

基于近红外光谱小波变换的温室番茄叶绿素含量预测

丁永军, 李民赞*, 郑立华, 赵瑞娇, 李修华, 安登奎

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

摘要 为了提高基于近红外光谱的温室番茄叶绿素含量预测精度, 采用小波变换消除光谱中的随机噪声。但是在去噪的同时, 也会降低有效信息量。因此, 引入平滑指数(SI)和时移指数(TSI)对去噪效果进行量化, 以控制变换尺度, 获得最佳变换效果。实验表明 $TSI < 0.01$ 且 $SI > 0.1004$ 时, 在去噪的同时, 也能保留反映生化参量的特征峰, 从而实现自适应小波去噪。通过小波变换反射率与叶绿素含量的相关分析, 提取了反映叶绿素含量变化的特征波段, 使用偏最小二乘法建立了叶绿素含量预测模型, 结果表明使用 384, 405, 436, 554, 675 和 693 nm 处的吸光度建立的模型, 预测系数 R_p 达到 0.892 6, 验证系数 R_v 达到 0.829 7, 可以作为温室番茄营养状态快速诊断的技术基础。

关键词 温室番茄; 光谱分析; 小波变换; 叶绿素
中图分类号: O657.3; S641.2 文献标识码: A

DOI: 10.3964/j.issn.1000-0593(2011)11-2936-04

引言

叶绿素是光合作用中最重要的有机分子及合成叶片氮素的重要组成部分, 其浓度直接影响植物生长和营养水平^[1]。大量研究表明植物营养状况与其光谱特性密切相关, 而温室材料对光线强度及光谱构成会产生一定影响^[2, 3], 因此如何快速、准确地获得温室栽培作物的光谱特性、测定营养含量在长势监测、产量估计等方面有重要意义。在温室黄瓜的研究方面 Kavdir 建立了黄瓜色度、色调、干物质与光谱特征间的校正模型^[4]; 田有文等利用高光谱图像技术检测黄瓜霜霉病和白粉病^[5]; 刘飞提取 556, 581, 698 和 715 nm 四个特征波长建立黄瓜叶片 SPAD 值的校正模型^[6]; 杨昊谕研究了荧光参数与叶绿素含量的相关关系^[7]。而关于温室番茄的研究, 虽取得了一定的成果, 但尚处于探索阶段。如 Polder 研究了番茄果实光谱特性与成熟度的关系^[8]; XU 研究了利用番茄冠层光谱检测作物病害^[9]; 赵瑞娇研究了番茄叶片含氮量、含水量与反射光谱的关系。

近红外光谱主要由化合物含氢基团的倍频和泛频吸收组成, 吸收强度较弱, 谱峰重叠严重^[10]; 又由于受到温室材料吸收太阳光、传感器材料、电阻热噪声及随机误差等仪器因素的影响, 使得目标物光谱含有噪声、甚至出现基线漂移。为了提高光谱信息提取的准确度, 需要对光谱数据进行以去

噪为主的预处理。其中 Savitzky-Golay 卷积在窗口尺寸选择不当时, 容易造成信号失真; 微分法当存在频带较宽的强吸收时, 可能放大微小噪声; 傅里叶变换以信号的时不变和统计特性平稳为前提条件, 仅适用于对平稳信号去噪处理。而小波变换具有良好的时频局域化特性, 能对信号进行多分辨率分析, 可以有效地消除慢漂移、高频噪声及倍频干扰^[11, 12]。

本试验以温室番茄为研究对象, 研究了作物光谱特性与叶绿素含量间的变化规律。使用小波变换进行光谱去噪处理, 提取了温室番茄叶片叶绿素的特征波长, 并在建立光谱预测模型过程中, 对这些特征波段的贡献做了相应分析。

