# 农业信息与电气技术·

# 小波分析与支持向量机结合的冬小麦白粉病遥感监测

黄林生<sup>1</sup>,刘文静<sup>1,2</sup>,黄文江<sup>1,2</sup>\*,赵晋陵<sup>1</sup>,宋富冉<sup>1,2</sup>

(1. 安徽大学安徽省农业生态大数据工程实验室, 合肥 230601;

2. 中国科学院遥感与数字地球研究所,数字地球重点实验室,北京 100094)

摘 要:为利用遥感影像数据在区域尺度上实现快速、准确地监测小麦白粉病的发生、发展情况,该研究基于环境与灾 害监测预报小卫星(HJ-1A/1B)数据对地表温度(land surface temperature, LST)进行反演、提取4个波段反射率数据 并构建7个植被指数。耦合 *K*-mean 和 Relief 算法对小麦白粉病遥感特征进行筛选。通过支持向量机(support vector machine, SVM)与小波特征(Gabor)结合 SVM(GaborSVM)的方法分别建立河北省晋州市小麦白粉病发生监测模型,并对2种模型的监测精度进行对比。结果表明:归一化植被指数(normalized difference vegetation index, NDVI)、比值植 被指数(simple ratio index, SR)和地表温度3种特征参量可较好地表征小麦白粉病的发生情况,GaborSVM的总体精度 达到 86.7%,优于 SVM 的 80%。因此,小波分析与支持向量机结合的方法可用于基于卫星遥感影像的大面积病害监测,对提高病害监测精度具有重要应用价值。

关键词:遥感;支持向量机;病害;白粉病;小波特征

doi: 10.11975/j.issn.1002-6819.2017.14.026

中图分类号: TP79 文献标志码: A

文章编号: 1002-6819(2017)-14-0188-08

黄林生,刘文静,黄文江,赵晋陵,宋富冉.小波分析与支持向量机结合的冬小麦白粉病遥感监测[J].农业工程学报, 2017,33(14):188-195. doi: 10.11975/j.issn.1002-6819.2017.14.026 http://www.tcsae.org Huang Linsheng, Liu Wenjing, Huang Wenjiang, Zhao Jinling, Song Furan. Remote sensing monitoring of winter wheat powdery mildew based on wavelet analysis and support vector machine[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2017, 33(14): 188-195. (in Chinese with English abstract)

doi: 10.11975/j.issn.1002-6819.2017.14.026 http://www.tcsae.org

# 0 引 言

小麦白粉病严重影响小麦产量,据统计,白粉病危 害一般可使小麦减产 5%~10%,严重区域可达 20%以上 <sup>[1]</sup>。准确获取病害发生状况和其空间分布对于病害防治是 十分必要的,传统的病虫害监测主要依靠植保人员的田 间调查和田间取样等方式,尽管这些传统方法的真实性 和可靠性较高,但耗时、费力,难以适应目前大范围的 病虫害实时监测和预报的需求,因此有必要建立基于遥 感影像的监测模型<sup>[2]</sup>。

目前,一些学者利用遥感宽波段数据对病虫害进行 了一系列研究。马慧琴等<sup>[1]</sup>利用 Landsat 8 遥感影像数据 结合气象数据采用相关向量机的模型实现小麦白粉病的 区域尺度监测,证明了遥感影像数据单独使用无法得到 满足需求的试验结果,需与气象数据结合分析。Huang 等<sup>[3-4]</sup>发现小麦白粉病由于光谱响应总体上较平滑,不同

收稿日期: 2017-01-03 修订日期: 2017-07-07

Email: linsheng0808@163.com

※通信作者:黄文江,博士,研究员,博士生导师。研究方向为植被定量遥 感。北京 中国科学院遥感与数字地球研究所,100094。

Email: huangwj@radi.ac.cn

于某些仅在较窄波段范围内发生响应的病害,因此采用 宽波段的光谱特征识别白粉病是可行的。Luo 等<sup>[5]</sup>利用 LST (land surface temperature)等数据构建二维特征空间 对小麦蚜虫进行了预测,发现 LST 对蚜虫是否发生起决 定性作用,是蚜虫发生发展的一个关键性因子。张竞成 等[6-8]研究了小麦白粉病主要的敏感波段及敏感植被指 数。以上分析表明宽波段植被指数对于病害识别是有可用 性的,但宽波段植被指数所含信息较为笼统,单独使用无 法得到满意的试验结果。因此尝试对宽波段植被指数进行 进一步细化研究,从而得到更多有用信息,提高分类精度。 小波分析是多种分析的结合算法,能够从多尺度、多方向 上分解数据<sup>[9]</sup>,实现对数据的细化分析。鲁军景等<sup>[10]</sup>利用 航空遥感高光谱数据采用小波分析的方法识别小麦白粉 病的敏感波段,并得到了较高的识别精度,高光谱数据包 含大量细节信息,证明小麦白粉病在细节信息中有较好体 现,但航空遥感高光谱数据在大范围尺度的监测上存在 一定局限性。以下分析采用不同的源数据,在进行识别分 类时利用小波分析来提取细节信息,均达到了提高识别精 度的目的。Chen 等<sup>[11]</sup>在对地震信号谱分解中先寻找小波 变换的最优旋转因子,再进行处理,降低了算法的运算复 杂度,并得到较好的试验结果。印勇等[12]在对人脸表情识 别时采用了 PCA (principal component analysis) 算法对小 波特征进行降维处理,提高了算法的运算效率及精度。 牛连强等<sup>[13]</sup>在表情识别试验中,利用 LBP (local binary patterns)算法结合小波变换的方法,大幅度降低了特征

