

DOI:10.3969/j.issn.1000-1298.2010.05.035

# 冬小麦叶绿素含量高光谱检测技术\*

王伟<sup>1</sup> 彭彦昆<sup>1</sup> 马伟<sup>2</sup> 黄慧<sup>1</sup> 王秀<sup>2</sup>

(1. 中国农业大学工学院, 北京 100083; 2. 国家农业信息化工程技术研究中心, 北京 100097)

**【摘要】** 以大田冬小麦叶绿素含量为研究对象,首先利用高光谱成像系统以线扫描方式获取其反射光谱图像,选择感兴趣区域(ROI)并计算出光谱平均反射率值;然后分别针对其原始光谱和一阶差分光谱,通过相关分析和逐步回归分析,得到能反映叶绿素含量变化的7个最佳优化波长;进而基于该优化波长采用多元线性回归(MLR)方法组建模型,通过假设检验剔除对模型贡献不显著的3个波长变量。选用剩余的4个波长即710.85、767.42、650和520 nm作为自变量重新建立模型,基于校正集和预测集模型的决定系数 $R^2$ 分别为0.8434和0.7093。研究表明,利用高光谱技术检测大田冬小麦叶绿素含量的方法是可行的。

**关键词:** 冬小麦 叶绿素含量 高光谱图像 多元线性回归

**中图分类号:** O433.4 **文献标识码:** A **文章编号:** 1000-1298(2010)05-0172-06

## Prediction of Chlorophyll Content of Winter Wheat Using Leaf-level Hyperspectral Data

Wang Wei<sup>1</sup> Peng Yankun<sup>1</sup> Ma Wei<sup>2</sup> Huang Hui<sup>1</sup> Wang Xiu<sup>2</sup>

(1. College of Engineering, China Agricultural University, Beijing 100083, China

2. National Engineering Research Center for Information Technology in Agriculture, Beijing 100097, China)

### Abstract

The leaf-level winter wheat hyperspectral response to its chlorophyll content was examined. Firstly, after the 316 scan line images were acquired, the cube image data was constructed and the region of interest (ROI) was selected, then after the average pixel intensity acquired, using correlation analysis combined with stepwise discrimination method for the origin reflective spectrum and the first derivative spectrum, the optimal wavelengths were selected respectively; the chlorophyll content model using multivariate linear regression (MLR) was constructed based on the above seven optimal wavelengths. After statistical significance testing, three wavelengths were abandoned, and the residual four wavelengths, i. e., 710.85, 767.42, 650 and 520 nm were used to construct chlorophyll content prediction model. The prediction results showed that the determination coefficient were  $R^2 = 0.8434$  and  $R^2 = 0.7093$  for the training dataset and the validation dataset respectively. All of these indicated that with the hyperspectral technology, chlorophyll content of winter wheat could be predicted precisely.

**Key words** Winter wheat, Chlorophyll content, Hyperspectral image, Multivariate linear regression

### 引言

叶绿素含量与植被的光合能力、发育阶段以及氮素状况有着较好的相关性<sup>[1-2]</sup>,通常是氮素胁迫、

光合作用能力和植被发育阶段的指示器<sup>[3]</sup>。因此,定量测定小麦叶片的叶绿素含量在小麦估产、农情监测等方面具有重要意义。测量叶绿素含量的传统方法是化学分析法<sup>[4]</sup>。该方法费时、耗力<sup>[5]</sup>,无法

收稿日期:2009-07-24 修回日期:2009-08-24

\* 国家“863”高技术研究发展计划资助项目(2006AA10A308、2006AA10A305-1)和“十一五”国家科技支撑计划资助项目(2007BAD89B04)