1 方法与材料

试验从 2009 年 10 月至 2010 年 4 月在中国农业大学水利与土木工程学院可控温室进行, 番茄栽培基质选用四种不同的草炭-蛭石比, 比例分别为 0%, 20%, 50% 和 80%, 分别对应无营养、少营养、适量、富营养四个施肥水平。先后实施 6 次番茄叶片光谱数据采集, 及叶绿素含量检测。

1.1 光谱数据的采集

采用 ASD FieldSpec HH (Analytical Spectral Devices Ins., USA) 光谱辐射仪实施光谱数据采集。考虑到天气对光谱构成的不利影响, 实验选择正午、少云天气进行。采集前

收稿日期: 2011-01-15, 修订日期: 2011-04-08

基金项目: 国家自然科学基金项目(30871453, 31071330)和高等学校博士学科点专项科研基金项目(20100008110030)资助

作者简介: 丁永军, 1978 年生, 中国农业大学博士研究生 e-mail: weakout@yahoo.com.cn

* 通讯联系人 e-mail: limz@cau.edu.cn

均使用标准板进行校正, 并且对每组样本测量三次, 其平均值作为光谱采集结果。

1.2 叶绿素含量的测定

采用分光光度法测定叶绿素含量, 将叶片去茎, 剪碎、混匀, 称取 0.4 g, 用 99% 丙酮和无水乙醇的 2:1 混合液 25 mL 浸泡 24 h, 利用分光光度计测量吸光度, 计算叶片叶绿素含量。

2 结果与分析

2.1 小波去噪评价及分解尺度确定

Daubechies 和 Symlets 小波不仅正交, 而且在消失矩一定的条件下有最小的支撑域, 即对信号局部特性的描述能力较强。同时, 消失矩不宜过大也不宜过小, 因此可选用 db3~db6 及 sym3~sym6 等小波函数。本试验采用 db4。在小波去噪过程中, 使用 Birge-Massart 和软阈值策略进行小波系数的阈值量化。图 1 为样本 17 原始光谱曲线及 1~7 层小波分解去噪光谱, 可以看出随着分解尺度的增加, 曲线变得越来越光滑, 消除了随机噪声的影响, 但同时反映生化参量的特征峰、谷也变得越来越不明显, 表明失去了有效信息。因此希望确定合理的分解尺度, 以获得理想的去噪光谱。

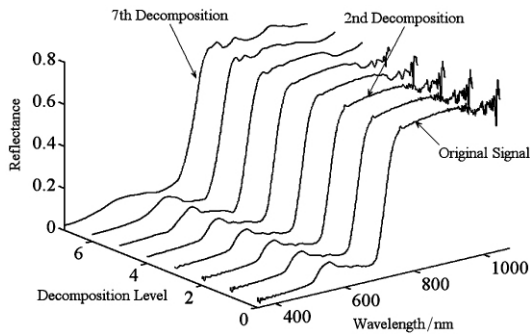


Fig 1 Spectra after denoising

目前广泛用于评价仿真信号去噪效果的指标有信噪比 (SNR)、均方根误差 (RMSE)、波形相似系数 (normalied correlation coefficient, NCC)。但事实上, 真实的漫反射光谱无法事先知道, 而且由图 2 可知, 随着分解尺度的增大, SNR、NCC 逐渐减小, RMSE 不断变大, 但是变化很不显著, 所以难以使用这些参数来确定合理的分解尺度。为此, 引入平滑指数 (SI) 和时移指数 (TSI), 对去噪效果进行量化, 以控制变换尺度, 获得最佳变换效果。

$$SI = \frac{\sum_{i=1}^{N-1} (f_{i+1} - f_i)}{\sum_{i=1}^{N-1} (s_{i+1} - s_i)} \quad (1)$$

$$TSI = |\lambda_{k-1} - \lambda_k| / \text{length}(\text{start}, \text{end}) \cdot \text{satisfy}(\lambda_{k-1}, \lambda_k \in (\text{start}, \text{end})) \quad (2)$$