基金项目: 安徽省自然科学基金(1608085MF139); 安徽省科技重大专项 (16030701091); 中国科学院国际合作局对外合作重点项目 (131211KYSB20150034);国家自然科学基金国际合作项目(61661136004); 国家重点研发计划项目(2016YFD0300702)

作者简介:黄林生,博士,副教授,研究方向为农业遥感技术与应用。合肥 安徽大学安徽省农业生态大数据工程实验室,230601。

的维数,并提高了特征提取的准确性,得到了100%的识别率。上述3组研究表明小波分析算法在图像识别领域有较高的应用价值,但目前尚未出现小波分析应用于宽波段植被指数提取病害信息方面的研究,因此尝试将小波变换(Gabor)应用于宽波段植被指数,并对得到的小波特征进行筛选,突出对病害敏感的因子,以提高病害识别精度。

支持向量机在机器学习领域通常用来模式识别、分类 及回归分析,此算法结构稳定,使用方便,相较于人工神 经网络等其他算法具有能够获得全局最优解的优点<sup>[14-15]</sup>。 Wang 等<sup>[16]</sup>利用支持向量机模型对小麦条锈病进行分类 和识别,获得了 97%的识别精度。袁莹等<sup>[17]</sup>利用 SVM (support vector machine)模型对玉米颗粒霉变程度进行 判别,准确率达到 91%。张录达等<sup>[18]</sup>利用 SVM 对小麦蛋 白质含量进行了预测,并得到了较好的分析结果。以上 SVM 模型在小范围病害数据分类识别中具有较高的应用 价值,但未尝试使用小波变换结合 SVM 在大区域尺度的 小麦白粉病识别中进行应用,本研究将二者结合进行试 验分析,实现区域尺度的小麦白粉病发生分布的监测。

基于以上分析,本文以河北晋州市为研究对象,尝 试仅利用环境星遥感数据经过小波变换及特征筛选后, 结合 SVM 算法建立的监测模型(GaborSVM),最终实时 准确获取大面积小麦白粉病发生的空间分布特征,为白 粉病防治提供依据,提高农药的使用效率,从而有助于 粮食产量的提高。

#### 1 材料与方法

小麦白粉病监测的具体操作流程如图1所示。



Fig.1 Flow chart of wheat powdery mildew monitoring

#### 1.1 研究区概况

本研究区位于河北省石家庄市的晋州(114°58′E~ 115°12′E,37°48′N~38°10′N)(如图 2),该区域属于黄 河流域白粉病易发气候区,晋州市地处滹沱河和滏阳河 冲积扇的交汇处,地势平缓开旷且全境皆平原,由于此 地地势平坦,气候单一,同时种植结构较单一,适合利 用遥感卫星影像来展开小麦白粉病监测。

#### 1.2 数据获取

研究所用数据主要包括遥感数据和小麦白粉病实地 调查数据。遥感数据为环境与灾害监测预报小卫星星座 A、B(HJ-1A/1B 星)数据,根据研究区天气状况,选择 质量较好、时间最接近地面调查的影像数据,即 2014 年 5月 29日的 CCD 光学数据和 IRS 热红外数据,其详细的 载荷信息如表 1 所示。小麦白粉病实地调查数据于 2014 年 5月 27、28日(小麦灌浆期)于晋州调查获得,在选 择调查点区域时,观察到晋州市内的白粉病发生情况类 似,故希望选择一处具有代表性的地块做调查,周家庄 是全国仅存的人民公社,此处处于种植区域的中部最大 且集中种植小麦的地块,满足卫星影像处理的要求,同 时此处的小麦白粉病发生情况较均匀,发生了不同等级 的小麦白粉病,相较于其他区域更具有代表性。野外调 查共得 40 个有效数据,调查方法参见文献<sup>[19]</sup>。





#### 表 1 HJ-1A/1B 卫星主要载荷参数信息 Table 1 Specifications of multispectral remote sensors

onboard HJ-1A and 1B satellites

HJ	J-1A/1B CCD	1&CCD2	HJ-1B IRS				
波段 Band	光谱范围 分辨率 Spectral Spatial range/µm resolution/m		波段 Band	光谱范围 Spectral range/µm	分辨率 Spatial resolution/m		
1	0.43~0.52	30	5	0.75~1.10	150		
2	0.52~0.60	30	6	1.55~1.75	150		
3	0.63~0.69	30	7	3.50~3.90	150		
4	0.76~0.90	30	8	10.50~12.50	300		

#### 1.3 数据处理

获取的环境星 CCD 影像、IRS 影像需要经过辐射定标、大气校正和影像裁剪,并结合 Landsat8 影像进行几何校正等预处理。环境星影像辐射定标公式如下:

$$L = DN/\beta + L_0 \tag{1}$$

式中L为辐射亮度, $\beta$ 为绝对定标系数增益, $L_0$ 为偏移量, DN为遥感影像像元亮度值。

辐射定标系数来源于中国资源卫星应用中心,之后 完成相应传感器的波普响应函数待用,并采用 ENVI5.1 软件中 FLAASH 模块完成影像的大气校正,最后对校正 后图像进行裁剪获取研究区影像。

预处理完成后,根据研究区的作物类型利用 NDVI、

数字高程模型(digital elevation model, DEM)、近红外反 射率数据,并结合 ENVI5.1 监督分类中的最大似然分类 提取冬小麦的种植区域<sup>[1]</sup>。

利用环境星影像数据选取对小麦白粉病较敏感的 7 个宽波段植被指数<sup>[5]</sup>(表 2)和红、绿、蓝和近红外 4 个 波段反射率数据,以及采用单通道算法反演得到的 LST 数据<sup>[20-21]</sup>作为白粉病监测模型的初选特征。

