

## 基于叶绿素荧光光谱分析的黄瓜霜霉病害预测模型

隋媛媛<sup>1</sup>, 于海业<sup>1\*</sup>, 张 蕾<sup>1</sup>, 曲剑巍<sup>1</sup>, 武海巍<sup>1,2</sup>, 罗 瀚<sup>1</sup>

1. 吉林大学生物与农业工程学院, 仿生工程教育部重点实验室, 吉林 长春 130022
2. 北华大学电气信息工程学院, 吉林 吉林市 132021

**摘 要** 为了实现对黄瓜病害的快速无损准确预测, 基于激光诱导叶绿素荧光光谱分析技术, 建立了温室黄瓜霜霉病害的预测模型。通过测定健康叶片、病菌接种 3 d 叶片和接种 6 d 叶片的光谱曲线, 采用一阶导数光谱预处理方法, 结合主成分分析数据降维方法对三组光谱数据进行特征信息提取后, 建立主成分得分散点图, 依据累积贡献率选取 10 个主成分代替导数光谱曲线, 再利用最小二乘支持向量机技术进行分类和预测。通过对三组光谱数据 105 个样本的训练, 对 44 个样本进行分类预测, 并对比了四种核函数的支持向量机的分类能力, 结果表明, 径向基核函数对黄瓜霜霉病害的分类预测能力达到了 97.73%, 具有很好的分类和鉴别效果。

**关键词** 荧光光谱; 主成分分析; 支持向量机; 黄瓜霜霉病

**中图分类号:** S123 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2011)11-2987-04

### 引 言

黄瓜疾病的综合管理是保证高品质和生产安全的关键, 因此对黄瓜病害的预测是对有害生物综合防治的先决条件<sup>[1]</sup>, 同时也是减少对其经济投入和使用农药过程中环境污染的有效途径。叶绿素荧光被誉为植物体的无损检测探针, 它是通过采集植物体内的微弱荧光, 达到判断植物体内光合速率、叶绿素含量等生理状况的目的<sup>[2]</sup>。近几年来叶绿素荧光在植物生理监测中的应用得到了迅速的发展<sup>[3-5]</sup>。使用叶绿素荧光光谱技术实现对黄瓜病害预测的快速、有效、正确的判断, 对指导农业生产具有重要的意义。

主成分分析的精髓在于较少损失信息的前提下将多个指标转化为少数几个综合指标<sup>[6]</sup>, 是通过线性变换<sup>[7]</sup>以尽可能少的正交向量表征数据信息特征的一种多元统计分析的方法<sup>[8]</sup>。支持向量机(support vector machine, SVM)是采用结构风险最小化(SRM)原理<sup>[9]</sup>, 在高维特征空间使用线性函数假设空间的学习系统, 由最优化理论的学习算法训练, 实现了由统计学习理论<sup>[10]</sup>导出的学习偏置<sup>[11,12]</sup>。

本研究的目的是采用反射式激光诱导叶绿素荧光光谱技术结合主成分分析和最小二乘支持向量机的数据挖掘方法建立黄瓜霜霉病害的预测模型。

### 1 实 验

#### 1.1 仪器

本实验使用荷兰 Avantes 公司生产的 AvaSpec-2048-USB2 型光纤光谱仪采集系统, 其光谱检测范围为 360~1100 nm, VA 光栅 300 线·mm<sup>-1</sup>, 分辨率 2.1 nm, 采样积分时间 1.1 ms。激光发射装置激发光强度为 7.5 mW。设置光谱仪采集探头与激光发射探头成 45°角, 贴近活体叶片表面进行光谱采集。

#### 1.2 样品光谱采集

实验于 2010 年 6—9 月于吉林大学南岭校区玻璃温室内进行。黄瓜实验品种为“长春密刺王”, 霜霉病菌采自吉林农业大学实验田。

诱导接种方法: 采摘带有霜霉病菌的叶片用无菌水冲洗, 双层纱布过滤, 配成孢子悬浮液。接种前先将薄层湿棉平铺于叶片上, 保湿 10 min; 接种后将自封袋套于叶片上, 保湿 24 h, 定期通风降温, 喷水加湿。在温室棚顶加盖遮阳网避免阳光直射。