其中, s 为原始光谱信号; f 为小波去噪后信号; λ_k 为第 k 层分解时极值点处波长; (start, end) 为考察的波长范围。考虑到在 500~700 nm 间, 绿色植被光谱必存在一个极大值点和一个极小值点, 试验中取 start=500, end=700。SI 反映了去噪曲线的平滑程度, 其值越小去噪效果越明显, 但同时也会

丢失更多的有效信息。试验表明当 $SI < 0.1004$ 时, 信息失真严重, 因此确定分解尺度上限为 7。TSI 反映了特征保持能力, 其值越大则信息损失越大。试验发现 $TSI < 0.01$ 且 $SI > 0.1004$ 时, 在消除随机噪声影响的同时, 也能保留反映生化参量的特征峰。由于各个样本的光谱曲线存在差异, 最优的分解尺度并非一成不变, 而是分布在 2~5 之间, 本试验采用的自适应小波去噪算法, 根据每条光谱曲线的特点, 自动选择相应的最优分解尺度。

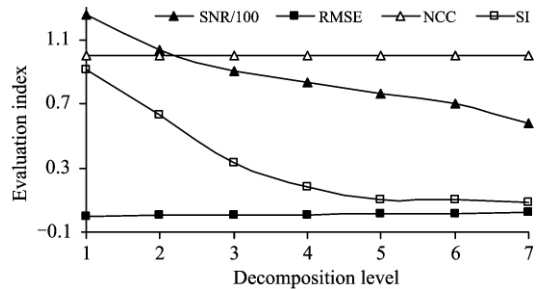


Fig 2 Changing tendency of evaluation index

2.2 番茄叶片反射光谱特征

通过对比苗期、出芽分化期、开花期、坐果期、绿熟期和成熟期的 6 次试验数据, 发现温室番茄在苗期和出芽分化期叶绿素含量不断增加, 从第 10 片真叶展开、第一花穗开花着果开始叶绿素含量呈下降趋势。受光合作用的影响, 蓝光波段和红光波段被叶绿素吸收, 在 400~500 和 600~700 nm 形成两个吸收谷, 而绿光吸收较少, 所以在 550 nm 附近形成一个反射峰。在 700 nm 左右反射率突然上升, 形成一个波状起伏的高反射平台。

图 3 为 4 个叶绿素含量不同的样本在 350~1 000 nm 波段的反射光谱曲线, 可以看出, 反射光谱在近红外及 350~560 nm 波段均有高频噪声, 而小波去噪后光谱曲线 (图 1 中曲线 2nd decomposition 为图 3 中曲线 III 的小波去噪结果) 变得比较平滑, 并且保留了 366 nm 对应的叶酸特征峰, 384, 405 和 437 nm 对应的叶绿素特征峰^[13]。

图 3 还反映了随着叶绿素含量的变化, 红边位置、绿峰位置、绿峰幅值、蓝光波段吸收谷和红光波段吸收谷等光谱特征的变化规律。当叶绿素含量越来越大, 作物的光合能力增强, 红边位置向长波长方向偏移, 而绿峰位置则向短波长方向偏移, 绿峰幅值变小, 蓝光波段吸收谷和红光波段吸收谷均有不同程度的降低。当作物光合能力减弱时红边位置、绿峰位置及绿峰幅值向相反方向变化。以上的结果表明可以用红边位置、反射峰及吸收谷等光谱特征表征作物的个体差异。光谱特征参量的提取有待进一步研究。

2.3 光谱反射率与叶绿素含量的相关性研究

对温室番茄结果期叶片反射率与其叶绿素含量进行了相关分析, 结果如图 4 所示。图 4(a) 是在整个波长范围的相关系数, 图 4(b) 是针对小波变换降噪效果明显的 350~450 nm 波长范围, 比较了降噪前后相关系数的变化趋势。从图 4 可以看出: ①在可见光波段 (400~730 nm) 光谱反射率与叶绿素含量呈负相关关系, 而进入近红外波段 (730~1 050 nm) 则与之相反, 而且在 690~710 nm 波段相关系数绝对值超过