表 2 宽波段植被指数 Table 2 Wide-band vegetation index

植被指数 Vegetation index	计算公式 Formula	参考文献 References
增强型植被指数	$2.5(\rho_{NRI} - \rho_R)/(\rho_{NRI} + 6\rho_R - 7.5\rho_B + 1)$	[22]
Enhanced vegetation		
index (EVI)		
改进的简单比值指数	$(\rho_{_{NPI}}/\rho_{_{P}}-1)/(\sqrt{\rho_{_{NPI}}/\rho_{_{P}}}+1)$	[23]
Modified simple		
ratio index (MSR)		
归一化植被指数	$(\rho_{\scriptscriptstyle NRI} - \rho_{\scriptscriptstyle R})/(\rho_{\scriptscriptstyle NRI} + \rho_{\scriptscriptstyle R})$	[24]
Normalized difference		
vegetation index (NDVI)		
优化土壤调节植被指数	$(\rho_{NRI} - \rho_R)/(\rho_{NRI} + \rho_R + 0.16)$	[25]
Optimized soil adjusted		
vegetation index (OSAVI)		
重归一化植被指数	$(\rho_{\rm MPI} - \rho_{\rm p})/(\sqrt{\rho_{\rm MPI}/\rho_{\rm p}})$	[26]
Re-normalized difference		
vegetation index (RDVI)		
土壤调节植被指数	$1.5(\rho_{NRI} - \rho_R)/(\rho_{NRI} + \rho_R + 0.5)$	[27]
Soil adjusted vegetation		
index (SAVI)		
比值植被指数	$\rho_{\scriptscriptstyle NRI}/\rho_{\scriptscriptstyle R}$	[28]
Simple Ratio Index (SR)		

注:  $\rho_{NRI}$ 为近红外波段的反射率;  $\rho_R$ 为红波段的反射率;  $\rho_B$ 为蓝波段的反射率。 Note:  $\rho_{NR}$  means near-infrared reflectance;  $\rho_R$  means red band reflectance;  $\rho_B$  means blue band reflectance.

#### 1.4 建模特征选择

在模型构建时选择最能反映病害发生发展状况的特 征变量可以有效提高模型的准确度。本文在建模特征选 择时主要包括2个部分内容。首先是针对提取的12个植 被指数特征数据进行特征优选,得出最佳特征组合,并 用于模型构建。其次,为了进一步凸显健康小麦与病害 小麦的区别,将筛选出的特征变量进行小波变换,得到 一组反映指数特征某种局部细节的小波特征集,同时对 其进行二次筛选得到一组最佳小波特征作为监测模型的 输入变量,具体的实现过程如下。

1.4.1 宽波段植被指数筛选

Relief 算法是一种特征权重算法, 通过计算特征与类 别间的相关性赋予特征不同的权重,但是 Relief 算法不能 识别类别间的冗余以及特征间的相互负作用, 聚类分析 可以得到不同特征对样本的聚类精度,可根据聚类分析 的精度来提取最高聚类精度的特征集合,因此研究采用 Relief 算法<sup>[29-30]</sup>结合 K-mean 算法<sup>[31-32]</sup>的方法对植被指数 特征进行筛选,得出最佳的特征组合,其中聚类分析通 过 MATLAB 中的 K-means 函数实现。具体的操作过程为 根据 Relief 算法将特征数据按权重降序排序,将排序后的 特征依次组合进行聚类分析,具体做法为:①选择 NDVI 进行聚类分析; ②选择 NDVI、SR 进行聚类分析; ③选择 NDVI、SR、SAVI 进行聚类分析,按照特征权重排序依次 类推。将取得的聚类精度最大的特征集合用于建模分析。 表 3 列举出了各个特征的 Relief 特征权重、K-mean 独个特 征聚类精度及特征组合聚类精度,由第三行数据可知在组 合到 SAVI 时精度开始下降,在 LST 时有上升。选择 NDVI、SR 和 LST 再次进行聚类分析,精度为 0.7451,选 择 NDVI 和 LST 与 SR 和 LST 分别进行聚类分析,精度均 为 0.6078。最终选择 NDVI、SR 和 LST 用于模型的构建。 1.4.2 Gabor 小波变换及小波特征筛选

尽管宽波段植被指数可以表达出小麦白粉病的相关 特征,但由于宽波段植被指数本身的特点,运用的波段 范围较大,存有误差,因此有必要对其细节信息进行进 一步的滤波提取。小波变换可以实现对数据的滤波、去 嗓等优化。张竞成<sup>[6]</sup>在对小麦白粉病叶片光谱特征的研究 中发现小波特征与病情严重度有较强的相关性,因此小 波特征在区域尺度白粉病识别上可以尝试利用。小波变 换具有多分辨率特性,采用多通道滤波,每个通道都可 以得到数据的某种局部细节特征,突出数据的敏感信息, 故而在某种程度上优化了对光谱信息的利用。将筛选得 到的 NDVI、SR 和 LST 进行小波变换,并从变换得到的 小波特征集中筛选出对有无小麦白粉病区别最大的小波

表 3 各个特征的 Relief 特征权重、K-mean 聚类精度、特征组合精度

Table 3         Feature weights by Relief, clustering precision by K-mean and precision by combined features												
特征筛选算法 Feature filtering algorithms	NDVI	SR	SAVI	$\rho_R$	$ ho_{NRI}$	EVI	RDVI	LST	$ ho_G$	OSAVI	MSR	$ ho_{B}$
Relief 特征权重 Feature weights by Relief	0.0812	0.0722	0.0659	0.0650	0.0619	0.0475	0.0415	0.0360	0.0038	0.0035	-0.0062	-1.2800
K-mean 聚类精度 Clustering precision by K-mean	0.7059	0.7059	0.7059	0.6078	0.5098	0.6667	0.7059	0.6078	0.2941	0.2941	0.2941	0.2353
特征组合精度 Precision by combined features	0.7059	0.7059	0.6078	0.4902	0.2941	0.2353	0.2353	0.7451	0.2353	0.2353	0.2353	0.2157

注: pG 为绿波段的反射率。

Note:  $\rho_G$  means green band reflectance.