在温室内随机选取生长健康的壮苗期叶片 50 片进行实验。实验分为接种前的健康叶片、病菌接种 3 和 6 d 的叶片共三组数据进行对比研究, 并对其进行叶绿素荧光光谱采

收稿日期: 2011-02-11, 修订日期: 2011-06-20

基金项目: 国家高技术研究发展计划(863 计划)项目(2007AA10Z203)资助

作者简介: 隋媛媛, 女, 1985 年生, 吉林大学生物与农业工程学院博士研究生 e-mail: sui yuan0115@126.com

\* 通讯联系人 e-mail: haiye@jlu.edu.cn

集。但在第 6 天进行光谱数据采集时人为损坏 1 片叶片，故样本数据为 149 个。

### 1.3 荧光光谱预处理

光谱预处理目的是去除光谱的随机噪声、基线漂移、样本不均匀、光散射等影响。经研究发现，一阶导数处理比较适合于荧光光谱分析，但是噪声较高的仪器不宜采用这种方法，因此在对荧光光谱预处理之前采用光谱仪自带软件 AVANTES Software 7.2 进行基本处理，以消除仪器因素对光谱数据的影响。

为消除光谱首末端波动的影响，选取荧光光谱 500~800 nm 波段进行分析。分别将健康叶片、病菌接种 3 和 6 d 叶片的荧光光谱进行一阶导数处理，得到一阶导数荧光光谱曲线如图 1 所示。对比健康叶片与接种后叶片的导数光谱发现，其强度表现出了一定的变化规律，但接种后叶片强度变化规律的效果并不明显。

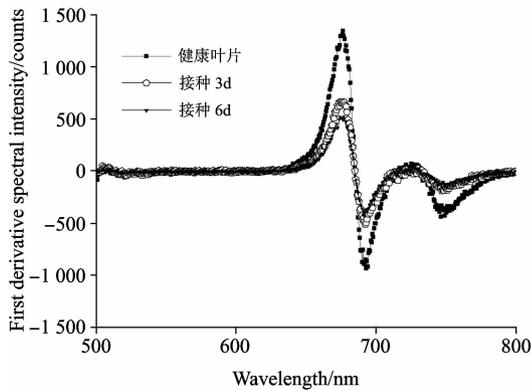


Fig. 1 First order derivative fluorescence spectra of healthy and inoculated leaves

## 2 试验结果与分析

### 2.1 荧光光谱的主成分分析

采用 Matlab 软件对上述光谱数据提取主成分，对提取的主成分累积贡献率进行分析，在得到的主成分中，前 10 个主成分累计贡献率达到 98.796%，因此采用前 10 个主成分代替导数光谱。主成分累计贡献率见表 1。

Table 1 Reliabilities of principal components

主成分	累计贡献率/%	主成分	累计贡献率/%
PC1	97.418	PC6	98.610
PC2	98.172	PC7	98.661
PC3	98.371	PC8	98.709
PC4	98.499	PC9	98.753
PC5	98.556	PC10	98.796

分别将光谱提取的第一主成分、第二主成分、第三主成分得分作为 X, Y 和 Z 轴的值，如图 2 所示。

图 2 中定性地说明了黄瓜健康叶片与病菌接种后 3 和 6 d 后的叶片光谱的变化情况，通过主成分分析降维处理后能够进行定性判别，但在分类边界处的区分不很明显。因此，

引入支持向量机数据挖掘方法对降维后的导数光谱进行辨别预测。

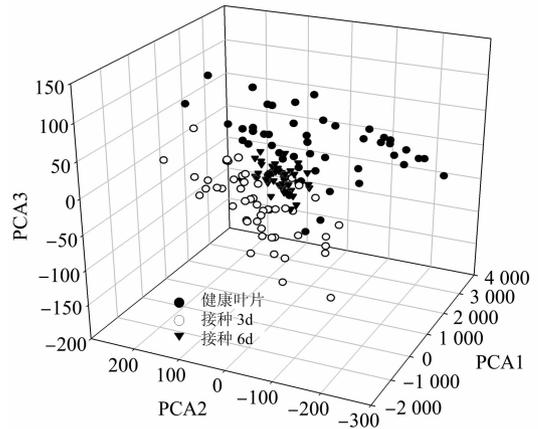


Fig. 2 Principal component scores scatter plot (PC1×PC2×PC3) of leaves

### 2.2 支持向量机的分析结果

支持向量机的输入数据为 149 片叶片的前 10 个主成分得分值，将每组数据随机分成训练集 (35 个) 和预测集 (15 个)，得到训练集样本 105 个，预测集样本 44 个。