0.8, 说明红边参数能够有效表征植株的营养信息。②在 350~450 nm 波段 Original signal 系列数据点杂乱无章, 说明高频噪声干扰了特征信息的提取。③去噪后信号的相关性曲线变得光滑, 而且在 366, 384, 405, 430 和 436 nm 均存在极值

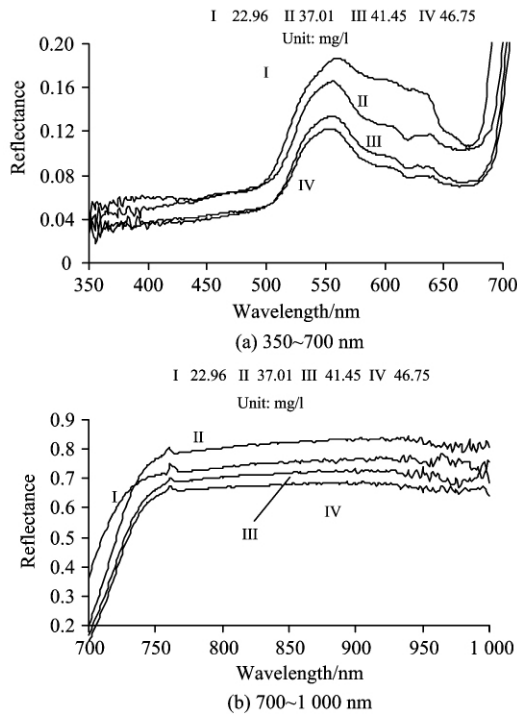


Fig 3 The reflectance of tomato leaves

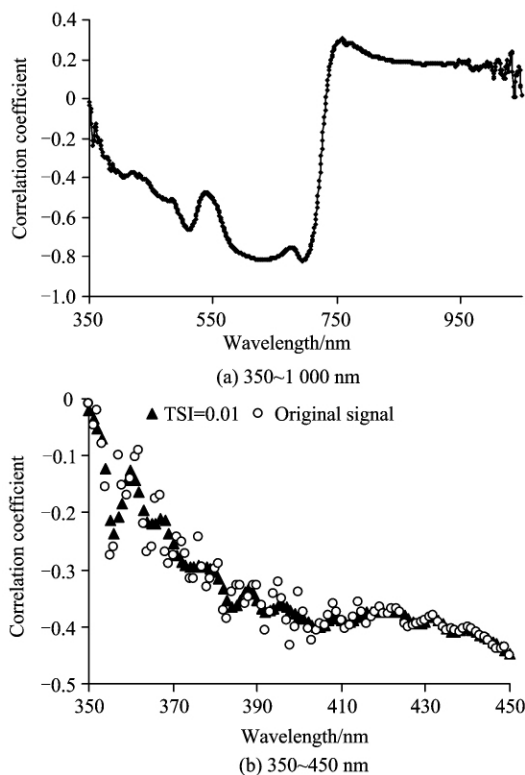


Fig 4 Correlation between spectral reflectance and chlorophyll content

点, 其中 366 nm 对应于叶酸的特征峰; 384 和 436 nm 对应于叶绿素 b 特征峰; 405 nm 对应于叶绿素 a 特征峰。

2.4 模型建立及评价

选取 40 个光谱样品作为校正集, 39 个样品作为预测集进行建模和模型验证。首先对提取的各个特征波段进行共线性诊断, 发现最大条件指数为 36.81, 最大方差膨胀因子为 133.78, 说明特征波段间存在一定的相关性。为了消除自变量间的共线性, 采用偏最小二乘回归 PLSR, 建立番茄叶片叶绿素含量的光谱预测模型。

