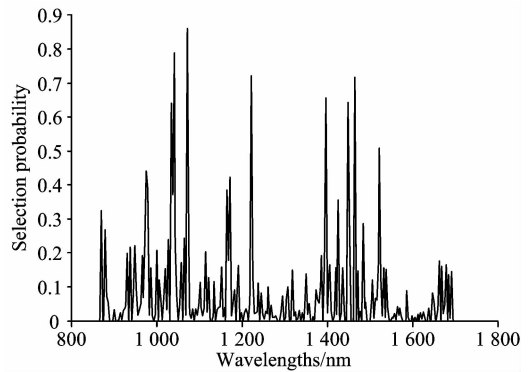


图 4 采用 RF 进行特征波长提取结果
Fig. 4 Results of selection variables using RF

2.4 PLSR 回归模型建立

分别将 CA 和 RF 算法提取到的特征波长对应的光谱反射率数据作为 X 变量,将马铃薯叶片的含水率值作为 Y 变量,建立马铃薯叶片含水率诊断 PLSR 模型。

为了消除光谱信息冗余和多重共线性对模型精度的影响,两种模型均采用留一交互验证法 (leave-one-out cross validation, LOOCV) 进行内部交互验证。结果为:CA-PLSR 模型在主成分个数为 10 时, RMSECV 最小为 3.14%, RF-PLSR 在主成分个数为 10 时, EMSECV 最小为 2.53%。CA-PLSR 和 RF-PLSR 的最优特征变量数 (optLV) 均为 10, 当变量数多于 optLV 时就会因数据冗余导致验证精度降低。

根据选取的主成分建立马铃薯叶片含水率诊断 CA-PLSR 和 RF-PLSR 模型, 得到建模集决定系数 (R_c^2) 分别为 0.975 5 和 0.983 2, RMSEC 分别为 2.81% 和 2.32%。然后将验证集中特征波长变量对应的光谱数据作为 X 变量代入到两个模型中得到验证集决定系数 (R_v^2) 分别为 0.933 2 和 0.947 1, RMSEV 分别为 2.31% 和 2.15%, 如表 2 所示。

表 2 两种算法得到的特征波长的叶片含水率 PLSR 模型
Table 2 Results of water content by PLSR model

模型	主成分数	建模集		验证集	
		R_c^2	RMSEC/%	R_v^2	RMSEV
CA-PLSR	10	0.975 5	2.81	0.933 2	2.31
RF-PLSR	10	0.983 2	2.32	0.947 1	2.15

比较 CA-PLSR 和 RF-PLSR 模型可知, RF-PLSR 模型建模集和验证集精度较高。分析其原因,从统计角度,由 CA 提取得到的 15 个波长在 1 406.8~1 445.8 nm 范围较为集中,分析它们之间的自相关性发现,自相关系数都高于 0.995 6,说明他们之间有严重的多重共线性;而 RF 算法提取的 15 个特征波长比较分散,跨度较广,包含有与含水率相关性低的波长如 869.8 和 979.8 nm ($|r| < 0.62$) 和相关系数较高的波长如 1 426.3, 1 465.2 和 1 523.2 nm ($|r| > 0.95$)。从光谱角度,CA 筛选得到的特征波长集中反映了 O—H 键在 1 410, 1 440 和 1 450 nm 附近的特征吸收波长;而 RF 算法筛选得到的特征波长包含的信息量较多,如 976.42 和 979.8 nm 体现了 970 nm 附近的水分子的吸收特征,

1 165.82 和 1 172.52 nm 位于作物水分敏感波长 1 150 nm 和 C—H 敏感波长 1 170 nm 之间, 以及 1 426.32 和 1 449.02 nm 体现了作物水分敏感波长 1 450 nm 附近的吸收特征。因此 RF 算法筛选得到的特征波长对水分的响应区间更为全面, 具有较高的代表性。

2.5 马铃薯叶片含水率分布图

2.5.1 叶片含水率分布可视化

利用 RF-PLSR 模型计算叶片各个像素点的叶片含水率, 随着不同烘干处理, 绘制叶片含水率分布如图 5 所示。图中黑色代表含水率为 0, 红色代表含水率为 100%, 从黑色到红色代表含水率的逐渐上升。观察可知, 新鲜叶片随着烘干时间的增加, 叶片中红色像素点逐渐减少, 绿色像素点逐渐增多; 继而在叶片边缘出现蓝色像素点, 逐渐增多并变成黑色, 说明叶片含水率逐渐下降。