作者简介:王伟,讲师,主要从事农畜产品品质和安全无损检测技术研究,E-mail: playerwxw@cau.edu.cn

实现快速、非破坏测量。而测量光谱反射率方法具有非破坏、快速的特点,可以作为植被生化成分十分有效的检测方法<sup>[6-8]</sup>。图像光谱技术集成了传统的图像和光谱技术,可同时获得被测物体的空间和光谱信息,尤其是高光谱图像技术因其高的光谱分辨率已成为农产品品质评估和安全检测的强有力工具<sup>[9]</sup>,国外研究表明通过将叶绿素含量、含水率等与高光谱反射图像数据建立关系,可以在植被叶片或冠层水平定量确定其生化成分及其含量<sup>[10-11]</sup>。本文利用实验室获取的冬小麦叶片线扫描图像构建 ENVI 格式的立方体图像,选择感兴趣区域并计算区域内反射光谱平均像素密度,利用逐步回归结合相关分析确定能够表征叶片叶绿素含量的最佳优化波长,并利用多元线性回归(multivariate linear regression,简称 MLR)结合假设检验构建并验证叶绿素含量预测模型,以评估采用高光谱图像数据预测大田冬小麦叶绿素含量的可行性,并评价所建立模型的预测精度。

## 1 材料与方法

### 1.1 试验材料

研究区域位于北京市昌平区小汤山镇国家精准农业示范基地。以冬小麦京冬2号作为研究对象,选择了来自4个氮胁迫区域的60片叶子,其中每个氮胁迫区域15片。剪下的叶子放入透明的聚乙烯采样袋,然后放进便携式冰箱中,并在2 h内尽快送回中国农业大学农畜产品无损检测实验室。剔除样本后共计44片叶子作研究用,选择属于1、2、3号氮胁迫区域的共计33片小麦叶片作为校正集样本,4号区域的11片小麦叶片作为验证集样本。

### 1.2 高光谱成像系统结构组成

采用的高光谱成像系统如图1所示。系统主要包括一个高性能的背照式 CCD 相机(Sencicam QE 型, Germany)及其控制单元,一台波长范围在400~1 000 nm、光谱分辨率为2.8 nm的图像光谱仪(ImSpector V10E 型, Spectral Imaging Ltd., Finland),由石英卤钨灯和稳压电源组成的光源系统(Oriel Instruments 66882 型, USA),试样载物台及其调节机构以及聚光镜等组成。使用前参考文献[12]的方法对高光谱图像系统进行空间和光谱维校准。试样载物台可沿高度方向进行调节并可在步进电动机的带动下向前或向后移动,光源为可以覆盖整个样本并使各处光强一致的平面光源。小麦叶片两端用橡皮筋压紧放置在载物台上,当光源照射在叶片表面时,叶片组织的漫反射光经聚光镜、光谱仪通过 CCD 相机形成光谱图像,经数据采集卡生成 16

位图像数据文件。当移动平台向前移动时,高光谱图像系统根据所设定曝光时间和读取时间逐条扫描样本,光谱仪将扫描线进行分光,分光后投射在 CCD 探测器上生成二维图像:其中一维代表扫描线的空间轴,另一维代表其光谱轴。系统所采用 CCD 相机的分辨率为1 376×1 040,利用 CCD 相机自带软件设定数据进仓值为2×2,则在扫描线图像的空间轴方向可表达每条扫描线的688个空间位置,而在光谱轴方向可将扫描线图像分光为520个波长。

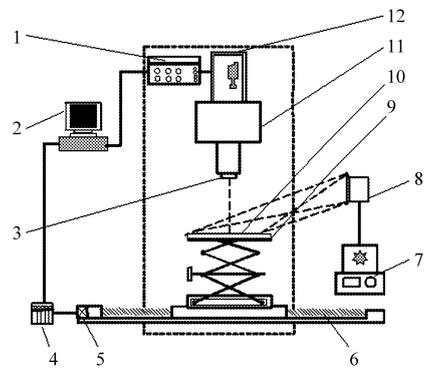


图1 高光谱图像系统架构

Fig. 1 Hyperspectral imaging system for measuring the chlorophyll content

1. 相机控制器 2. 计算机 3. 聚光镜 4. 步进电动机驱动器
5. 步进电动机 6. 丝杠 7. 光源控制器 8. 面光源 9. 载物台
10. 小麦 11. 光谱仪 12. CCD 相机

### 1.3 叶绿素含量检测

所有叶片的光谱测量完成后,采用湿化学方法测定叶片叶绿素含量作为建模的标准参照值,过程如下:按照2:1比例配制丙酮与乙醇混合液,将配好的溶液倒入试管中,体积为25 mL,取新鲜小麦叶片,去除叶脉后将叶片剪成细丝,称取0.4 g,将其装入装有25 mL丙酮/乙醇混合液的试管中,封口,放在暗处24 h,中间振荡3次。摇匀后用紫外分光光度法分别测出663、645和652 nm处的吸光度,然后依据叶绿素a和叶绿素b的浓度与吸光度的关系参照文献[13]方法计算出总的叶绿素含量。