为了明确各特征波段对模型及预测精度的贡献, 把它们划分为 3 个集合 $a = \{366\}$, $b = \{384, 405, 436\}$, $c = \{554, 675, 693\}$ 。其中, 366 nm 是叶酸的特征波长; 384, 405 和 436 nm 是叶绿素 a 和叶绿素 b 的特征波长, 554 nm 是叶绿素的强反射峰, 675 nm 是叶绿素敏感波段, 693 nm 与叶绿素含量具有最大相关性。分别使用 c , $c \cup b$, $c \cup b \cup a$ 处的吸光度作为输入变量建立 PLSR 模型并使用预测集进行模型预测能力检验。

Table 1 Evaluating parameters of calibration and prediction by PLS

| 模型 | 建模特征波集合 | RMSE | | R_c/R_v | |
|----|-------------------|---------|---------|-----------|---------|
| | | 原始信号 | 去噪信号 | 原始信号 | 去噪信号 |
| 校正 | c | 6.569 2 | 6.421 3 | 0.867 6 | 0.889 4 |
| | $c \cup b$ | 6.459 9 | 6.332 9 | 0.871 7 | 0.892 6 |
| | $c \cup b \cup a$ | 6.460 4 | 6.311 9 | 0.873 4 | 0.893 4 |
| 预测 | c | 6.724 6 | 6.525 7 | 0.800 4 | 0.812 4 |
| | $c \cup b$ | 6.532 2 | 6.451 9 | 0.817 1 | 0.829 7 |
| | $c \cup b \cup a$ | 6.695 3 | 6.507 8 | 0.805 2 | 0.815 1 |

从表 1 可知, 不同波长组合的校正模型的相关系数都较高。关于预测精度, 当使用 554, 675 和 693 nm 建模时, 预测模型的相关系数 R_v 达到 0.8124, 能够满足实际应用的需求。而加入 384, 405, 436 nm 后预测精度提高到 0.8297, 说明叶绿素 a 和 b 特征波对预测模型有一定贡献, 但模型相对复杂。另外, 当叶酸特征波长加入预测模型后, 预测效果反而下降, 其原因可能是叶酸无助于提高模型的预测精度; 或者是因为验证集样本来自番茄结果晚期采集的叶片光谱, 其叶酸特征已不显著或被淹没于高频噪声, 具体原因还有待进一步分析研究。从表 1 还可以看出, 小波去噪后模型的均方根误差 (RMSE) 得到了明显改善。

3 结论

本研究分析了温室番茄叶绿素含量和响应光谱的变化趋势和相关关系, 引入平滑指数 (SI) 和时移指数 (TSI) 对小波去噪效果进行量化, 从而实现自适应小波变换。对提取的特征波段, 使用偏最小二乘法建立了叶绿素含量预测模型。研究结果表明:

(1) 温室番茄在苗期和出芽分化期叶片叶绿素含量呈增加趋势, 从第 10 片真叶展开、第一花穗开花着果开始叶绿素

含量呈下降趋势。当作物光合能力增强时红边位置向长波长方向偏移,绿峰位置向短波长方向偏移,绿峰幅值减小,蓝光波段吸收谷和红光波段吸收谷均有所降低。当作物光合能力减弱时红边位置、绿峰位置及绿峰幅值向相反方向变化。

(2) 在 Birge-Massart 和软阈值策略条件下,当满足 $TSI < 0.01$ 且 $SI > 0.1004$ 时,在消除随机噪声的同时保留了 350 ~ 450 nm 范围内反映叶酸、叶绿素 a 及叶绿素 b 的特征峰,

从而获得理想的小波去噪光谱,实现了自适应小波去噪。

(3) 当使用 554, 675 和 693 nm 建模时,预测模型的相关系数 R_v 达到 0.812 4,能够满足实际应用的需求。而加入 384, 405, 436 nm 后预测精度提高到 0.829 7,说明叶绿素特征波长对预测模型有一定贡献。小波去噪后模型的均方根误差 (RMSE) 得到了明显改善。