将样本中的数据映射到高维空间中，最终建立一个线性可分的超平面来实现支持向量机的分类功能是通过核函数完成的。采用不同的核函数可以得到不同的算法和计算精度。在计算中对比了四种经典的核函数分别是线性核函数 (LINEAR)、多项式核函数 (POLY)、径向基核函数 (RBF)、多层感知高斯核函数 (SIGMOID)。四种核函数的表达式如下：

$$\text{LINEAR: } K(x, y) = xy$$

$$\text{POLY: } K(x, y) = [(xy) + 1]^d$$

$$\text{RBF: } K(x, y) = \exp\{-\gamma \|x - y\|^2\}$$

$$\text{SIGMOID: } K(x, y) = \tanh[s(xy) + u]$$

在黄瓜霜霉病的预测中采用一类对余类的分类方法，算法如下：

设样本数据的训练集为  $T = \{(x_1, y_1), \dots, (x_n, y_n)\} \in (X \times Y)^n$ ，其中  $x_i \in X = R^n$ ， $y_i \in Y = \{1, \dots, l\}$ ， $i = 1, \dots, n$ 。当针对第  $z$  ( $z = 1, \dots, l$ ) 类进行训练时，把其余  $l-1$  的类看作负类。设样本数据的分类函数为

$$f^z(x) = \text{sgn}\left\{\sum_{i=1}^n y_i \alpha_i^z K(x_i, x) + b^z\right\}$$

其中  $x$  为预测集中的样本。

Table 2 Parameters of SVM model with different kernel function

核函数	C	d	γ	s	u	分类准确率/%	预测准确率/%
LINEAR	2	/	/	/	/	100	61.36
POLY	2	3	/	/	/	100	95.45
RBF	2	/	0.25	/	/	100	97.73
SIGMOID	2	/	/	0.02	0	100	72.72

采用 Matlab 7.0.1 软件完成支持向量机的四种算法程序,通过网格搜索法<sup>[13]</sup>确定核参数,对比分析发现,使用 POLY 和 RBF 进行预测的效果优于 LINEAR 和 SIDMOID,分类准确率能够达到 95%以上,其中使用 RBF 的预测效果更优,达到 97.73%。结果见表 2。

研究表明,采用主成分分析结合支持向量机数据挖掘方法对黄瓜霜霉病害的预测能力效果较好,使用 RBF 霜霉病预测能力达到 97.73%,优于另外三种核函数的处理结果。因此选择径向基函数结合主成分分析方法,可以建立温室黄瓜霜霉病害预测模型。

### 3 结 论

采用激光诱导叶绿素荧光光谱分析技术对黄瓜霜霉病害预测模型进行了深入的研究。通过采用一阶导数光谱预处理方法,结合主成分分析和径向基核函数最小二乘支持向量机的数据挖掘方法建立了黄瓜霜霉病的预测模型,是一种无损、快速、定量的检测方法,经实验证明预测能力可以达到 97.73%,该模型的预测效果较好,可以用于指导生产,为植物病虫害的无损检测提供了新的方法。