### 1.4 图像获取和预处理

利用自编的步进电动机平移台及相机曝光控制软件控制图像的获取过程,采用 Matlab 和 ENVI 进行图像处理与分析。为提高扫描效率,一次将来自4个不同施肥区域的小麦叶片作为一组,叶片两端分别用橡皮筋压紧并平行放置在黑色背景板上,然后按照小麦叶片长度方向与光谱仪狭缝垂直方向将背景板放置在载物台上,这样,一次线扫描可对4片小麦叶片同时进行扫描并成像在同一幅图像中,依据小麦样本的长度范围,对每一组小麦叶片共进行316次线扫描,每次扫描的结果储存为一幅图像。

扫描的图像数据均通过 USB 接口送入计算机作进一步处理用。

扫描试样漫反射图像过程中,关上试验箱门以防止外部环境光影响。设定 CCD 的曝光时间为 0.025 s,调整光密度使其最大值为相机总动态范围的 1/3。在采集所有试样图像之前,为校正相机暗电流和室内照明光对图像的影响,分别采用盖住镜头和采集标准白板 (Spectralon by Labsphere, North Sutton, NH, U. S. A) 图像的方法获取黑、白图像,相对反射密度计算公式为

$$R = (R_s - R_d) / (R_r - R_d)$$

式中  $R$ ——相对反射密度

$R_s$ ——试样原始图像反射密度

$R_r$ ——标准参考白板反射密度

$R_d$ ——黑色图像反射密度

所有的光谱数据在各单个波长沿扫描线上取平均值。包括图像参数在内的图像数据均以 16 位二进制格式文件存放。

为便于用 ENVI 软件分析线扫描图像,首先采用 VC++ 编写应用程序将每组小麦样本的 316 次线扫描图像转换为 BSQ 格式的立方体图像数据。图 2 显示经分光后 3 个不同波长的原始反射图像及

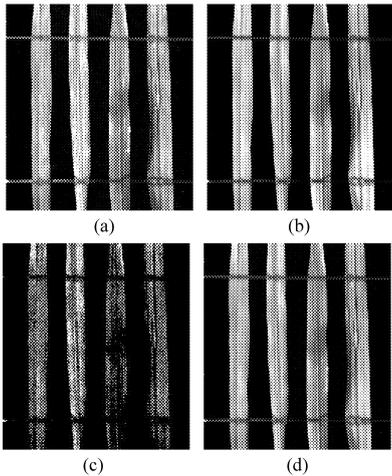


图 2 原始反射光谱图像

Fig. 2 Original reflectance image

(a) 420 nm (b) 550 nm (c) 900 nm (d) 伪彩色图像

用此 3 个波长图像合成的伪彩色图像。进一步通过选择感兴趣区域 (ROI) 获取光谱数据集,此处选择平均像素点数为 900 的远离叶脉的长方形区域作为感兴趣区域 (如图 3 所示),将感兴趣区域内所有像素反射密度的平均值作为该叶片样本的最终反射密度。

## 2 优化波长选取

为克服光谱仪波谱范围两端的干扰,仅取波段

450 ~ 800 nm 用作叶绿素含量定量分析。即便如此,若采用该范围内全部波段的数据直接进行建模则数据量大、计算耗时长,最重要的是许多波长可能都与叶绿素、水等某个生化信息存在某种程度的关联性,因此,若直接采用剩余全部波长进行建模,波

长之间多重共线性的影响会十分显著,反而不利于获得较好的建模精度。为解决该问题,本文将建立在高分辨率基础上的高光谱问题 (提供高的光谱分辨率以使信息的精细辨识成为可能) 转化为简单易行且精度高的多光谱问题,达此目的关键在于找到能够反映叶绿素含量绝大多数方差变动的几个最关键的波长。

