

基于多光谱成像的番茄叶片叶绿素含量预测建模方法研究

姜伟杰, 孙明^{1*}

中国农业大学信息与电气工程学院, 北京 100083

摘要 传统的光谱分析技术预测植物的叶绿素含量的精度较低, 而基于 3CCD 的多光谱摄像机的叶绿素预测研究存在其摄像机本身成本昂贵和无法调整的波长通道数等局限性。文章提出了基于多光谱图像技术利用敏感波长(532, 610 和 700 nm)下番茄叶片的灰度值来预测其叶绿素含量的研究方法。利用多元线性回归分析、主成分分析和偏最小二乘回归分析等方法建立了预测模型, 取得了较好的预测效果, 其相关系数 R_1^2 与 R_2^2 均达到了 0.9 左右。表明该方法用于番茄叶绿素的预测是有效和可行的, 也为作物的长势检测仪器的发展奠定了基础。

关键词 番茄; 多光谱成像; 叶绿素含量; 偏最小二乘回归分析

中图分类号: S641.2 **文献标识码**: A **DOI**: 10.3964/j.issn.1000-0593(2011)03-0758-04

引言

我国目前的温室栽培面积已达 250 万公顷, 而番茄在温室栽培中占有很大的比重。当植株缺乏营养元素时, 会严重影响产量。因此, 改变传统的水氮管理, 在保证农产品和环境安全的前提下合理有效地变量施肥, 促进设施蔬菜优质生产和可持续发展是当前我国设施蔬菜发展过程中亟待解决的问题^[1]。

在利用传统的光谱技术对于植株长势的检测方面, 刘良云等^[2]、薛利红等^[3]对于监测冬小麦长势进行了研究。李映雪等^[4]测定了小麦叶片氮积累量与冠层反射光谱指数的定量关系。李民赞等^[5]建立了温室黄瓜叶片氮素、磷素含量的预测模型。赵国富等^[6]建立了基于光谱反射率的叶绿素含量预测模型。冯洁等利用多光谱成像技术识别黄瓜病害, 而利用 3CCD 多光谱相机, 张浩等^[7]对水稻氮含量进行了诊断。现有研究中, 基于传统光谱技术对植株的检测相关性较低, 而基于 3CCD 多光谱摄像机对植株的检测有摄像机本身成本高和维修费用昂贵, 同时也有无法调整的波长通道数(限制为三个)的局限性。多光谱图像技术利用敏感波长段, 能有效反映样品光谱和空间特征的信息, 但是建模方法的合理与否对最终预测能力有着十分明显的影响。

本文在多光谱成像技术的基础上, 探讨番茄叶片叶绿素含量预测建模方法, 为作物的长势检测仪器的开发奠定了基础。

础。

1 材料与方法

1.1 实验仪器

本实验采用美国 ASD 公司推出的 FieldSpec 手持式光谱辐射仪来测量叶片的反射率。用日本 KONICA MINOLTA 公司生产的 SPAD502 叶绿素仪测量叶片的 SPAD 值。基于所筛选敏感波长构建了如图 1 所示的多光谱采集系统, 以获取番茄叶片的灰度值。系统主要包括加拿大 Point Grey 公司的 FFMV-03MTM-CS 型号摄像机, Computar 公司的 M0814-MP 型号镜头, 日本 SIGMA KOKI 公司的 NDWH-15W 滤光片回转轮, 中心波长为 532, 610 和 700 nm, 半带

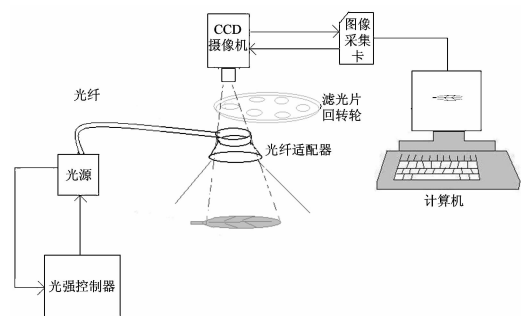


Fig. 1 Multi-spectral imaging system configuration

收稿日期: 2010-05-10, 修订日期: 2010-08-20

基金项目: 国家高技术研究发展计划(863 计划)项目(2006AA10Z201)资助

作者简介: 姜伟杰, 1984 年生, 中国农业大学信息与电气工程学院硕士研究生

e-mail: jwj20003@163.com

* 通讯联系人 e-mail: sunming@cau.edu.cn

宽为 10 nm 的滤光片,日本 Moritex 公司的 MHAB-150W 的卤素光源、MRG53-1000S 引导光纤、MD-04 扩散型照明适配器;试样载物台。

