

玉米叶片铜铅污染元素种类光谱判别的 EC-PB 规则

吴 兵, 杨可明*, 高 伟, 李艳茹, 韩倩倩, 张建红

中国矿业大学(北京)地球科学与测绘工程学院, 北京 100083

摘 要 随着人类生活质量的提高,农产品重金属污染问题备受关注。农作物中的重金属元素会通过食物链侵害人体健康,而不同重金属元素对人体毒害差别较大,因此农作物中含有重金属元素的类别识别至关重要。传统重金属元素检测方法存在环节多、耗时长、成本高等缺点,但高光谱遥感技术具有信息使用量大,理化反演能力强,分析速度快,无损监测等优势,逐渐成为农作物重金属污染分析的重要手段之一。以不同 $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ 和 $\text{Pb}(\text{NO}_3)_2$ 浓度梯度土壤胁迫下典型农作物玉米生长的叶片光谱为研究对象,引入光谱包络线去除(CR)、光谱比值(SR)、分数阶微分(FOD)同时结合改进红边比值指数(MSR)构建铜铅元素识别指数(CLI);通过挑选与铜铅元素种类相关性最强的三个分数阶微分阶数的 CLI 值建立铜铅元素判别特征点(CLDFP);再利用欧式聚类(EC)将训练集样本分为铜污染与铅污染两类并结合圆心连线的垂直平分线(PB),建立基于 EC-PB 识别铜铅元素种类的二维坐标系下判别规则线(CLDRL)和三维坐标系下判别规则面(CLDRP),从而实现玉米叶片光谱信息重金属铜铅元素种类准确识别。研究表明,CR-SR-FOD 光谱变换处理增加了玉米叶片光谱信息与铜铅元素种类之间的相关性;各阶次 FOD 对应的 CLI 与铜铅元素种类相关系数各不相同,随着阶次的增加,相关性呈现先递增后递减的趋势,其中相关系数最高的三个阶次分别为 1.2 阶,0.7 阶,1.0 阶;在二维坐标系下训练集样本判别正确率为 78.95%,验证集样本判别正确率为 75.0%;在三维坐标系下训练集样本判别正确率为 76.32%,验证集样本判别正确率为 75.0%,证明了基于 EC-PB 构建的二维 CLDRL 和三维 CLDRP 光谱判别规则可以有效识别玉米叶片中铜铅污染元素种类。

关键词 光谱分析;玉米叶片;光谱变换;重金属元素识别;欧式聚类;垂直平分线

中图分类号: TP75 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2022)10-3256-07

引 言

工业化和城市化发展以及一些自然或人为干扰,使人类赖以生存的土壤中重金属浓度逐渐增大,当土壤中重金属含量超过了当地环境背景值时,其累积效应会影响土地承载力,最终导致土壤重金属污染^[1-2],引发生态环境恶化和粮食安全危机。重金属污染物带有毒性并且不可降解,一旦通过生态系统或食物链进入人体,就会对人类身体健康构成严重威胁^[3-4]。传统的重金属污染检测方法通常是对野外采集的样品开展实验室化学分析,该方法测量精度高、准确性强,但由于检测环节多、耗时长、成本高,很难快速获取大区域的污染物含量及分布信息^[5]。快速判别重金属污染的元素种类以及污染程度,为污染区的针对性治理计划及时提供依据,成为当前亟待解决的关键步骤^[6-7]。

由于高光谱遥感具有光谱信息量丰富、技术成本低、理化反演能力强,分析操作实时便捷、可非接触式无损检测等优点^[8],从而如何将光谱分析技术应用于重金属污染检测已成为了现今遥感领域研究的一个热点,并取得了一定进展。Hou 等利用可见-近红外光谱数据分析山东省邹城市某煤矿周边的土壤重金属浓度,发现利用光谱数据建立的偏最小二乘回归模型预测土壤中 Cu 和 Pb 重金属浓度精度较高^[9]。也有一些学者研究表明,土壤中的重金属能对植物的生理结构特征产生影响,当土壤中的重金属含量过高时,会导致植物叶片变黄、分枝、落叶甚至死亡^[10],从而改变植物的光谱特性。在农作物重金属污染方面,杨可明等利用玉米叶片光谱的分数阶微分(fractional order differential, FOD)信息构建玉米叶片光谱红边位置的铜铅敏感指数集群,实现铜铅胁迫识别^[11];黄钟霆等以湖南省某典型关停锰矿区为研究对象,采集矿区周边的农作物及其对应的土壤样品,探讨土壤及对应