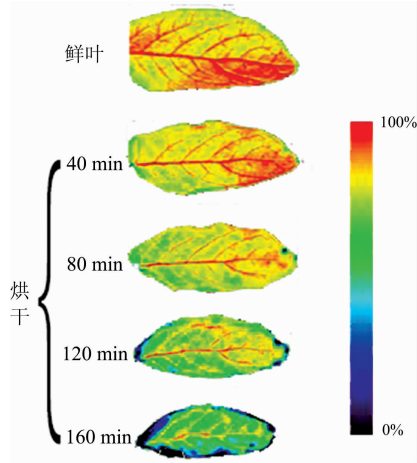


图 5 马铃薯叶片含水率分布可视化图

Fig. 5 Water content visualization distribution map of potato leaves

此外, 从图 5 中可以明显发现, 从鲜叶到烘干的 160 min 含水率下降的过程中, 叶肉部分的颜色变化最明显, 叶脉次之, 叶茎颜色的变化最缓慢; 叶片的面积逐渐减小, 且颜色总是从叶片边缘位置开始变化。如在烘干至 120 min 时, 叶片的叶茎和叶脉的颜色还保持红色, 而绝大多数叶肉部分已经变成绿色, 叶片的边缘部分已经变为蓝色。

以上现象说明当马铃薯处于干旱状态时, 其叶茎和叶脉仍保持较高的含水率, 进一步表征水分通过叶茎在叶脉间传输至整个叶片。在水分胁迫的过程中, 叶片会呈现卷曲枯萎等生理变化。这与农学家阐述的马铃薯受到水分胁迫, 其叶片会发生卷曲以减少水分流失的现象一致。

2.5.2 基于 HSV 模型对叶片含水率分布图分割

为了进一步探究马铃薯叶片在受到水分胁迫后含水率的分布情况, 用 HSV 彩色模型根据颜色的差异对图像进行分割。由于当含水率低于 70% 时, 叶片失去生理活性, 所以只计算鲜叶和烘干后叶片的含水率值大于 90%, 80% 和 70% 的像素点占整个叶片的面积比例, 分割结果如图 6 所示, 具体的面积比例数据见表 3。

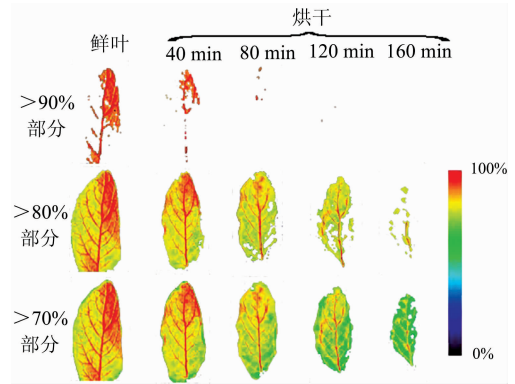


图 6 叶片中含水率大于 90%, 80% 和 70% 的部分可视化图

Fig. 6 Visualization map of water content more than 90%, 80% and 70% in potato leaves

表 3 含水率大于 70% 的部分占整个叶片比例的统计结果

Table 3 Statistical results of the percentage by the section with water content greater than 70%

	鲜叶	烘干 40 min	烘干 80 min	烘干 120 min	烘干 160 min
>90%	24.65	11.35	1.02	0.15	0
80%~90%	69.38	75.51	77.10	47.81	10.43
>80%	94.03	86.86	78.12	47.96	10.43
70%~80%	4.33	7.63	14.79	35.36	46.94
>70%	98.36	94.49	92.91	83.32	57.37

由图 6 可直观的看出, 叶片含水率随着烘干的次数增加而依次降低, 例如鲜叶中含水率大于 90% 的部分随着烘干时间的增加逐渐较少, 直至烘干 160 min 后为 0, 具体数据如表 3 所示。

鲜叶中含水率大于 90% 的部分占 24.65%, 鲜叶中的含水率大于 80% 的部分和大于 70% 的部分所占比例很相近, 比例分别为 94.03% 和 98.36%, 说明鲜叶中含水率大于 80% 的部分占绝大多数。从鲜叶到烘干 80 min, 含水率在 80%~90% 的像素点数所占比例一直增加, 但是从烘干 80 min 后开始减少, 直到烘干 160 min 时减少到 10.43%; 而含水率在 70%~80% 的像素点数所占比例。从鲜叶至烘干 160 min 一直处于增加的趋势。

