

基于扫描图像光谱特征和模式识别的水稻叶片磷素诊断研究

晓东, 石媛媛, 路 雪, 邓劲松, 沈掌泉, 王 珂*

浙江大学农业遥感与信息技术应用研究所, 环境修复与生态健康教育部重点实验室,
浙江省农业遥感与信息技术应用重点实验室, 浙江 杭州 310029

摘 要 磷的缺乏对水稻产量有很大影响, 磷的无损快速营养诊断对缺磷水稻生产有十分重要的意义。该文以水稻不同磷营养水平的顶部三张完全展开叶图像为研究对象, 综合提取图像光谱信息表现出的颜色、纹理和形状三类共 26 个叶片特征指数并进行单因子特征分析, 结合 CfsSubsetEval+ Scattersearch 方法对 26 个特征属性进行优化组合、评价和选择, 根据不同叶位的特征指数选择结果, 利用粗糙集理论将不同磷营养水平叶片图像样本分为三类: 极缺、微缺、正常。由识别精度可知, 严重缺磷样本识别率最高, 第三叶为水稻磷营养诊断的最佳叶位。

关键词 扫描图像; 光学特征; 缺磷; 诊断

中图分类号: S123; S127 文献标识码: A DOI: 10.3964/j.issn.1000-0593(2011)05-1336-04

引 言

磷是水稻生长所需的大量元素之一, 是生长代谢过程中不可缺少的元素, 对作物的正常生长和稳定产量有重要作用。磷肥的施用在生产上虽受重视, 但因磷在土壤中存在和贮存的形式及特定的化学行为, 导致其利用率很低, 磷的缺乏历来是植物生长的主要限制性因子之一^[1]。我国缺磷土壤面积很大, 特别是我国南方水稻土缺磷情况更为严重, 对水稻产量产生很大的影响^[2]。因此磷素诊断是保证水稻营养水平的关键环节之一。目前磷素诊断基本以实验室内农化分析为主, 成本高且费时费力, 需要发展高效快速且低成本的磷素诊断方法。研究表明水稻缺磷后生长明显缓慢, 叶片细瘦, 叶色呈暗绿或灰绿色且缺乏光泽^[3], 影响叶片在可见光等波段的反射光谱特征。因此可根据叶片在形状、颜色和纹理等方面的差异, 采用光谱分析技术对缺磷的程度进行评价。

数码相机拍摄的彩色图像目前被用于作物氮素营养诊断研究^[4-6], 在磷素诊断方面少有涉及。这种数据采集方法设备简单便携、图像获取方便, 将会更多地用于作物快速营养诊断。虽然该方法获取方式方便, 但图像冗余信息多, 噪声大, 影响数据分析; 此外数码相机拍摄的图像不便于校正,

不利于准确表达尺寸、形状等的特征。图像扫描作为一种封闭环境下的成像方式, 可提供相同的扫描环境以保证真实的色彩和尺寸还原, 最大程度降低背景带来的影响, 可作为新型的图像获取方式进行营养诊断。目前利用扫描图像进行营养诊断的研究尚不多见^[7], 而涉及到磷素诊断的研究还没有报道。此外拥有更强的移动性和功能性的便携性扫描仪已开始普及, 这也将拓展扫描图像在营养元素诊断方面的应用空间。

本文使用彩色高分辨率水稻叶片扫描图像, 根据空间和光学特征提取水稻叶片的颜色、纹理、形状和其他光谱特征属性, 采用 CfsSubsetEval+ Scattersearch 法对众多属性进行选择 and 约简, 根据粗糙集理论生成分类规则, 对水稻叶片磷素的缺失程度进行评价。这将为水稻叶片缺磷程度的快速评价提供有效手段, 也将为植物营养元素的诊断提供技术基础。

1 材料与方法

1.1 数据获取

水稻水培试验 2009 年在浙江大学玻璃温室进行, 供试水稻品种分别为中浙优 1 号。为避免环境因素的影响, 本试验采取水培方式进行, 设 3 个磷肥水平, 分别为 0, 2 5,

收稿日期: 2010-07-19, 修订日期: 2010-10-08

基金项目: 国家自然科学基金项目(30571112, 30800703), 浙江大学青年科技创新专项项目(2009QNA6016)和国家(863计划)项目(2006AA10Z204)资助