本文采用相关分析与有进有出逐步回归分析相结合的方法确定优化波长。除了原始反射光谱,通常反射光谱数据的差分变换也能较好地反映方差变动。图 4a、4b 分别为 44 个小麦叶片的原始反射光谱曲线和一阶差分光谱曲线,从图 4a 叶片的原始光谱可以看出小麦叶片具有绿色植物所共有的绿峰、红谷、近红外高的反射平台等典型特征,小麦叶片 44 个样本差异信息集中体现在 510 ~ 550 nm,以及 620 ~ 690 nm 波段范围;而从图 4b 可以看出 44 个样本的差异信息集中体现在 480 ~ 510 nm 以及 700 ~ 770 nm 波段范围,即一阶差分光谱高的峰值上,对应原始反射光谱的相应位置两者均反映出光谱曲线出现了拐点,前者似乎包含有蓝边的变化信息,而后者似乎与红边现象有关。

对 44 个样本原始反射光谱及一阶差分光谱分别与叶绿素含量进行相关分析,进一步证实了上述方法的正确性。对应于原始反射光谱曲线,图 5a 反映出在 510 ~ 550 nm 波段范围以及 650 nm 附近波段范围具有明显高的相关系数,尤其在 520 ~ 540 nm 波段范围的相关系数均大于 0.1,而 650 nm 附近为 -0.3。同样,对应于一阶差分光谱曲线,从图 5b 可以看到 480 ~ 510 nm 以及 700 ~ 770 nm 波段范围的相关系数明显高于其他波段范围,在 480 ~ 510 nm 波段,相关系数接近 0.6,而 710 ~ 720 nm 和 760 nm 附近的相关系数也分别达到 -0.4 和 0.4。

进一步采用 Matlab 软件进行有进有出逐步回归分析确定优化波长,设定因子进入与退出模型的显著性水平为 0.15。表 1 和表 2 分别表示针对原始光谱和一阶差分光谱各个优化波长选进选出时决