1.2 材料与敏感波长选取

本研究采集的样本为中国农业大学水利与土木工程学院温室中栽培的番茄的叶片。番茄品种为中杂 9 号,2009 年 11 月中旬定植,测量期为 2010 年 3 月下旬到 4 月中旬,采用基质栽培方式。为了对番茄实施氮胁迫,根据蛭石和草炭的不同比例配置了四种基质,温室采用滴灌方式。

选择光照适宜的情况下,利用 ASD 手持式光谱辐射仪测量了番茄叶片在 325~1 075 nm 的光谱反射率曲线如图 2 所示。由图可见,在 550 nm 波长附近是叶绿素的强反射峰区 750 nm 波长以上具有波状起伏和高反射率,吸收率极低。

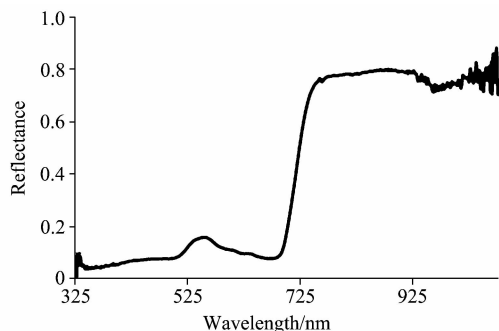


Fig. 2 Spectral reflectance of tomato leave

利用原始光谱进行相关性分析。番茄叶片的叶绿素含量与其光谱特性间的相关系数由式(1)所得, x_j 表示波长 j 处的光谱反射率, y 表示所测得的叶绿素含量。

$$r(x_j, y) = \text{cov}(x_j, y) / [\text{var}(x_j)\text{var}(y)]^{1/2} \quad (1)$$

1.3 获取叶片灰度值

为了建立敏感波长下番茄叶片的灰度值与其叶绿素含量的预测模型,基于 Visual C++ 平台编制了在多光谱系统下采集番茄叶片图像并计算每个叶片灰度值的软件。叶片的叶绿素含量则是以 SPAD502 叶绿素测量仪所测的值作为标准值。

1.4 建立预测模型的方法

拟采用多元线性回归分析(multiple linear regression analysis, MLRA)、主成分分析(principal component analysis, PCA)和偏最小二乘回归分析(partial least squares regression analysis, PLSRA)等常用的典型分析方法来建立预测模型。

多元线性回归分析在光谱学中最基本的,也是最常用的处理方法,而且它还是主成分分析和其他一些分析方法的基础^[8]。但也存在多光谱图像信息重叠严重,模型存在符号反常现象,没有从根本上消除多重共线等问题。为解决这些问题,一些研究者提出主成分回归分析和偏最小二乘回归分析等具有代表性的方法。

主成分分析的原理是使少数几个新变量成为原变量的线性组合,这些变量要尽可能多地表征原变量的数据结构特征而不丢失信息。新变量要求正交^[9]。

计算敏感波长下主成分的贡献率,前两个主成分的累计方差贡献率达 99.42%,因此可以用前两个主成分来建立多元线性回归模型。

偏最小二乘法回归分析的基本思路是一种逐步回归,即逐步提取光谱数据中的成分,逐步增加变量(信息不全时),逐步检验模型的显著性,一旦满足要求立即停止计算。

利用叶片灰度值与叶片的 SPAD 值进行了偏最小二乘回归分析。本研究中样本数量较多,因此采用交叉证实法中的外部证实法。计算模型的预测误差 PRESS 如表 1 所示。可见,当主成分数为 1 时,预测误差 PRESS 已经是最小值,这表明如果再加入主成分将会产生过度拟合而降低模型的预测能力。

Table 1 Result of PRESS

主成分数	PRESS	主成分数	PRESS
1	672.59	2	1 955.75

2 结果与讨论

2.1 敏感波长的选择

光谱反射率数据与叶片的叶绿素含量的相关系数如图 3 所示。可见,在 520~540 nm 之间为一个波谷,在 600~620 nm 及 690~710 nm 又有两个波谷。从上述三个敏感波段中选取三个敏感波长分别为 532, 610 和 700 nm,利用这三个敏感波长进行多元线性回归分析(MLR),所得的多元回归模型如式(2)

$$y = 26.8 + 1\ 271.4x_{532} - 1\ 326.9x_{610} - 299.5x_{700} \quad (2)$$

式(2)中, x_i 为波长 i 处的光谱反射率数据, y 为叶片的叶绿素含量的计算结果。该模型验证后相关系数 R^2 为 0.800,表明所选取的敏感波长与叶片的 SPAD 值有比较显著的相关性,为后文预测建模的准确性奠定了基础。