作者简介: 丁晓东, 1982年生, 浙江大学农业遥感与信息技术应用研究所博士研究生 e-mail: tedxd@163.com

* 通讯联系人 e-mail: kwang@zju.edu.cn

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

5.0, 7.5 mg · L⁻¹, 其中 7.5 mg · L⁻¹ 是水稻正常生长所需磷浓度, 培养液中的其他营养元素水平参照国际水稻所营养液配方。

研究选择 EPSON GT20000 型平板扫描仪, 采用彩色 CCD 方式扫描, 光学分辨率为 600 × 1 200 DPI, 色彩转换方式为 24 Bit 彩色。为避免叶片枯萎卷曲产生颜色变化, 取植株顶部三张完全展开叶片鲜样即刻扫描。叶片各属性指数均在 MATLAB 的软件环境中提取。为避免噪声的影响, 采用中值滤波对图像进行预处理。

1.2 叶片颜色指数的选取和计算

数字图像处理中, 面向硬件设备的最常用彩色空间是 RGB 空间, 而面向色彩处理的最常用颜色空间是 HIS 空间, 其中, HIS 空间更符合人的视觉特性。本文利用两个彩色空间的六个分量(R, G, B, H, S, I)作为颜色特征指数^[8]。此外, 考虑到植物叶片对蓝、绿波段的敏感性, 增加两波段的比值 G/B 作为特征指数。

1.3 叶片纹理指数的选取和计算

有规则的颜色变化形成了图像纹理, 图像灰度直方图特征可以提供图像中不同灰度值的量的变化, 而对表示灰度变化的灰度差进行统计分析可以得到一种描述纹理空间分布特征的方法。因此, 本文计算了图像灰度直方图和灰度差分统计法的 9 种特征指数来描述叶片纹理变化, 各纹理指数的具体计算方法见表 1。

1.4 叶片形状指数的选取和计算

将 RGB 图像进行二值分割, 分割后叶片目标内部产生的小孔洞通过膨胀和填充处理进行消除, 统计前景区域的像素点数得到水稻叶片面积; 计算叶片边缘每两个相邻点的距离的和得到叶片的周长。试验中不同磷素水平、不同位置的叶片形状存在一定差异, 例如有些叶片短而宽, 有些则较为细长, 因此本文引入圆形度这一形状作为特征属性之一。各形状指数的计算公式见表 1^[9]。

Table 1 Equations for calculating texture and shape indices

纹理指数		形状指数	
名称	公式	名称	公式
均值	$\bar{b} = \frac{1}{L} \sum_{b=0}^{L-1} bP(b)$	偏心率	$Eccentricity = \frac{AxisLength_{long}}{AxisLength_{short}}$
差	$\sigma_b^2 = \frac{1}{L} \sum_{b=0}^{L-1} (b - \bar{b})^2 P(b)$	矩形度	$Rectangularity = \frac{Area_{object}}{Area_{bounding-box}}$
倾斜度	$b_n = \frac{1}{\sigma_b^2} = \frac{1}{L} \sum_{b=0}^{L-1} (b - \bar{b}) P(b)$	面积凹凸比	$Area_convexity = \frac{Area}{Convex_Area}$
峭度	$b_k = \frac{1}{\sigma_b^4} = \frac{1}{L} \sum_{b=0}^{L-1} (b - \bar{b})^4 P(b) - 3$	圆形度	$Circularity = \frac{R_{inscribedcircle}}{R_{extrcle}}$
能量	$b_N = \frac{1}{L} \sum_{b=0}^{L-1} [P(b)]^2$	形状参数	$Form_factor = \frac{4\pi \times Area}{Perimeter^2}$
对比度	$COM = \sum_i i^2 p_{\Delta}(i)$	周长 × 圆形度	$PC = Perimeter \times Circularity$
方向角度二阶矩	$ASM = \sum_i [p_{\Delta}(i)]^2$	面积/周长	$AP_ratio = \frac{Area}{Perimeter}$
熵	$ENT = - \sum_i p_{\Delta}(i) \lg p_{\Delta}(i)$		