Gabor 小波能同时对时间和频率进行局部分析,这使 得对平稳信号的分析更加容易,对 Gabor 利用傅立叶展 开,就是利用时间和频率同时定义一个时间函数的方法, 而 Gabor 小波变换就是求解 Gabor 的展开系数<sup>[33]</sup>。本研 究采用高斯核函数作为母小波构建小波核函数将植被指 数特征与小波核函数进行卷积运算,卷积后的幅值作为 建模特征信息,在农业应用方面通常采用高斯函数作为 母小波函数构建小波核函数<sup>[34]</sup>:

 $\begin{cases} g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right) \\ h(x, y) = g(x, y) \exp(2\pi j W x) \\ H(u, v) = \exp\left(-\frac{1}{2}\left(\frac{(u-W)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}\right)\right) \end{cases}$ (2)

式中g(x, y)为高斯调制函数, $\sigma_x$ 和 $\sigma_y$ 为其在(x, y)2个

坐标轴上的标准差, h(x, y) 为小波函数, W 为复正弦函数在横轴上的频率, H(u,v) 为小波函数的傅立叶变换形式, u 为频域中的自变量频率, v 为对应频率信号的幅度值,  $\sigma_u$  和 $\sigma_v$  为其在(u,v) 2 个坐标轴上的标准差。

$$S(x, y) = \sqrt{(h_R * I)^2 + (h_I * I)^2}$$
(3)

式中(h\*I)表示滤波器h与数据I的卷积, $h_R$ 表示滤波器h的实部, $h_I$ 表示滤波器h的虚部,S(x,y)即为经过Gabor滤波器得到的特征。以h(x,y)为母小波,对其进行尺度和旋转变换,可以得到一组自相似的滤波器:

$$h_{mn}(x, y) = \alpha^{-m} h(x', y')$$
 (4)

式中  $x' = \alpha^{-m}(x\cos\theta + y\sin\theta), y' = \alpha^{-m}(-x\cos\theta + y\sin\theta),$  $\alpha > 1, \theta = n\pi/K, \alpha^{-m}$ 为尺度因子, t为尺度的数目 m=0,1,...,t-1, K为方向的数目  $n=0,1,\cdots,K-1$ 。

本试验中 *t*=5, *K*=8,  $\alpha = \sqrt{2}$ 。

因此,研究共构建 5 个尺度 8 个方向共 40 个小波核 函数,使得小波变换后的数据量扩大为原来的 40 倍。为 了找到最佳分类的小波特征,同时去除特征维数过多对 模型运算效率的影响<sup>[35]</sup>,对小波特征进行了进一步筛选。 研究采用独立样本 *T* 检验的方式对小波特征进行筛选处 理<sup>[10,33,36]</sup>。

经过独立样本 *T* 检验选择对有无病害具有显著性差异(*P*<0.001)且*T*统计的相伴概率最小的小波特征构建监测模型,并得到其对应的小波核函数。

m、n及对应的尺度因子和旋转角度如表4所示。

Tuble	I urum	etens of optim	indi wavelet i	anetion
植被指数 Vegetation index	尺度 Scale	方向 Direction	尺度因子 Scale factor	旋转角度 Rotation angle
LST	4	2	1/4	π/4
NDVI	1	0	$1/\sqrt{2}$	0
SR	4	1	1/4	$\pi/8$

表 4 最佳小波函数的参数 Table 4 Parameters of optimal wavelet function

### 1.5 模型构建

支持向量机是基于统计学习理论的一种机器学习方法,它的核心思想是结构风险最小化,通过核函数把输入线性不可分的数据映射到高维空间,构造超平面,使得不同样本之间的类间隔最大,类内间隔最小,它具有结构简单、适应性强、全局最优等特点,能较好的解决高维特征、非线性、过学习与不确定性等问题,广泛应用于遥感影像分类中<sup>[37]</sup>。该模型的判别函数为

$$f(x) = \operatorname{sgn}\left(\sum_{x_i \in S_V}^n a_i y_i k(x_i, x) + b\right)$$
(5)

式中 $a_i$ 为 Lagrange 乘子,  $S_V$ 为支持向量,  $x_i$ 、 $y_i$ 为 2 类中的支持向量, b为阈值, 其中 $k(x_i, x)$ 为满足 Mercer 定理的正定核函数<sup>[37-40]</sup>。

本研究基于 SVM 算法共建立 3 个模型, 第一个模型 利用全部的 12 个植被指数结合 SVM 算法建立对照试验, 第二个模型对 12 个植被指数进行指数筛选,用得到 3 个 植被指数结合 SVM 算法建立模型,第三个模型在第二个 模型的基础上对 3 个植被指数进行 Gabor 小波变换,再 对得到的小波特征进行筛选,用得到的小波特征结合 SVM 建立模型 (GaborSVM),利用 Gabor 小波变换在敏 感特征提取方面的优势和支持向量机在小样本分类中的 优势以提高监测模型的精度和效率。