### References

- [1] Ming L, Chunjiang Z, Daoliang L, et al. *Computer and Computing Technologies in Agriculture*, 2008, 2(259): 1375.
- [2] HE Yan-hong, GUO Liang-sheng, TIAN You-liang(何炎红, 郭连生, 田有亮). *Acta Botanica Boreali-Occidentalia Sinica(西北植物学报)*, 2005, 25(11): 2226.
- [3] Chatzistathis T A, Papadakis I E, Therios I N, et al. *Journal of Plant Nutrition*, 2010, 34(1): 98.
- [4] Kim J, Moon Y R, Lee M H, et al. *International Journal of Low Radiation*, 2010, 7(4): 1477.
- [5] Floerl S, Druebert C, Aroud H I, et al. *Journal of Plant Pathology*, 2010, 92(3): 693.
- [6] LI Jing-ping, XIE Bang-chang(李静萍, 谢邦昌). *Multivariate Analysis: Methods and Applications(多元统计分析: 方法与应用)*. Beijing: China Renmin University Press(北京: 中国人民大学出版社), 2008.
- [7] YUE Tian-li, PENG Bang-zhu, YUAN Ya-hong, et al(岳田利, 彭邦柱, 袁亚宏, 等). *Transactions of the Chinese Society of Agricultural Engineering(农业工程学报)*, 2007, 23(6): 223.
- [8] ZENG Jie, SUN Jun-liang, LI Guang-lei, et al(曾洁, 孙俊良, 李光磊). *Journal of Shenyang Agricultural University(沈阳农业大学学报)*, 2009, 40(1): 53.
- [9] Cortes C, Vapnik V. *Machine Learning*, 1995, 20(3): 273.
- [10] Vapnik V N. *The Nature of Statistical Learning Theory*. New York: Springer-Verlag, 1995.
- [11] Cristianini N, Shawe-Taylor J. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods(支持向量机导论)*. Beijing: Publishing House of Electronics Industry(北京: 电子工业出版社), 2004.
- [12] DENG Nai-yang, TIAN Ying-jie(邓乃扬, 田英杰). *New Method of Data Mining Support Vector Machine(数据挖掘中的新方法——支持向量机)*. Beijing: Science Press(北京: 科学出版社), 2004.
- [13] WANG Xing-ling, LI Zhan-bin(王兴玲, 李占斌). *Periodical of Ocean University of China(中国海洋大学)*, 2005, 35(5): 859.

## Cucumber Downy Mildew Prediction Model Based on Analysis of Chlorophyll Fluorescence Spectrum

SUI Yuan-yuan<sup>1</sup>, YU Hai-ye<sup>1\*</sup>, ZHANG Lei<sup>1</sup>, QU Jian-wei<sup>1</sup>, WU Hai-wei<sup>1,2</sup>, LUO Han<sup>1</sup>

1. Key Laboratory of Bionic Engineering, Ministry of Education, School of Biological and Agricultural Engineering, Jilin University, Changchun 130022, China
2. Electrical and Information Engineering College of Beihua University, Jilin 132021, China

**Abstract** In order to achieve quick and nondestructive prediction of cucumber disease, a prediction model of greenhouse cucumber downy mildew has been established and it is based on analysis technology of laser-induced chlorophyll fluorescence spectrum. By assaying the spectrum curve of healthy leaves, leaves inoculated with bacteria for three days and six days and after feature information extraction of those three groups of spectrum data using first-order derivative spectrum preprocessing with principal components and data reduction, principal components score scatter diagram has been built, and according to accumulation contribution rate, ten principal components have been selected to replace derivative spectrum curve, and then classification and prediction has been done by support vector machine. According to the training of 105 samples from the three groups, classification and prediction of 44 samples and comparing the classification capacities of four kernel function support vector machines, the consequence is that RBF has high quality in classification and identification and the accuracy rate in classification and prediction of cu-

cumber downy mildew reaches 97.73%.

**Keywords** Fluorescence spectrum; Principal components analysis; Support vector machine; Cucumber downy mildew

\* Corresponding author

(Received Feb. 11, 2011; accepted Jun. 20, 2011)

## 第十七届全国分子光谱学学术会议 (第一轮通知)

由中国化学会和中国光学学会联合主办, 韶关学院和韶关市化学化工学会联合承办的“第 17 届全国分子光谱学术报告会”, 将于 2012 年 10 月 19—22 日在广东韶关召开。