表中 b 为量化值, L 为量化值范围, $P(b)$ 为被测窗口内灰度值为 b 的像素所占比例 $p_{\Delta}(i)$ 为灰度差分取不同灰度值时的概率

1.5 属性约简和选择

属性在经过离散化处理后, 采用 CfsSubsetEval+ Scattersearch 的方法进行约简和选择。Cfs(correlation-based feature selector) 是一种基于相关性的启发式评价函数来选择属性子集的过滤算法^[10]。CfsSubsetEval 则是基于该属性子集评价方法^[11]。Merit 是根据属性间相关测度对子集进行评价的函数, 用于属性空间的搜索和属性子集的选择。Scattersearch 算法的基本计算过程是, 首先生成多样性初始解, 并用新解改进法对每一个解进行改进, 加入初始种群 P ; 根据解的质量和多样性兼顾原则, 从 P 选择若干个解构成初始参考集; 利用子集生成法从参考级中系统化生成一系列子集, 对这些子集中的解利用组合法策略地生成新的解; 利用新解改进法对该新解进行改进, 并利用该解对参考集进行更新, 反复执行上述过程, 直到满足结束准则^[12, 13]。

1.6 模式识别方法

本文采用粗糙集理论进行模式识别, 该方法在不需任何

何初始或附加信息的情况下就可对不精确、不一致、不完整的数据直接进行分析处理, 发现数据之间的关系, 提取出有用特征, 得到简明扼要的知识表达形式^[14, 15]。

在粗糙集中所研究对象的集合称作论域(U), 对论域中对象的分类就是建立等价关系, 也称属性集合(R), V 是属性值的集合, $f: U \times R \rightarrow V$ 是一个函数, 它指定了 U 中每一个对象的属性值, $\langle U, R, V, f \rangle$ 为信息表知识表达系统(S)。本文模式识别的过程即为建立水稻叶片属性信息系统 $\langle U, R, V, f \rangle$ 的过程。

由于不同特征属性间存在差异, 在进行模式识别前, 需对各属性值分别进行归一化处理消除差异。之后将属性值用 Entropy/MDL 方法离散化处理, 并用 CfsSubsetEval+ Scattersearch 方法进行属性约简, 删除决策表中不相容属性信息, 根据约简结果生成分类规则, 最后用检验数据对分类规则进行精度检验。

2 结果与分析

2.1 单因子特征分析

根据样本图像提取水稻叶片纹理、颜色、形状特征，然后对特征指数进行归一化处理，统计结果见图 1，极度缺磷、微量缺磷、正常磷的三个水平分别以 P1、P2 和 P3 表示。不同磷素水平和不同位置的水稻叶片样本特征之间存在差异。

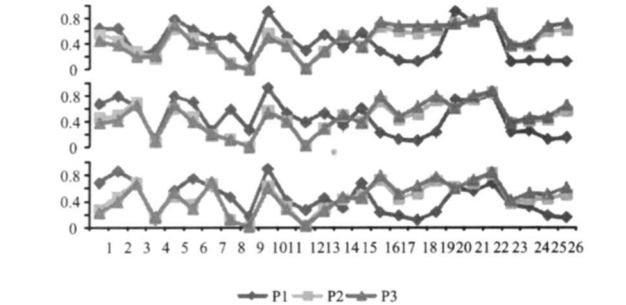


Fig 1 Analysis of different feature indices
1: R; 2: G; 3: B; 4: H; 5: S; 6: I; 7: G/B; 8: 均值; 9: 方差; 10: 倾斜度; 11: 峭度; 12: 能量; 13: 对比度; 14: 二阶矩; 15: 熵; 16: 长轴; 17: 短轴; 18: 面积; 19: 周长; 20: 偏心率; 21: 矩形度; 22: 面积凹凸比; 23: 圆形度; 24: 形状参数; 25: 周长圆形度; 26: 面积/周长

各类别的指数中，形状指数几乎对所有磷素水平和叶位的叶片都有更好的区分度；纹理指数对于区分 P1 水平作用较为明显。各磷素水平的水稻叶片中，P1 水平在大部分形状指数，部分纹理和颜色指数方面差异明显，说明严重缺磷的水稻叶片最容易区分；P2 与 P3 水平的区分度较小，但可作为区分两个磷素水平的依据。各叶位的水稻叶片中，R、G 和 I 颜色指数对三个水平都有一定的区分度，这可能会对区分三个水平起到重要作用。此外第三叶在形状特征方面的与其他两个位置的叶片差异较大，具体表现在偏心率、矩形度、面积凹凸比等指数方面。