本试验共获得 40 个发生白粉病的调查点数据,调查 时将其分为0(无病害)、1(轻度)、2(中度)、3(重度)、 4(特重)共5个等级。由于无病害与轻度较难区分,故 将以上5类样本分为健康(无病害、轻度)和病害(中 度、重度和特重)2类。其中25个为训练样本构建模型, 15个为测试样本用于模型的验证。

# 2 结果与分析

#### 2.1 研究区小麦白粉病监测

利用 2014 年 5 月 29 日的影像数据,以单个像元为 单位,利用 Relief 算法与 K-mean 聚类相结合的方式筛选 出3个特征指数 NDVI、LST 和 SR,分别利用 SVM 模型 及 GaborSVM 模型得到 2014 年 5 月 29 日的小麦白粉病 发生分布情况如图 3 所示,图 3a 为将全部 12 个植被指数 运用 SVM 模型预测结果,图 3b 为经过特征筛选后采用 NDVI、LST、SR 的 SVM 监测结果,图 3c 为分别对 NDVI、 LST、SR 小波变换后,利用经过独立样本 T 检验得到的 最优小波特征进行 SVM 的监测结果。从图 3 中可以看出 3种监测模型的白粉病发生情况的总体空间分布相似,东 部发病情况较西部严重。而发病面积占总种植面积的百 分比图 3a 为 49%、图 3b 为 45%和图 3c 为 38%。图 3a 与其余 2 幅图相比, 白粉病的发生情况较为零散。图 3b 与图 3c 大体相同,呈现整块的区域分布,仅在某些细小 部位存在区别:在图 3b 中处于健康区域内的小块病害区 域,在图 3c 中部分被分为健康区域。而小麦白粉病是由 布氏白粉菌引起的,具有繁殖快,传播面广的特点<sup>[41]</sup>。 因此,在小麦灌浆期白粉病零散发生的概率较低。由此 可以间接得出 SVM 结合特征筛选模型与 GaborSVM 结合 特征筛选模型的可信度高于SVM未经特征筛选模型并且 GaborSVM 结合特征筛选模型相较于 SVM 结合特征筛选 模型有了一定的改善。

从整体上看 3 个模型的空间分布相类似,为了进一步观察监测结果,可以从局部图中查看。图 4 为小麦白粉病在调查点区域的监测结果图。1 号、2 号区域实际调查结果均为健康区域。图 4a 在利用 SVM 建模在 1 号区域中得到的监测结果为染病,在 2 号区域中的到的监测结果为染病病区域大于健康区域;图 4b 在特征筛选后利用 SVM 建模在 1 号区域中得到的监测结果为健康,在 2 号区域中的到的监测结果为染病;图 4c 利用 GaborSVM 建模在 1 号区域与 2 号区域中得到的监测结果均为健康。GaborSVM 模型的监测结果与实际调查结果最相似,在局部图中 4a 中小麦白粉病分布零散,而图 4b、c 中小麦白粉病分布较集中,这与总体分析结果相一致。因此GaborSVM 模型可适用于小麦白粉病监测。







图 4 小麦白粉病监测结果空间局部分布图 Fig.4 Monitoring spatial local map of wheat powdery mildew

### 2.2 模型的评估与验证

采用独立的样本数据对模型进行验证能够更好地体现实际模型的精度<sup>[12]</sup>。本试验采用 2014 年 5 月 27、28 日的白粉病的地面调查点数据对模型监测结果进行评价。表 5 中列出了 SVM 模型与 GaborSVM 结合 Relief 算法与*K*-mean 聚类的用户精度、总体精度和 Kappa 系数。从结果中可以看出 3 组试验都获得了较好的试验结果。

从总体精度上看,SVM 模型的总体精度低于 GaborSVM 模型,说明 Gabor 小波特征相较于原始植被指 数特征对病害的识别率较高,GaborSVM 模型的 Kappa 系数也达到 0.583,高于 2 个 SVM 模型的 0.286 和 0.444。 并且在 2 个 SVM 模型中,通过特征筛选的 SVM 模型精 度高于未进行特征筛选的 SVM 模型,可以推测是由于去 除了冗余特征及负相关特征所致。从用户精度来看,3 组 模型中病害的用户精度分别为 50%、83.3%、91.7%,表 明 3 个模型对病害的识别精度在不断提高,GaborSVM 的 用户精度达到 91.7%,表明此模型能较为准确的识别病害 样本。以上结果表明,小波特征能提高监测模型对健康 与病害的区分精度,且特征筛选有助于提高模型精度。

表 5 总体验证结果

	Table 5	Overa	all veri	ficati	on resul	ts				
		精确度指标 Accuracy indices								
模型 Models		健康 Healthy	病害 Disease	总 和 Sum	用户 精度 User Accuracy /%	总体 精度 Overall accuracy /%	Kappa 系数 Kappa coefficient			
	健康 Healthy	3	0	3	100					
支持问量机 SVM	病害 Disease	6	6	12	50	60	0.286			
	总和 Sum	9	6	15						
特征筛选 结合 SVM	健康 Healthy	2	1	3	66.7					
SVM with	病害 Disease	2	10	12	83.3	80	0.444			
Relief and <i>K</i> -mean	总和 Sum	4	11	15						
特征筛选结	健康 Healthy	2	1	3	66.7					
GaborSVM	病害 Disease	1	11	12	91.7	86.7	0.583			
with Relief and K-mean	总和 Sum	3	12	15						