本次会议是我国分子光谱界的又一次聚会和高水平、高信息容量的学术交流。本着继往开来、与时俱进的精神, 本次会议将全力展示我国在分子光谱及相关领域所取得的最新研究进展及成果, 增进广大分子光谱科学工作者和支持分子光谱事业的人们之间的交流与合作, 促进我国分子光谱事业的发展。届时会议将邀请国内外知名专家学者就分子光谱有关学术领域的前沿热点问题作大会报告, 同时会议还将组织各类专题讨论和学术交流。竭诚欢迎全国高等院校、科研机构 and 产业部门从事分子光谱研究和应用开发的同事和朋友们来广东韶关参加会议, 交流最新研究成果, 推进分子光谱基础研究和应用技术在国民经济和高新技术开发中的广泛应用和技术转化。我们真诚期待着国内外同行在 2012 年 10 月相聚在美丽的韶关。

### 征文范围

分子光谱理论研究, 红外光谱、拉曼光谱、荧光光谱、磷光光谱、紫外一可见吸收光谱、激光光谱、光谱成像等各类光谱技术在物理、化学、生物、材料科学、表面/界面科学、医药、环境、工业过程、催化学、地学、农林及其他领域的基础理论与应用研究的最新科研成果。同时也欢迎相关的光谱技术(如质谱、核磁共振等)的最新研究成果。

### 论文要求

1. 论文内容必须是未在期刊杂志上发表过或其他全国或国际会议宣读过。
2. 提交论文扩展摘要一份, 纸张大小用 A4 纸版式(用 Office word 软件排版, 页边距为 2 cm, 单倍行距)。
3. 扩展摘要按以下顺序排版: 文题(三号黑体居中); 作者(四号仿宋居中); 单位(小四号宋体居中, 含所在省市、邮政编码、电子邮址(如有)); 论文的创新性, 研究意义与结果(五号宋体); 关键词和主要参考文献(自版芯左起, 五号宋体)。文稿中可穿插主要论据的图、表和照片, 图题、图注和表题、表注一律用英文表述。摘要的字数, 包括图、表, 参考文献, 总共不能超过 4 500 字。
4. 具体投稿要求可参看模板及《光谱学与光谱分析》征稿简则。稿件一经录用, 将由《光谱学与光谱分析》以增刊形式发表会议论文摘要集。
5. 论文截稿日期: 2012 年 4 月 30 日, 尽量通过电子版, 一般不接受手稿。

### 论文提交方式

欢迎大家通过网站提交论文, 请您注册登陆中国光谱网(<http://www.sinospectroscopy.org.cn>) 点击论文提交上传您的论文。论文将采用网上评审的方式, 你可在网上浏览对您论文的评审结果和修改意见。请您通过电子邮件方式提交论文。请发往: [gphy2012@126.com](mailto:gphy2012@126.com); 若以信件方式投稿, 请以挂号邮寄, 信封上请注明“第十七届全国分子光谱学学术会议征文”字样; 收稿地址: 邮编 512005 广东省韶关市浈江区大塘路九公里韶关学院化学系; 黄冬兰收。

### 报告形式

为充分利用会议时间, 提高学术交流的效率, 本次会议仍采用“口头报告”和“墙报展示”两种方法进行学术交流。无论是口头报告还是墙报展示, 均属大会同等学术交流, 无水平高低之分。对内容好, 制作精良的报展颁发奖励。为尊重个人意见和便于组委会的安排, 请投稿人注明选择自己的稿件为“口头报告”或“墙报”的字样。在安排“口头报告”和“墙报”时, 将充分考虑作者的意见。同时会议还将邀请国内外知名专家学者就分子光谱有关学术领域的前沿热点问题作大会报告和主题报告。主要报告形式有:

1. 大会邀请报告: 主要邀请国内外知名专家学者报告光谱分析的前沿技术在各个领域的最新进展。
2. 论坛主题报告: 本次会议将选择光谱技术的热点应用领域, 开设多个专题论坛, 邀请在该领域的知名专家作论坛主题报告。
3. 论坛邀请报告: 邀请专家学者围绕论坛主题进行学术交流。
4. 墙报展示: 作为本次会议的主要交流和展示形式之一, 会议将统一安排墙报的讲解时间, 希望作者能按时到位。

(下转 3030 页)