颜色、纹理和形状指数能够区分不同叶位和不同磷素水平的水稻叶片，但是由于因子众多，很难将冗余的指数去除来归纳对每个叶位的叶片最敏感的指数。而且各指数之间可能存在相互内在联系，不能根据单个指数效果的优劣而随意删减指数。因此，需要结合有效的指数约简方法选择最有效的指数集合，以达到对磷素缺失程度进行分类，进而进行评价的目的。

2.2 特征组合的优化与选择

利用 CfsSubsetEval+ Scattersearch 方法对特征选择在本质上是一个重新组合优化的问题，在原有多种特征组合中根据兼顾解的质量和多样性的原则选择出较优的几个特征子集（表 2）。表 2 中子集的特征指数以数字代替，具体参照图 1。

从特征属性选择来看，颜色、纹理和形状指数对特征子集的选择都起到一定作用。第一叶所选特征子集中各类指数的比例较为均衡，经优化的第二叶特征子集主要通过颜色和形状指数反映出来；第三叶特征子集中则是颜色和纹理指数的作用明显大于形状指数，说明第三叶在光学和光谱特性

方面较其他叶位的叶片更为明显。

2.3 模式识别与精度评价

根据特征属性选择的结果，本文利用粗糙集的数据分析方法对不同磷素营养水平的水稻叶片图像特征数据进行模式识别，分为三种模式：极缺、微缺、正常。扫描数据分成两部分，一部分作为训练样本，另外一部分作为测试样本，用来对生成的识别规则进行检验。识别精度见表 3。

Table 2 Selected feature subset of rice leaves indifferent position					
第一叶		第二叶		第三叶	
Merit	子集	Merit	子集	Merit	子集
0 637 5	[10]	0 793 7	[2, 6, 14]	0 728 4	[4, 10]
0 616 3	[10, 13, 14]	0 792 9	[2, 6, 12, 14]	0 680 5	[1, 3, 4, 10]
0 563 1		0 777 1	[2, 13, 14, 15]	0 679 3	[4]
0 446 9		0 764 2	[12, 14]	0 589 4	[1, 4, 5, 15]
0 422 3	[8, 12-14]	0 750 1	[2, 12, 13]	0 456 9	[1, 3]
0 392 5	[2, 5, 6, 15]	0 727 6	[2, 15]	0 420 3	[6, 7]

Table 3 Accuracy of pattern recognition			
	P1/ %	P2/ %	P3/ %
第一叶	100	58.8	41.2
第二叶	100	80.0	68.4
第三叶	100	82.3	73.7

表 3 中识别精度的结果显示：不同叶位，极度缺磷状态的叶片识别率均为 100%，微缺和正常叶片的识别率从第一叶至第三叶的顺序递增，因此，严重缺磷时，顶部三张叶片的表现症状均可作为诊断依据，也就是说，三个叶位的叶片均能识别出极度缺磷的状态，而微量缺磷或者正常状态时，用高叶位的叶片进行识别的准确率较高；从不同缺磷水平来看，三个叶位的识别率都是从极缺到正常状态依次降低，由此可知，缺磷程度越严重，各个叶位的症状表现也越明显，而对微量缺磷和正常状态的水稻叶片，识别率也越低。

3 结果与讨论

本文以不同磷营养水平下水稻顶部前三张完全展开叶为研究对象，综合叶片图像的光谱信息，提取缺磷及正常营养状态下叶片特征，分析不同磷素水平下，叶片颜色、纹理和形状特征的敏感性差异，并结合 CfsSubsetEval+ Scattersearch 方法对 26 种特征进行优化组合和选择。由结果可知不同颜色、纹理和形状因子都对特征子集的选择提供了重要依据。

根据不同叶位的不同特征组合，利用粗糙集对不同磷营养模式进行识别，并对识别结果进行精度评价，识别模式为三种：极缺、微缺、正常。由精度评价的结果可以看出严重缺磷样本的识别精度比较高，而随着磷营养水平的升高，识别精度降低，这与 Osborne 在玉米磷营养诊断中的结论存在一致性^[15]。从不同叶位的识别精度来看，三个叶位对严重缺磷的样本识别率都比较高，而综合三种不同磷营养模式的识别精度，第三叶应该为水稻磷营养诊断的最佳叶位，这一结论也符合磷在作物体内从老叶到新叶转移的生理特性^[16]。