### 3 结 论

1)特征筛选结合 SVM 模型比 SVM 模型总体精度从 60%提高到 80%,说明去除特征间的冗余性确实可以提高 模型精度与效率; GaborSVM 模型的总体精度达到 86.7%, 此模型与特征筛选结合 SVM 模型相比总体精度从 80%提 高到 86.7%,说明小波变换应用于植被指数可以提高小麦 白粉病的监测精度。

2)GaborSVM 模型的 Kappa 系数为 0.583 为 SVM 模型 0.286 的 1 倍多,GaborSVM 模型的 Kappa 系数与特征 筛选结合 SVM 模型相比也有明显提高,说明 GaborSVM 模型的一致性有了明显的提高,在实际应用中更为可靠。此实验结果满足农业部门对小麦白粉病发生及时监测的 需求,可指导其及时制定相关的应对治理措施,减少产量损失,提高经济效益。

训练样本的质量和数量会影响模型的精度,本研究 开展时由于成本限制,地域限制,采样量不足,仅获取 了河北石家庄的小部分区域的数据,因此模型的通用性 有待进一步验证。初步选取的植被特征指数不完整,可 选择更多敏感指数进行试验,以提高试验精度。选用的 因子均来自遥感数据,未选用其他影响病害发生发展的 气象因子、农田管理信息等,这一方面降低了数据获取 难度,但另一方面可能会影响模型的精度,在今后的研 究中可以融合更多数据,对比分析试验,找到既易获取、 易处理又能提高精度的数据来构建监测模型。研究在建 立监测模型时使用了 GaborSVM 算法,一方面由于 GaborSVM 模型的测试地点处于平原区域,在复杂地形的 情况下仍需进一步测试,模型的可移植性仍需验证。另 一方面 Gabor 特征在建模前需对其进行进一步筛选,而 筛选算法可以有进一步改进测试,以得到更多有效特征, 提高识别精度和数据处理效率。

#### [参考文献]

- 马慧琴,黄文江,景元书. 遥感与气象数据结合预测小麦 灌浆期白粉病[J]. 农业工程学报, 2016, 32(9): 165-172. Ma Huiqin, Huang Wenjiang, Jing Yuanshu. Wheat powdery mildew forecasting in filling stage based on remote sensing and meteorological data[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2016, 32(9): 165-172. (in Chinese with English abstract)
- [2] Sankaran S, Mishra A, Ehsani R, et al. A review of advanced techniques for detecting plant diseases[J]. Computers and Electronics in Agriculture, 2010, 72(1): 1–13.
- [3] Huang W, Lamb D W, Niu Z, et al. Identification of yellow rust in wheat using in-situ spectral reflectance measurements and airborne hyperspectral imaging[J]. Precision Agriculture, 2007, 8(4/5): 187–197.
- [4] Devadas R, Lamb D W, Simpfendorfer S, et al. Evaluating ten spectral vegetation indices for identifying rust infection in individual wheat leaves[J]. Precision Agriculture, 2009, 10(6): 459-470.
- [5] Luo Juhua, Zhao Chunjiang, Huang Wenjiang, et al. Discriminating wheat aphid damage degree using 2-dimensional

feature space derived from landsat 5 TM[J]. Sensor Letters, 2012, 10(1/2): 608-614.

- [6] 张竞成. 多源遥感数据小麦病害信息提取方法研究[D]. 杭州:浙江大学, 2012.
  Zhang Jingcheng. Methods for Information Extraction of Wheat Disease Based on Multi-source Remote Sensing Data[D]. Hangzhou: Zhejiang University, 2012. (in Chinese with English abstract)
- [7] Zhang Jingcheng, Pu Ruiliang, Wang Jihua, et al. Detecting powdery mildew of winter wheat using leaf level hyperspectral measurements[J]. Computers and Electronics in Agriculture, 2012, 85: 13-23.
- [8] Zhang Jingcheng, Pu Ruiliang, Yuan Lin, et al. Monitoring powdery mildew of winter wheat by using moderate resolution multi-temporal satellite imagery[J]. PloS one, 2014, 9(4): e93107.
- [9] 崔海滨,杨柯,张龙,等.基于 Gabor 变换的 TDLAS 检测信号的降噪研究[J].光谱学与光谱分析,2016,36(9):2997-3002.
  Cui Haibing, Yang Ke, Zhang Long, et al. Tunable diode

laser absorption spectroscopy (TDLAS) detection signal denoising based on gabor transform[J]. Spectroscopy and Spectral Analysis, 2016, 36(9): 2997-3002. (in Chinese with English abstract)

[10] 鲁军景,黄文江,张竞成,等.基于小波特征的小麦白粉病与条锈病的定量识别[J].光谱学与光谱分析,2016,36(6):1854-1858.
 Ly, Juring, Hung, Wanilang, Zhang, Jingshang, et al.

Lu Junjing, Huang Wenjiang, Zhang Jingcheng, et al. Quantitative identification of yellow rust and powdery mildew in winter wheat based on wavelet feature[J]. Spectroscopy and Spectral Analysis, 2016, 36(6): 1854– 1858. (in Chinese with English abstract)