3 结论

为了检测马铃薯叶片含水率, 获取了马铃薯叶片光谱成像数据, 利用 CA 和 RF 算法各筛选了 15 个特征波长, 建立了马铃薯叶片含水率 PLSR 诊断模型, 绘制了马铃薯叶片含水率可视化图像, 主要结论如下:

CA 筛选得到的 15 个特征波长集中在 1 406.8~1 445.8 nm, 自相关系数高于 0.995 6, 存在严重的多重共线性; RF 算法筛选得到的 15 个特征波长跨度较广, 包含的信息量较多。利用 CA 和 RF 算法各筛选出的 15 个特征波长, 建立马铃薯叶片含水率 PLSR 诊断模型。CA-PLSR 模型的建模集精度 (R_c^2) 为 0.975 5、RMSEC 为 2.81%, 验证集精度 (R_v^2) 为

0.933 2、RMSEV 为 2.31%；RF-PLSR 模型的建模集精度 (R_c^2) 为 0.983 2、RMSEV 为 2.32%，验证集精度 (R_v^2) 为 0.947 1、RMSEV 为 2.15%。RF-PLSR 模型的精度较高。

选取 RF-PLSR 模型绘制马铃薯叶片含水率灰度和伪彩色图像，实现含水率分布可视化；采用 HSV 彩色模型对

90%、80%和 70%不同梯度含水率图像进行分割，并提取面积百分比，结果表明采用高光谱成像可以实现马铃薯叶片的含水率分布可视化，为直观监测马铃薯生长状况以及叶片含水率分布提供理论根据。

References

- [1] HOU Xian-qing, LI Rong, HE Wen-shou, et al(侯贤清, 李 荣, 何文寿, 等). Journal of Zhejiang University(浙江大学学报), 2015, 41(5): 558.
- [2] XUE Jun-wu, REN Wen-jiang, YAN Chang-rong(薛俊武, 任稳江, 严昌荣). Chinese Journal of Agrometeorology(中国农业气象), 2014, 35(1): 74.
- [3] QIN Jun-hong, CHEN You-jun, ZHOU Chang-yan, et al(秦军红, 陈有君, 周长艳, 等). Chinese Journal of Eco-Agriculture(中国生态农业学报), 2013, 21(7): 824.
- [4] Cheng T, Rivard B, Sánchez-Azofeifa A. Remote Sensing of Environment, 2011, 115(2): 659.
- [5] LI Lei-lei, WANG Hai-xia, LIN Ping, et al(李蕾蕾, 王海霞, 林 萍, 等). Chinese Journal of Experimental Traditional Medical Formulae(中国实验方剂学杂志), 2013, 19(19): 104.
- [6] Zhao C, Li H, Li P, et al. IEEE Transactions on Geoscience & Remote Sensing, 2016, 55(1): 236.
- [7] Dejonge K C, Mefford B S, Chávez J L. International Journal of Remote Sensing, 2016, 37(10): 2294.
- [8] Gerhards M, Rock G, Schlerf M, et al. International Journal of Applied Earth Observation & Geoinformation, 2016, 53: 27.
- [9] CHEN Xiang, LI Min-zan, SUN Hong, et al(陈 香, 李民赞, 孙 红, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2017, 33(S1): 137.
- [10] Das B, Sahoo R N, Pargal S, et al. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2018, 192: 41.
- [11] SUN Hong, ZHENG Tao, LIU Ning, et al(孙 红, 郑 涛, 刘 宁, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2018, 34(1): 149.
- [12] Yang H, Zhang D Y, Dong Y Y, et al. International Journal of Agriculture & Biology, 2013, 15(4): 694.
- [13] Mirik M, Anskey R J, Steddom K, et al. Journal of Applied Remote Sensing, 2014, 8(1): 083661.
- [14] Li H D, Xu Q S, Liang Y Z. Analytica Chimica Acta, 2012, 740(31): 20.