References

- [1] SHEN Ping-song, ZHANG Ding-hua(沈乒松, 张鼎华). Journal of Fujian Forestry Science and Technology(福建林业科技), 2005, 32(1): 76.
- [2] LI Qing-kui, ZHU Zhao-liang, YU Tian-ren(李庆逵, 朱兆良, 于天仁). Fertilizer Problems in Agricultural Sustainable Development in China(中国农业持续发展中的肥料问题). Nanchang: Science and Technology Press of Jiangxi(南昌: 江西科学技术出版社), 1988. 52.
- [3] WANG Cui-lian(汪翠莲). Serves of Agricultural Technology(农技服务), 2008, 25(7): 61.
- [4] Graeff S, Schepers J S, Schubert S. Journal of Plant Nutrition and Soil Science, 2001, 164: 445.
- [5] Jia L, Chen X, Zhang F, et al. Journal of Plant Nutrition, 2004, 27(3): 441.
- [6] XIAO Yan-bo, JIA Liang-liang, CHEN Xin-ping, et al(肖焱波, 贾良良, 陈新平, 等). Chinese Agricultural Science Bulletin(中国农学通报), 2008, 24(8): 448.
- [7] WANG Juan, LEI Yong-wen, ZHANG Yong-shuai, et al(王娟, 雷咏雯, 张永帅, 等). Chinese Journal of Eco Agriculture(中国生态农业学报), 2008, 16(1): 145.
- [8] Shettigara V K. Photogrammetric Engineering and Remote Sensing, 1992, 58(5): 561.
- [9] WANG Xiao-feng, HUANG De-shuang, DU Ji-xiang(王晓峰, 黄德双, 杜吉祥, 等). Computer Engineering and Applications(计算机工程与应用), 2006, 42(3): 190.
- [10] Mark A. Hall Correlation-Based Feature Selection for Machine Learning. The University of was Kato, 1999. 51.
- [11] Vasconcelos J A, Maeiel J H R D, Parreiras R O. IEEE Transactions on Magnetics, 2005, 4(5): 1804.
- [12] Herrem F, Lozano M, Molina D. European Journal of Operational Research, 2006, 16(9): 450.
- [13] Glover F, Laguna M, Marta R. Control and Cybernetics, 2000, 39(3): 653.
- [14] Pawlak Z, Skowron A. Information Sciences, 2007, 177(1): 3.
- [15] Skowron A. Rough Sets and Boolean Reasoning. New York: Physica Verlag, 2001. 95.
- [16] Osborne S L, Schepers J S, Francis D D, et al. Agronomy Journal, 2002, 94: 1215.

Diagnosis Study of Rice Leaf under Phosphorus Insufficiency Based on Spectral Features of Scan Image and Pattern Recognition

DING Xiao-dong, SHI Yuan-yuan, LU Xue, DENG Jin-song, SHEN Zhang-quan, WANG Ke*

Zhejiang Key Laboratory of Remote Sensing & Information Technique, Ministry of Education Key Laboratory of Environmental Remediation, Ecological and Health, Institute of Agricultural Remote Sensing & Information Technique, Zhejiang University, Hangzhou 310029, China

Abstract Insufficiency of phosphorus could greatly effect rice production, thus it is significant to adopt quick and nondestructive diagnosis of phosphorus content. The present paper focused on first expanded leaves with different phosphorus fertilization levels, comprehensively extracted 26 features' spectral information such as color, texture and shape etc. Single feature index analysis was conducted. Then features were collected to integrate CfsSubsetEval+ Scattersearch method for optimizing, evaluation and choosing. Based on the feature selection for different leave positions, leaves in different phosphorus fertilization levels were finally classified into three grades (extremely insufficient, significant insufficient and normal) according to rough set theory. Results showed that the accuracy of recognition was very high while few phosphorus contained in the leaves. Moreover, the third expanded leaf is the best part for phosphorus-nutrient diagnosis.

Keywords Scan image; Spectral feature; Phosphorus insufficiency; Diagnosis

(Received Jul. 19, 2010; accepted Oct. 8, 2010)

* Corresponding author