- [11] Chen Y P, Peng Z M, He Z H, et al. The optimal fractional Gabor transform based on the adaptive window function and its application[J]. Applied Geophysics, 2013, 10(3): 305– 313.
- [12] 印勇,史金玉,刘丹平.基于 Gabor 小波的人脸表情识别[J].光電工程,2009,36(5):111-116.
  Ying yong, Shi Jingyu, Liu Danping. Facial expression recognition based on Gabor wavelet transform[J]. Photoelectric engineering, 2009, 36(5): 111-116. (in Chinese with English abstract)
- [13] 牛连强,赵子天,张胜男.基于 Gabor 特征融合与 LBP 直方图的人脸表情特征提取方法[J]. 沈阳工业大学学报, 2016, 38(1): 63-68.
  Niu Lianqiang, Zhao Zitian, Zhang Shengnan. Extraction method for facial expression features based on Gabor feature fusion and LBP hiatogram[J]. Journal of Shenyang University of Technology, 2016, 38(1): 63-68. (in Chinese with English abstract)
- [14] 王道明,鲁昌华,蒋薇薇,等.基于粒子群算法的决策树 SVM 多分类方法研究[J].电子测量与仪器学报,2015, 29(4):611-615.
  Wang Daoming, Lu Changhua, Jiang Weiwei, et al. Study on PSO-based decision-tree SVM multi-class classification

method[J]. Journal of Electronic Measurement and Instrumentation, 2015, 29(4): 611-615. (in Chinese with English abstract)

- [15] 陈中杰,蔡勇,蒋刚. 复高斯小波核函数的支持向量机研究[J]. 计算机应用研究, 2012, 29(9): 3263-3265.
  Chen Zhongjie, Cai Yong, Jiang Gang. Study on SVM of complex Gaussian wavelet kernel function[J]. Journal of Computer Applications, 2012, 29(9): 3263 3265. (in Chinese with English abstract)
- [16] Wang H G, Ma Z H, Wang T, et al. Application of hyperspectral data to the classification and identification of severity of wheat stripe rust (in Chinese) [J]. Spectroscopy and Spectral Analysis, 2007, 27(9): 1811–1814.
- [17] 袁莹, 王伟, 褚璇, 等. 光谱特征波长的 SPA 选取和基于 SVM 的玉米颗粒霉变程度定性判别[J]. 光谱学与光谱分析, 2016, 36(1): 226-230.
   Yuan Ying, Wang Wei, Chu Xuan, et al. Selection of

characteristic wavelengths using SPA and qualitative discrimination of mildew degree of corn kernels based on SVM[J]. Spectroscopy and Spectral Analysis, 2016, 36(1): 226–230. (in Chinese with English abstract)

- [18] 张录达,金泽宸,沈晓南,等. SVM 回归法在近红外光 谱定量分析中的应用研究[J].光谱学与光谱分析,2005, 25(9): 1400-1403.
  Zhang Luda, Jin Zechen, Shen Xiaonan, et al. Applied study on support vector machine (SVM) regression method in quantitative analysis with near-infrared spectroscopy[J].
  Spectroscopy and Spectral Analysis, 2005, 25(9): 1400-1403. (in Chinese with English abstract)
- [19] Luo J, Wang D, Dong Y, et al. Developing an aphid damage hyperspectral index for detecting aphid (Hemiptera: Aphididae) damage levels in winter wheat[C]//Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International. Ieee, 2011: 1744-1747.
- [20] 徐涵秋. 新型 Landsat8 卫星影像的反射率和地表温度反 演[J]. 地球物理学报, 2015, 58(3): 741-747.
  Xu Hanqiu. Retrieval of the reflectance and land surface temperature of newly-launched Landsat 8 satellite[J]. Chinese Journal of Geophysics, 2015, 58(3): 741-747. (in Chinese with English abstract)
- [21] Jiménez-Muñoz J C, Sobrino J A. A generalized singlechannel method for retrieving land surface temperature from remote sensing data[J]. Journal of Geophysical Research: Atmospheres, 2003, 108(D22): ACL2.1-ACL2.9.
- [22] Verstraete M M, Pinty B, Myneni R B. Potential and limitations of information extraction on the terrestrial biosphere from satellite remote sensing[J]. Remote Sensing of Environment, 1996, 58(2): 201–214.
- [23] Chen S F, Goodman J. An empirical study of smoothing techniques for language modeling[C]//Proceedings of the 34th annual meeting on Association for Computational Linguistics. Association for Computational Linguistics, 1996: 310-318.
- [24] Rouse B T, Wells R J H, Warner N L. Proportion of T and B lymphocytes in lesions of Marek's disease: theoretical

implications for pathogenesis[J]. The Journal of Immunology, 1973, 110(2): 534-539.