收稿日期: 2021-08-22, 修订日期: 2022-01-21

基金项目: 国家自然科学基金项目(41971401), 中央高校基本科研业务费专项资金(2021YJSDC17)资助

作者简介: 吴 兵, 1993 年生, 中国矿业大学(北京)地球科学与测绘工程学院博士研究生 e-mail: wubing9362@163.com

* 通讯作者 e-mail: ykm69@163.com

农作物间重金属迁移规律, 并进行生态风险评价^[12]。上述研究在重金属元素含量预测以及识别方面均得到了较好的结果, 但光谱处理繁琐, 算法较为复杂, 计算量大, 不易复现。

以不同铜铅离子浓度梯度胁迫生长下的玉米叶片光谱为研究对象, 使用 ASD FieldSpec 4 便携式地物光谱仪获取不同胁迫梯度下穗期玉米叶片光谱数据, 基于常规的包络线去除(continuum removed, CR), 光谱比值(spectral ratio, SR)以及分数阶微分(FOD)光谱预处理方法构建铜铅元素识别指数(copper and lead identification index, CLI), 联合欧式聚类(euclidean cluster, EC)与垂直平分线(perpendicular bisector, PB)法建立二维(2D)与三维(3D)铜、铅污染元素判别规则, 实现依据被污染玉米叶片光谱的铜铅元素准确识别, 为重金属元素快速识别, 污染区域治理提供新思路。

1 实验部分

1.1 试验设计

玉米是我国重要的农作物之一, 也是我国种植面积最大的农作物。为了探究重金属对农作物的影响, 试验以盆栽种植的玉米植株为研究对象, 设置 $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ 和 $\text{Pb}(\text{NO}_3)_2$ 含量分别为 0, 50, 100, 150, 200, 300, 400, 600, 800 和 1 000 $\mu\text{g} \cdot \text{g}^{-1}$ 的污染胁迫梯度土壤用于培育植株生长, 培育场所室温可控, 水气环境适宜, 营养元素充足, 不同重金属胁迫梯度设置 3 个平行试验样本组。光谱数据采集在暗室进行, 使用波段范围为 350~2 500 nm 的 ASD FieldSpec 4 便携式地物光谱仪获取不同胁迫梯度下穗期玉米叶片光谱数据, 每个样本测量 3 次光谱后剔除异常信息取平均光谱, 使用电感耦合等离子发射光谱仪测定玉米叶片 Cu^{2+} 和 Pb^{2+} 含量(如表 1 所示)并用于研究结果验证。

表 1 不同浓度梯度下玉米叶片中 Cu^{2+} , Pb^{2+} 含量
Table 1 Cu^{2+} and Pb^{2+} contents in corn leaves under different concentration gradients

土壤胁迫梯度/ ($\mu\text{g} \cdot \text{g}^{-1}$)	叶片 Cu^{2+} 含量/ ($\mu\text{g} \cdot \text{g}^{-1}$)	叶片 Pb^{2+} 含量/ ($\mu\text{g} \cdot \text{g}^{-1}$)
50	5.92	2.28
100	6.56	3.60
150	8.30	5.70
200	9.26	0.28
300	8.76	0.46
400	9.19	11.18
600	12.80	17.36
800	14.17	32.16
1 000	16.59	33.24

本试验共设置 9 个不同铜铅浓度梯度胁迫下生长的玉米植株, 每个浓度梯度设置 3 个平行组, 共 54 组样本数据, 由 Python 中的 sklearn 库随机划分样本数据的 70%(38 组)为训练集数据, 30%(16 组)为验证集数据。

1.2 光谱变换

(1)包络线去除(CR)。CR 也称连续统去除, 是一种突出光谱曲线的吸收和反射特征, 并将反射率归一化的光谱预处理方法。根据 Clark 提出的外壳系数法^[13], 通过计算光谱曲线上的最大极大值点作为包络线的一个端点, 计算该点与波长增长方向各个极大值点连线的斜率, 以斜率最大点作为包络线的下一个端点, 以此循环, 将所有端点连接, 形成光谱曲线的包络线, 用实际光谱曲线的反射率值除以包络线上相应波段的反射率值即可得到原始光谱曲线去除包络线后的结果, 如式(1)和式(2)所示。

$$S_i = \frac{\rho_i}{\rho_s + K(\lambda_i - \lambda_s)} \quad (1)$$

$$K = \frac{\rho_e - \rho_s}{\lambda_e - \lambda_s} \quad (2)$$

式中, S_i 为波段 i 的包络线去除值, K 为极大值始点和端点之间的斜率, ρ_i 为波段 i 的原始光谱反射率, ρ_s 和 ρ_e 为极大值始点和端点的原始光谱反射率, λ_s 和 λ_e 为 ρ_s 和 ρ_e 对应的波段值。本研究中所有玉米叶片光谱数据均进行了 CR 处理。