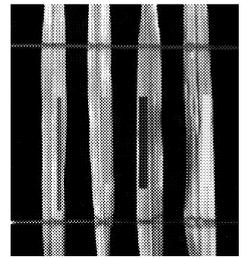


图 3 选择图像感兴趣区域  
Fig. 3 Region of interest (ROI) selected

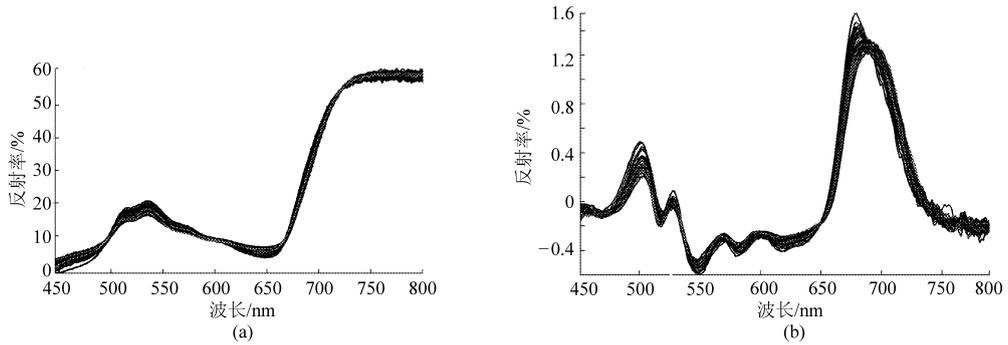


图 4 原始反射光谱相对反射率曲线及其相应一阶差分光谱曲线

Fig. 4 Curves of original spectrum and first derivative spectrum

(a) 原始反射光谱 (b) 一阶差分光谱

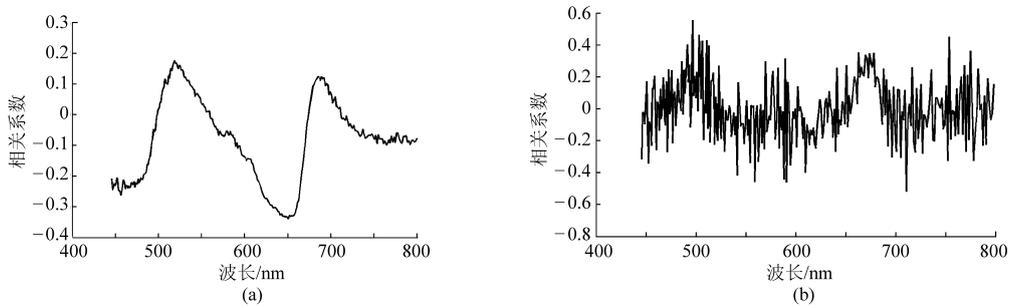


图 5 原始反射光谱与一阶差分光谱波长与叶绿素含量的相关性分析

Fig. 5 Correlation coefficient between chlorophyll content and wavelengths

(a) 原始光谱 (b) 一阶差分光谱

定系数和均方根误差的变化情况,对原始反射光谱进行逐步回归获得的优化波长点为 650(对应于叶绿素 b 的强吸收)、520、542.5 和 535.58 nm(叶绿素绿色强反射峰附近);而对一阶差分光谱进行逐步回归获得的优化波长点为:496.68、710.85 和 767.42 nm。这一结果与上述直观观察和相关分析基本一致。

表 1 逐步回归法对原始光谱进行波长选择过程

Tab.1 Wavelength selection course of stepwise regression for the origin spectrum

	逐步加入的波长/nm				
	650	542.5	520	535.58	650(out)
$R^2$	0.115 0	0.310 0	0.401 6	0.473 6	0.452 6
均方根误差	0.437 2	0.390 7	0.368 3	0.349 8	0.352 3

表 2 逐步回归法对一阶差分光谱进行波长选择过程

Tab.2 Wavelength selection course of stepwise regression for the first derivative spectrum

	逐步加入的波长/nm		
	496.68	710.85	767.42
$R^2$	0.305 5	0.423 5	0.497 6
均方根误差	0.387 2	0.357 1	0.337 5

### 3 模型的建立与验证

考虑到波长选取后变量数目已较少,采用 MLR

方法建立叶绿素含量的预测模型。首先单独采用基于原始光谱所获得的 4 个优化波长即 650、520、542.5、535.58 nm 和基于差分光谱所获得的 3 个优化波长 496.68、710.85、767.42 nm 作为自变量分别建立模型,试验表明两个模型的预测结果均欠佳,表明单独用这些优化波长进行建模有效信息不够充分。

将原始光谱及一阶差分光谱所获得的所有 7 个波长组合在一起,即波长 496.68、710.85、767.42、650、520、542.5、535.58 nm 作为自变量来建立模型。利用 Excel 进行回归统计的结果表明决定系数  $R^2$  仅为 0.567 2,调整的决定系数进一步降低,仅为 0.483 0。首先对模型总体的线性度进行假设检验,计算  $F = 6.7395 > F(7, 25) = 2.40$ ,可以看出线性模型拟合不够显著。进而对每个回归系数的显著性进行检验的结果如表 3 所示,通过第 4 列的  $t$  分布概率密度可见变量 1(与单独采用一阶差分光谱优化波长分析的结果一致)、变量 6、变量 7 所对应的概率值均高于通常的显著性水平  $\alpha = 0.05$ ,即变量 1 和变量 6、7 对模型的贡献不显著,应该从模型中剔除。

因此只选用剩余的 4 个波长即 710.85、767.42、650 和 520 nm 所对应数据作为自变量重新建立模型,最终得到的模型为  $y = 1.917 354 + 25.597 23\lambda_{710.85} - 17.081 8\lambda_{767.42} - 52.343 7\lambda_{650.00} +$

41.4550  $4\lambda_{520.00}$ , 回归统计的结果为决定系数  $R^2 = 0.8917$ , 调整的决定系数为 0.8806, 两者均已较高且非常接近, 表明模型预测能力有了较大增强。通过假设检验对模型总体的线性度进行检验, 如表 4 所示,  $F = 80.29267 \gg F(4, 28) = 2.71$ , 可见线性模

型已十分显著。

对每个回归系数的显著性检验的结果如表 5 所示, 包括常数项在内, 所有回归系数所对应的概率值远小于通常的显著性水平  $\alpha = 0.05$ , 表明用剩余的 4 个波长作为预测变量进行建模是合理的。