Visualization of Water Content Distribution in Potato Leaves Based on Hyperspectral Image

SUN Hong¹, LIU Ning¹, WU Li¹, ZHENG Tao¹, LI Min-zan^{1*}, WU Jing-zhu²

1. Key Laboratory of Modern Precision Agriculture System Integration Research, China Agricultural University, Beijing 100083, China

2. Beijing Key Laboratory of Big Data Technology for Food Safety, Beijing Technology and Business University, Beijing 100048, China

Abstract In order to quickly detect the water content of potato leaves and explore the change of leaf water content under drought stress, the hyperspectral imaging technology was utilized to detect and visualize the moisture content of potato leaves in this paper. 71 leaves were collected and the water gradient of the leaves was controlled by a drying method. A total of 355 samples were obtained. The hyperspectral sorting instrument was used to collect potato leaves spectral and image data of 862.9~1 704.2 nm (256 wavelengths). The water content was measured by weighing method. According to a certain proportion, Sample Set Partitioning Based on Joint X-Y Distance (SPXY) algorithm was used to divide the sample into a model set and a validation set. For the calibration set, the feature wavelengths were extracted by using Coefficient Analysis (CA) and Random Frog (RF) algorithms respectively, and Partial Least Squares Regression (PLSR) models were established respectively. The calibration set and validation set determines coefficient R^2 and the RMSE (Root Mean Square Error) were used as the evaluation index. The gray image of the potato leaves water content was calculated using the results of the detection model. The visualization analysis of potato leaves water content was realized based on the pseudo color image transformation and segmentation. The average reflectance of each sample leaf was calculated by ENVI software, obtaining a total of 355 sample's spectral data. According to the

proportion of 2 : 1, the total samples were divided into calibration set (240 samples) and validation set (115 samples) by SPXY algorithm. Spectral characteristics of the collected data were analyzed. Two algorithms, CA and RF, were used to select 15 characteristic wavelengths, respectively. Based on the CA, the selected 15 wavelengths with the correlation coefficient higher than 0.96 were 1 406.82, 1 410.12, 1 403.62, 1 413.32, 1 416.62, 1 419.82, 1 400.32, 1 423.12, 1 426.32, 1 429.62, 1 432.82, and 1 441.12, 1 493.32, 1 442.52 and 1 445.8 nm. Based on the RF algorithm, the 15 feature wavelengths that the selected probability higher than 0.3 were 1 071.62, 1 041.12, 1 222.52, 1 465.22, 1 397.02, 1 449.02, 1 034.32, 1 523.22, 976.42, 1 172.52, 979.82, 1 165.82, 1 037.72, 1 426.32 and 869.8 nm. The PLSR model was established using the characteristic wavelengths filtered by the CA and RF algorithms, marked RF-PLSR and RF-PLSR model respectively. The water content of potato leaves was analyzed visually using the more precise model. First, the each pixel water content of the potato leaf image was calculated to obtain a gray image. Then, the gray image was pseudo-color transformed to draw a visual color image of the leaf water content. In order to reflect the change of potato leaves water content in the drying process, HSV model was used to segment the pseudo-color image of sample leaf. The segmentation image results, showing the proportion of leaf area in a certain water content interval, were obtained. The results showed that the 15 wavelengths selected by the CA algorithm were in the range of 1 400.3 to 1 450.0. The calibration accuracy of CA-PLSR was 0.975 5, and the RMSEC (Root Mean Square Error of Calibration) was 2.81%, and the validation accuracy was 0.933 2, and the RMSEV (Root Mean Square Error of Validation) was 2.31%. The range of characteristic wavelengths selected by the RF algorithm, with local "peak valley" characteristic, was wider than that of the CA algorithm. The calibration accuracy of RF-PLSR model was 0.983 2, and the RMSEC was 2.32%, and the validation accuracy was 0.947 1, and the RMSEV was 2.15%. The RF-PLS model is selected to calculate the water content of each pixel in the potato leaf images. According to the pseudo-color image, it could be seen intuitively that the water content gradually decreased with the drying time increasing. From the perspective of leaf tissue structure, with the strengthening of water stress, the leaves began to lose water from the edge and gradually spread to the middle, in which the water content of leaf stems and veins was higher than that of other parts. The pseudo-color image was segmented by HSV model based on the color difference of the image. The proportion of pixels with water content greater than 90%, 80%, and 70% in the leaf pseudo-color image to the entire leaf image. Using the hyperspectral imaging technology can realize the water content detection and distribution visualization of potato leaves, which provides a new theoretical basis for the potato growth analysis and potato leaf water content analysis.

Keywords Water content; Potato leaf; Hyperspectral imaging; Image segmentation; Visualization

(Received May 15, 2018; accepted Oct. 12, 2018)

* Corresponding author