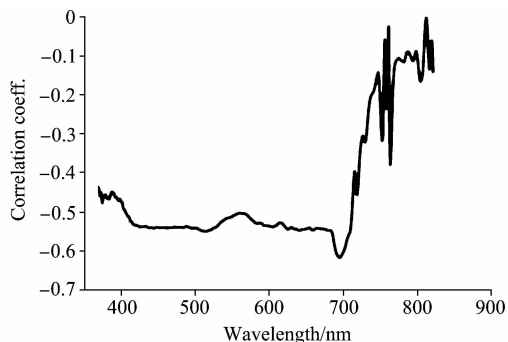


Fig. 3 Correlation coefficient between spectral reflectance and chlorophyll content

2.2 预测模型的建立与验证

本实验利用选取 532, 610 和 700 nm 三个敏感波长分别对样本数据进行多元线性回归分析、主成分分析和偏最小二乘回归分析,其结果分别如图 4—图 6 所示。式(3)—式(5)中, y 为叶片的叶绿素含量的预测结果, x_i 为敏感波长 i 下

的叶片的灰度值。

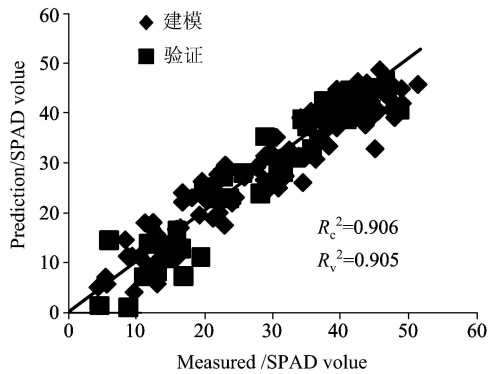


Fig. 4 Result of MLRA

图 4 为利用敏感波长下的叶片的灰度值建立的多元线性回归模型的分析结果。其回归模型如式(3)

$$y = 107.38 - 0.879x_{532} + 0.4445x_{610} - 0.368x_{700} \quad (3)$$

定标集样本的相关系数 R_c^2 达 0.906, 其验证集样本相关系数 R_v^2 达 0.905, 表明该模型具有很好的预测性。

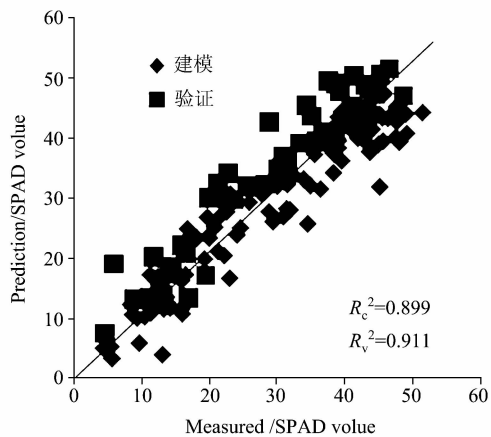


Fig. 5 Result of PCA

图 5 为利用前两个主成分进行还原, 然后经过多元线性回归分析, 得到主成分分析结果, 其模型如式(4)

$$y = 91.651 - 0.0766x_{532} + 0.0961x_{610} - 0.4372x_{700} \quad (4)$$

其验证结果也显示了较好的相关性。但主成分分析回归

一般适用于样本数量小于因变量个数的情况, 而本研究的目标是为了进行偏最小二乘回归分析选取主成分数, 计算 PRESS 值。

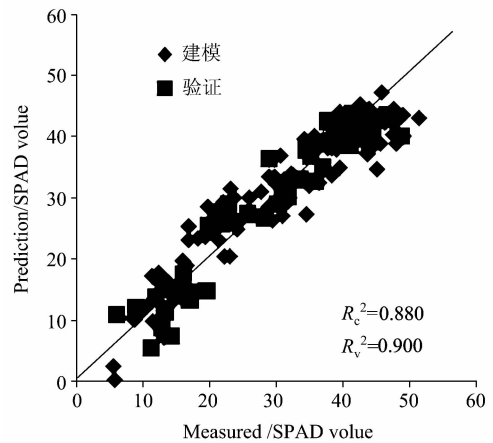


Fig. 6 Result of PLSRA

图 6 为利用 1 个主成分与番茄叶片的 SPAD 值做回归分析的结果, 其模型如式(5)

$$y = 97.994 - 0.2476x_{532} - 0.2458x_{610} - 0.2445x_{700} \quad (5)$$

可见, 该模型的 R_c^2 为 0.880, R_v^2 为 0.900, 表明该敏感波长下的叶片灰度值与其 SPAD 值有显著的线性关系, 且系数都为正, 已经不存在符号反常现象, 表明消除了多重共线性问题。