表 3 基于 7 个优化波长所建模型的总体线性度假设检验

Tab. 3 Linearity hypothesis test of model based on 7 optimal wavelengths

	系数	均方根误差	统计量 $t$	概率 $P$	95% 置信下限
截距	1.9188	0.6091	3.1502	0.0033	0.6835
变量 1	-23.6921	16.6987	-1.4188	0.1646	-57.5586
变量 2	28.6612	13.2981	2.1553	0.0379	1.6915
变量 3	-21.2825	9.2224	-2.3077	0.0269	-39.9864
变量 4	-129.7850	34.8379	-3.7254	0.0007	-200.4390
变量 5	56.7166	19.5347	2.9034	0.0063	17.0985
变量 6	35.6190	19.1259	1.8623	0.0707	-3.1701
变量 7	66.0275	36.5652	1.8057	0.0793	-8.1302

表 4 基于 4 个优化波长所建模型的总体线性度假设检验

Tab. 4 Linearity hypothesis test of model based on 4 optimal wavelengths

	自由度	离差平方和	平均离差平方和	统计量 $F$	大于 $F$ 的概率
回归分析	4	4.0059	1.0015	80.2927	$2.74 \times 10^{-18}$
残差	39	0.4864	0.0125		
总计	43	4.4923			

表 5 基于 4 个优化波长所建模型回归系数的显著性检验

Tab. 5 Regression coefficient remarkable test of model based on 4 optimal wavelengths

	系数	均方根误差	统计量 $t$	概率 $P$	95% 置信下限
截距	1.9174	0.1691	11.3408	$6.45 \times 10^{-14}$	1.5754
变量 1	25.5972	2.9097	8.7973	$8.46 \times 10^{-11}$	19.7119
变量 2	-17.0818	1.8720	-9.1248	$3.21 \times 10^{-11}$	-20.8683
变量 3	-52.3437	4.1004	-12.7655	$1.66 \times 10^{-15}$	-60.6375
变量 4	41.4550	4.5065	9.1989	$2.58 \times 10^{-11}$	32.3397

利用所建立的模型分别针对校正集和预测集的建模结果如图 6a、6b 所示, 决定系数  $R^2$  分别为 0.8434 和 0.7093, 校正均方根误差 RMSEC 和验证均方根误差 RMSEV 分别为 0.1117 和 0.2551。可见其基于校正集和验证集的模型预测结果均已比较

理想。

## 4 结束语

本文采用线扫描方式对冬小麦叶片进行扫描, 经过数百次扫描获得小麦叶片的面光谱图像, 该方

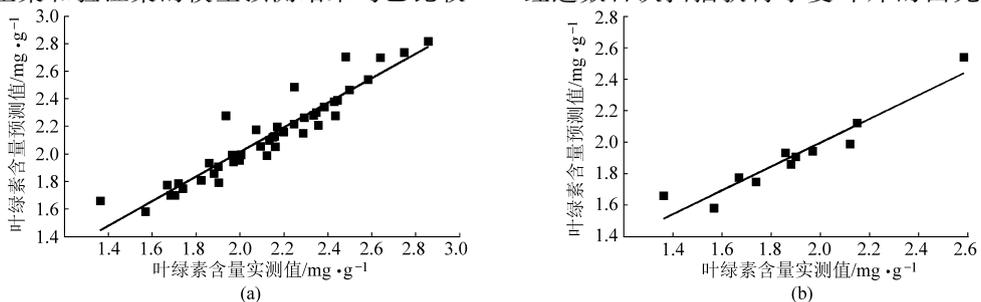


图 6 基于 4 个优化波长的 PLSR 模型性能

Fig. 6 Performance of PLSR model based on data of 4 bands

(a) 校正集 (b) 验证集

法在反映样本生化信息的同时,充分涵盖了叶片的空间差异信息,克服了传统地物高光谱方法只包含光谱信息的缺憾。通过在远离叶脉处选择感兴趣区域,并计算 ROI 内光谱反射率的平均值,可部分克服叶片局部不平整所带来的反射光谱光强变化的问题。基于高光谱图像光谱数据,分别针对大田冬小麦原始光谱和一阶差分光谱,通过直观观察、相关分析和逐步回归分析方法,得到能够反映不同施氮水

平下冬小麦叶片叶绿素含量方差最大变化的7个优化波长。然后基于上述波长采用 MLR 组建模型,并对所建立的模型进行假设检验,结果表明 496.68、535.58 和 542.5 nm 3 个波长变量对模型的贡献不显著,因此利用剩余 4 个波长变量重新建立模型,模型的预测结果显示,小麦叶片叶绿素含量的预测值和标准值之间的决定系数  $R^2$  分别为 0.843 4 和 0.709 3。