(2)光谱比值(SR)。SR 最早应用于消除遥感图像中的阴影。随着高光谱技术的成熟, 比值技术在混合高光谱信号解混中表现出巨大潜力, 其核心思想是通过相同波段不同光谱数据反射率的比值进行 SR 变换, SR 技术可以抑制作为分母的光谱特征, 而突出分子光谱的影响。本研究中为了突出重金属元素铜铅的光谱特征, 以相同环境下受不同程度铜铅污染的玉米叶片光谱作为分子, 无污染的正常的玉米叶片光谱作为分母进行 SR 处理。

(3)分数阶微分(FOD)。光谱微分可以消除光谱数据之间的系统误差, 减弱大气辐射、散射和大气吸收对待测物体光谱的影响, 突出待测光谱曲线特征吸收峰和斜率的微小变化, 光谱的 FOD 是整数阶微分的推广, 根据 Grünwald-Letnikov 定义形式^[14], FOD 计算方式如式(3)所示。

$$\frac{d^q}{d\lambda^q} f(\lambda) = \lim_{h \rightarrow 0} \frac{1}{h^q} \sum_{a=0}^{\lfloor \frac{n-m}{h} \rfloor} (-1)^m \cdot \frac{\Gamma(q+1)}{\Gamma(a+1)\Gamma(q-a+1)} f(\lambda - ah) \quad (3)$$

式(3)中, λ 表示光谱曲线某一波段值, $[m, n]$ 为波段区间, $\lambda \in [m, n]$, h 为光谱采样间隔, a 为常数参数, q 为任意阶数, Γ 为 Gamma 函数, 其积分定义如式(4)所示。

$$\Gamma(z) = \int_0^{\infty} e^{-t} t^{z-1} dt \quad (4)$$

本研究中对经过 CR、SR 处理后的玉米叶片光谱按照 0.1 的间隔进行 0~2 的各阶次 FOD 处理。

1.3 铜铅污染元素判别特征点(CLDLP)

作物中所含的重金属元素属于微量元素, 在光谱曲线中即使经过了增强处理也很难仅通过光谱曲线进行重金属元素识别, 因此需要建立一种判别规则对作物样本中的重金属元素进行准确的识别。本研究基于改进红边比值指数(modified red edge simple ratio index, MSR)^[15]建立一种铜铅元素识别指数(CLI)如式(5)所示。

$$\text{CLI} = \frac{\rho(\lambda_a)/\rho(\lambda_b) - 1}{\sqrt{\rho(\lambda_a)/\rho(\lambda_b) + 1}} \quad (5)$$

式(5)中, $\rho(\lambda_a)$ 和 $\rho(\lambda_b)$ 分别表示波长为 λ_a 和 λ_b 处的光谱值。本研究中将经过 CR-SR-FOD 处理后的各组玉米叶片光谱数据入 CLI 中, 挑选出与铜铅元素种类相关性最强的三个阶数的 CLI, 分别为 CLI1, CLI2 和 CLI3。以 CLI1 和 CLI2 为坐标分量构建 2D 坐标系, (CLI1 为 x 轴, CLI2 为 y 轴); 以 CLI1, CLI2 和 CLI3 为坐标分量构建 3D 坐标系, (CLI1 为 x 轴, CLI2 为 y 轴, CLI3 为 z 轴) 构建铜铅元素判别特征点 (copper and lead discriminant feature points, CLDFP), 如式(6)和式(7)所示。

$$\text{CLDFP}_{2D} = (\text{CLI1}, \text{CLI2}) \quad (6)$$

$$\text{CLDFP}_{3D} = (\text{CLI1}, \text{CLI2}, \text{CLI3}) \quad (7)$$

1.4 铜铅污染元素判别规则

欧式聚类 (EC)^[16] 即基于欧几里得距离的一种聚类方法, 针对 CLDFP 需要确定不同元素各样本点内最大的两个欧式距离 d_1 和 d_2 , 2D 与 3D 坐标系下的欧式距离公式如式(8)和式(9)所示。