- [25] Rondeaux G, Steven M, Baret F. Optimization of soil-adjusted vegetation indices[J]. Remote Sensing of Environment, 1996, 55(2): 95-107.
- [26] Roujean J L, Breon F M. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements[J]. Remote sensing of Environment, 1995, 51(3): 375-384.
- [27] Huete A R. A soil-adjusted vegetation index (SAVI) [J]. Remote sensing of environment, 1988, 25(3): 295-309.
- [28] Person R S, Kudina L P. Discharge frequency and discharge pattern of human motor units during voluntary contraction of muscle[J]. Electroencephalography and Clinical Neurophysiology, 1972, 32(5): 471–483.
- [29] Kohavi R, John G H. Wrappers for feature subset selection[J]. Artificial Intelligence, 1997, 97(1): 273-324.
- [30] Kira K, Rendell L A. The feature selection problem: Traditional methods and a new algorithm[C]//AAAI. 1992, 2: 129-134.
- [31] 孙吉贵, 刘杰, 赵连宇. 聚类算法研究[J]. 软件学报, 2008, 19(1): 48-61.
- [32] 王千,王成,冯振元,等.K-means 聚类算法研究综述[J].电 子设计工程, 2012, 20(7): 21-24.
- [33] 叶超. 基于 Gabor 小波和 SVM 的人脸识别算法研究
  [D]. 太原: 中北大学, 2014.
  Ye Chao. The Optimization Algorithm of Face Recognition Based on Gabor Wavelets and SVM[D]. Taiyuan: North University of China, 2014. (in Chinese with English abstract)
- [34] 鲁军景,黄文江,蒋金豹,等.小波特征与传统光谱特征 估测冬小麦条锈病病情严重度的对比研究[J].麦类作物学 报,2015,35(10):1456-1461.
  Lu Junjing, Huang Wenjiang, Jiang Jinbao, et al. Comparison of wavelet features and conventional spectral features on estimating severity of stripe rust in winter wheat[J]. Journal of Microbiology, 2015, 35(10):1456-1461. (in Chinese with English abstract)
- [35] 张丽新. 高维数据的特征选择及基于特征选择的集成学习研究[D]. 北京:清华大学,2004.
  Zhang Lixin. Study on Feature Selection and Ensemble Learning Based on Feature Selection for High-Dimensional Datasets[D]. Beijing: Tsinghua University, 2004. (in Chinese with English abstract)
- [36] 邓曾,李丹,柯樱海,等. 基于改进 SVM 算法的高分辨 率遥感影像分类[J]. 国土资源遥感, 2016, 28(3): 12-18.
- [37] 何东健,乔永亮,李攀,等.基于 SVM-DS 多特征融合的杂草识别[J]. 农业机械学报,2013,44(2):182-187.
  He Dongjiang, Qiao Yongliang, Li Pan, et al. Weed recognition based on SVM-DS multi-feature fusion[J]. Journal of agricultural machinery, 2013, 44(2):182-187. (in Chinese with English abstract)
- [38] 张学工.关于统计学习理论与支持向量机[J].自动化学报, 2000, 26(1): 32-42.
- [39] Huang D Y, Hu W C, Chang S H. Vision-based hand gesture recognition using PCA+ Gabor filters and SVM[C]// Intelligent Information Hiding and Multimedia Signal

Processing, 2009. IIH-MSP'09. Fifth International Conference on. IEEE, 2009: 1–4.

- [40] Sun Z, Bebis G, Miller R. On-road vehicle detection using Gabor filters and support vector machines[C]//Digital Signal Processing, 2002. DSP 2002. 2002 14th International Conference on. IEEE, 2002, 2: 1019-1022.
- [41] 郑秋红,杨霏云,朱玉洁.小麦白粉病发生气象条件和气象预报研究进展[J].中国农业气象,2013,34(3):358-365.
  Zheng Qiuhong, Yang Feiyun, Zhu Yujie. Meteorological factors and its forecast on wheat powdery mildew occurrence in China[J]. Chinese Journal of Agrometeorology, 2013, 34(3):358-365. (in Chinese with English abstract)

# Remote sensing monitoring of winter wheat powdery mildew based on wavelet analysis and support vector machine

Huang Linsheng<sup>1</sup>, Liu Wenjing<sup>1,2</sup>, Huang Wenjiang<sup>1,2\*</sup>, Zhao Jinling<sup>1</sup>, Song Furan<sup>1,2</sup>

 Anhui Engineering Laboratory of Agro-Ecological Big Data, Anhui University, Hefei 230601, China;
 Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China)

Abstract: Wheat powdery mildew is one of the main serious diseases for winter wheat. A fast and accurate monitoring of the disease at a regional scale plays a vital role in reducing yield loss. Remote sensing data has great advantages over traditional data in disease monitoring, including simpler operation, more real-time and higher resolution. In this study, Chinese HJ-1A/1B data with high revisit frequency and 30 m spatial resolution was used to inverse Land Surface Temperature (LST), extract four-band reflectance data, and build seven vegetation indices. These indices should be filtrated to improve accuracy of the model due to redundancy of them. Then, we implemented screening features with the combination of Relief and K-mean algorithm. Relief algorithm which can provide the basis for feature evaluation, so features were ranked in descending order judged by feature weights in preparation for the next process. Clustering accuracy obtained by K-mean algorithm. According to the weight of the feature, the features clustered in turn to perform K-mean analysis. Then the cluster with the highest precision was picked out, and we finally got the normalized difference vegetation index (NDVI), Simple vegetation index (SR) and surface temperature (LST) as the feature set. Wavelet feature can decompose the data in multi-scale and multi-direction, which can highlight the sensitive factor of vegetation index to a certain extent. Forty wavelet functions were constructed from five scales and eight directions, and made them convolve with features. Because there were too many wavelet features after convolved, the independent T-test samples were used to obtain the most sensitive wavelet feature of disease and the corresponding wavelet kernel function. After this process, three features corresponding to vegetation indices were available. These three wavelet features were used as input variables of the model. Support vector machine is a kind of machine learning method based on statistical learning theory. Its core idea is to minimize the structural risk by mapping the input linear indivisible data to the high dimensional space, which makes the difference between different samples. The class interval is the largest while the intra-class interval is the smallest, then the hyper plane is constructed to classify data. The monitoring model of wheat powdery mildew in Jinzhou City of Hebei Province was established by using support vector machine (SVM) with three groups of features. The first group used twelve vegetation indices as the input variables of the model, which served as a control group. The second one used three features after feature selection and the third used three features of the wavelet transform. Then the monitoring precision of the three models was compared and analyzed. The experimental results showed that the overall accuracy and the kappa coefficient of the third model (called GaborSVM) were 86.7% and 0.583, respectively, performing better over the first model (60%, 0.286) and the second model (80%, 0.444). These results also showed that the combined method of wavelet analysis with SVM (GaborSVM) can be applied to large area disease monitoring based on satellite remote sensing image, and has important application value in improving the accuracy of disease monitoring. Keywords: remote sensing; support vector machine; diseases; powdery mildew; wavelet feature