3 结论

本文分析了番茄叶片的反射光谱和叶绿素含量的相关关系, 筛选出的敏感波长分别为 532, 610 和 700 nm。基于这三个敏感波长所构建的多光谱成像系统, 采集了番茄叶片的图像, 探讨叶片灰度值与表征其叶绿素含量的 SPAD 值之间的预测模型。本研究采用了多元线性回归分析、主成分分析和偏最小二乘回归分析等三种常用方法进行建模和验证分析, 结果表明在这三个敏感波长下的番茄叶片的灰度值与其 SPAD 值存在显著的线性关系。相关系数 R_c^2 和 R_v^2 均达到 0.9 左右, 该方法可以作为预测番茄叶绿素的有效手段, 同时本研究也为后续的温室番茄生长状况的信息监测及其长势检测仪器的开发奠定了基础。

References

- [1] JIANG Hui-min, ZHANG Jian-feng, YANG Jun-cheng, et al(姜慧敏, 张建峰, 杨俊诚, 等). Plant Nutrition and Fertilizer Science(植物营养与肥料学报), 2010, 16(1): 158.
- [2] LIU Liang-yun, WANG Ji-hua, HUANG Wen-jiang, et al(刘良云, 王纪华, 黄文江, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2004, 20(1): 172.
- [3] XUE Li-hong, CAO Wei-xing, LUO Wei-hong, et al(薛利红, 曹卫星, 罗卫红, 等). Chinese Journal of Plant Ecology(植物生态学报), 2004, 28(2): 172.
- [4] LI Ying-xue, ZHU Yan, TIAN Yong-chao, et al(李映雪, 朱艳, 田永超, 等). Acta Agronomica Sinica(作物学报), 2006, 2: 203.
- [5] ZHANG Xi-jie, LI Min-zan, ZHANG Yan-e, et al(张喜杰, 李民赞, 张彦娥, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2004, 20(6): 11.

- [6] ZHAO Guo-fu, WANG Jian-xing(赵国富, 王建兴). Journal of Agricultural Mechanization Research(农机化研究), 2008, (11): 33.
- [7] ZHANG Hao, YAO Xu-guo, ZHANG Xiao-bin, et al(张浩, 姚旭国, 张小斌, 等). Chinese Journal of Rice Science(中国水稻科学), 2008, 22(5): 555.
- [8] LI Min-zan, HAN Dong-hai, WANG Xiu(李民赞, 韩东海, 王秀). Spectral Analysis Techniques and Applications(光谱分析技术及其应用). Beijing: Science Press(北京: 科学出版社), 2006. 135.
- [9] XU Lu, SHAO Xue-guang(许禄, 邵学广). Methods of Chemometrics(化学计量方法). Beijing: Science Press(北京: 科学出版社), 2004. 130.

Research on Predicting Modeling for Chlorophyll Contents of Greenhouse Tomato Leaves Based on Multi-Spectral Imaging

JIANG Wei-jie, SUN Ming*

College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China

Abstract Traditional spectrum analysis technology has low accuracy for forecasting chlorophyll content of plants. Research based on 3CCD camera has the limitations of high cost and the number of sensitive wavelengths not adjustable. The present paper develops a new approach to forecasting the chlorophyll content of tomato leaves by the image gray value of the selected sensitive wavelengths (532, 610 and 700 nm). Three common methods such as multi-linear regression, principal component analysis and partial least square regression were employed in forecast modeling, the good results were obtained, and both R_c^2 and R_v^2 reached about 0.9. The method has proven effective and feasible for prediction of chlorophyll contents of tomato leaves, which also lays the foundation for the development of testing instruments for the growing of crops.

Keywords Tomato; Multi-spectral imaging; Chlorophyll content; Partial least square regression(PLSR)

(Received May 10, 2010; accepted Aug. 20, 2010)

* Corresponding author