#### 参 考 文 献

- 1 孟卓强,胡春胜,程一松. 高光谱数据与冬小麦叶绿素密度的相关性研究[J]. 干旱地区农业研究, 2007, 25(6): 74~79.  
Meng Zhuoqiang, Hu Chunsheng, Cheng Yisong. Study on correlation between chlorophyll density of winter wheat and hyperspectral data[J]. Agricultural Research in the Arid Areas, 2007, 25(6): 74~79. (in Chinese)
- 2 张金恒,王珂,王人潮. 高光谱评价植被叶绿素含量的研究进展[J]. 上海交通大学学报: 农业科学版, 2003, 21(1): 74~80.  
Zhang Jinheng, Wang Ke, Wang Renchao. Study on hyperspectral remote sensing in estimate vegetation leaf chlorophyll content[J]. Journal of Shanghai Jiaotong University: Agricultural Science, 2003, 21(1): 74~80. (in Chinese)
- 3 赵祥,刘素红. 基于高光谱数据的小麦叶绿素含量反演[J]. 地理与地理信息科学, 2004, 20(3): 36~39.  
Zhao Xiang, Liu Suhong. Method for inverting chlorophyll content of wheat using hyperspectral[J]. Geography and Geo-information Science, 2004, 20(3): 36~39. (in Chinese)
- 4 Arnon D L. Copper enzymes in isolated chloroplasts. Polyphenoloxidase in *Beta vulgaris* [J]. Plant Physiology, 1949, 24(1): 1~15.
- 5 吉海彦,王鹏新,严泰来. 冬小麦活体叶片叶绿素和水分含量与反射光谱的模型建立[J]. 光谱学与光谱分析, 2007, 27(3): 514~516.  
Ji Haiyan, Wang Pengxin, Yan Tailai. Estimations of chlorophyll and water contents in live leaf of winter wheat with reflectance spectroscopy[J]. Spectroscopy and Spectral Analysis, 2007, 27(3): 514~516. (in Chinese)
- 6 Jongschaap R E E, Booij R. Spectral measurements at different spatial scales in potato: relating leaf, plant and canopy nitrogen status[J]. Int. J. Appl. Earth Observ. Geoinform, 2004, 5(3): 205~218.
- 7 Vianney H, Martine G, Bruno M. Elaboration of a nitrogen nutrition indicator for winter wheat based on leaf area index and chlorophyll content for making nitrogen recommendations[J]. Europ. J. Agronomy, 2007, 27(1): 1~11.
- 8 杨昊谕,于海业,张蕾,等. 基于激光诱导荧光光谱分析的黄瓜叶片叶绿素含量检测[J]. 农业机械学报, 2009, 40(10): 169~172.  
Yang Haoyu, Yu Haiye, Zhang Lei, et al. Detecting of chlorophyll content of cucumber leaves based on laser-induced fluorescence spectrum analysis technique[J]. Transactions of the Chinese Society for Agricultural Machinery, 2009, 40(10): 169~172. (in Chinese)
- 9 Kim M S, Daughtry C S T, Chappelle E W. The use of high spectral resolution bands for estimating absorbed photosynthetically active radiation (APAR)[C]//Proceedings of the 6th Symp. on Physical Measurements and Signatures in Remote Sensing. Val D'Isere, France, 1994: 299~306.
- 10 乔欣,马旭,张小超,等. 大豆叶绿素和钾素信息的冠层光谱响应[J]. 农业机械学报, 2008, 39(4): 108~111.  
Qiao Xin, Ma Xu, Zhang Xiaochao, et al. Response of coronary spectrum on chlorophyll and K information of soy[J]. Transactions of the Chinese Society for Agricultural Machinery, 2008, 39(4): 108~111. (in Chinese)
- 11 Gitelson A A, Gritz Y, Merzlyak M N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves[J]. Journal of Plant Physiology, 2003, 160(3): 271~282.
- 12 Lu R, Peng Y K. Hyperspectral scattering for assessing peach fruit firmness[J]. Biosystems Engineering, 2006, 93(2): 161~171.
- 13 李民赞. 光谱分析技术及其应用[M]. 北京: 科学出版社, 2006: 45~48.