References

- [1] Blackburn G A. Journal of Experimental Botany, 2007, 58(4): 855.
- [2] Impron I, Hemming S, Bot G. Biosystems Engineering, 2008, 99(4): 553.
- [3] Donald T K, David H C, Roman M M. Photochemistry and Photobiology, 2005, 81(1): 1047.
- [4] Kavdir I, Lu R, Ariana D. Postharvest Biology and Technology, 2007, 44(2): 165.
- [5] TIAN You-wen, LI Tian-lai, ZHANG Lin(田有文, 李天来, 张琳). Transactions of the Chinese Society of Agricultural Engineering (农业工程学报), 2010, 26(5): 202.
- [6] LIU Fei, WANG Li, HE Yong(刘飞, 王莉, 何勇). Journal of Infrared and Millimeter Waves(红外与毫米波学报), 2008, 28(4): 272.
- [7] YANG Hao-yu, YU Hai-ye, ZHANG Lei(杨昊谕, 于海业, 张蕾). Transactions of the Chinese Society for Agricultural Machinery(农业机械学报), 2009, 40(10): 169.
- [8] Polder G, van der Heijden G W. Postharvest Biology and Technology, 2004, 34(2): 117.
- [9] XU H R, Ying Y B. Biosystems Engineering, 2007, 96(4): 447.
- [10] LIU Zhe-gang, TANG Guo-xin(刘则刚, 唐国新). Journal of Tianjin University(天津大学学报), 2004, 37(6): 535.
- [11] LI Min-zan(李民赞). Spectral Analysis Technology and Application(光谱分析技术及其应用). Beijing: Science Press(北京: 科学出版社), 2006. 45.
- [12] Blackburn G A, Ferwerda J G. Remote Sensing of Environment, 2008, 112(4): 1614.
- [13] PAN Rui-chi, WANG Zhi-guan, DONG Yu-de(潘瑞炽, 汪止馆, 董愚得). Plant Physiology(植物生理学). Beijing: Higher Education Press(北京: 高等教育出版社), 1958. 145.

Prediction of Chlorophyll Content of Greenhouse Tomato Using Wavelet Transform Combined with NIR Spectra

DING Yong-jun, LI Min-zan*, ZHENG Li-hua, ZHAO Rui-jiao, LI Xiu-hua, AN Deng-kui

Key Laboratory of Modern Precision Agriculture System Integration Research, Ministry of Education, China Agricultural University, Beijing 100083, China

Abstract In quantitative analysis of spectral data, noises and background interference always degrade the accuracy of spectral feature extraction. The wavelet transform is multi-scale decomposition used to reduce the noise and improve the analysis precision. On the other hand, the wavelet transform denoising is often followed by destroying the efficiency information. The present research introduced two indexes to control the scale of decomposition, the smoothness index (SI) and the time shift index (TSI). When the parameters satisfied $TSI < 0.01$ and $SI > 0.1004$, the noise of spectral characteristic was reduced. In the meanwhile, the reflection peaks of biochemical components were reserved. Through analyzing the correlation between denoised spectrum and chlorophyll content, some spectral characteristics parameters reflecting the changing tendency of chlorophyll content were chosen. Finally, the partial least squares regression (PLSR) was used to develop the prediction model of the chlorophyll content of tomato leaf. The result showed that the prediction model, which used the values of absorbance at 366, 405, 436, 554, 675 and 693 nm as input variables, had higher predictive ability (calibration coefficient was 0.892 6, and validation coefficient was 0.829 7) and better potential to diagnose tomato growth in greenhouse.

Keywords Greenhouse tomato; Spectral analysis; Wavelet transform; Chlorophyll

* Corresponding author

(Received Jan. 15, 2011; accepted Apr. 8, 2011)