$$d_{2D} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (8)$$

式(8)中, x_1, y_1, x_2 和 y_2 分别为 2D 坐标系下不同污染元素各样本点内最大欧式距离的 CLI 值。

$$d_{3D} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \quad (9)$$

式(9)中, x_1, y_1, z_1, x_2, y_2 和 z_2 分别为 3D 坐标系下不同污染元素各样本点内最大欧式距离的 CLI 值。

在 2D 坐标系中, 以最大欧式距离 d_{2D1} 和 d_{2D2} 为直径作圆(圆心为两点的中点), 将训练集样本分为铜污染和铅污染两类, 连接两圆圆心, 以圆心连线的垂直平分线 (PB) 作为铜铅元素的判别规则线 (copper and lead discriminant rule lines, CLDRL_{2D}); 同理, 在 3D 坐标系中, 以最大欧式距离 d_{3D1} 和 d_{3D2} 为直径作球(球心为两点的中点), 将训练集样本分为铜污染和铅污染两类, 连接两球球心, 在 3D 坐标系中将垂直平分线进一步推广为垂直平分面, 以球心连线的垂直平分面作为铜铅元素的判别规则面 (copper and lead discriminant rule planes, CLDRP_{3D}) 实现玉米叶片光谱重金属铜铅元素种类的准确识别。

2 结果与讨论

2.1 光谱变换处理与相关性分析

不同胁迫梯度下的玉米叶片原始光谱数据如图 1 所示。通过图 1 可以看出不同胁迫梯度的玉米叶片原始光谱曲线均表现出典型植被光谱特征, 不同 Cu^{2+} 和 Pb^{2+} 浓度胁迫下仅在反射率中表现出部分差异, 规律性不强, 很难通过原始数据实现铜铅元素类别的区分, 因此需要进行不同的光谱变换处理增强光谱特征以及不同污染元素光谱曲线之间的差异性。

以土壤中 $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ 含量为 $50 \mu\text{g} \cdot \text{g}^{-1}$ 下培育的玉米植株叶片光谱数据为例进行不同光谱变换处理, 结果如图 2 所示。通过图 2 可以看出经过包络线去除 (CR) 处理后的光谱曲线峰谷特征表现的更加明显; 光谱比值 (SR) 处理后的光谱曲线已无植被的光谱特征, 进一步突出了铜元素的光谱特

征; 1.2 阶分数阶微分 (FOD) 处理后的光谱曲线消除了光谱数据中的无用信息, 使得光谱数据中的有用信息能够更有效的被利用。

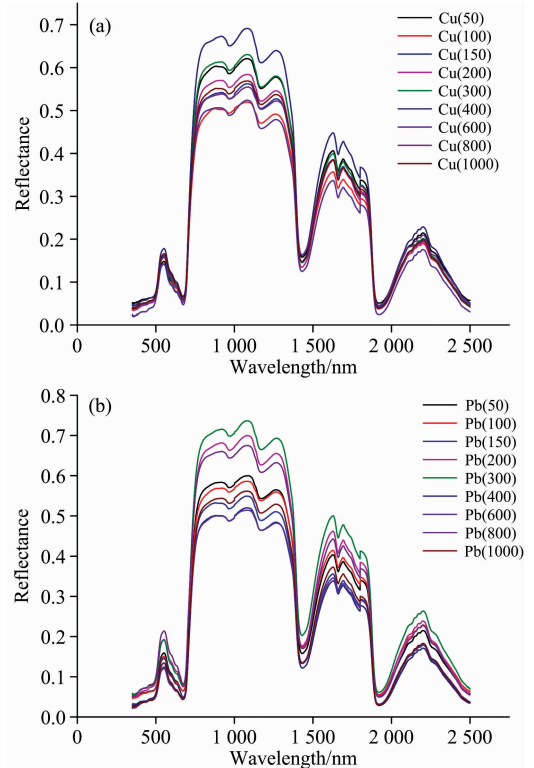


图 1 不同胁迫梯度下的玉米叶片原始光谱数据

(a): 不同 Cu^{2+} 浓度胁迫下玉米叶片原始光谱曲线;
(b): 不同 Pb^{2+} 浓度胁迫下玉米叶片原始光谱曲线

Fig. 1 Original spectral data of corn leaves under different stress gradients

(a): Original spectral curve of corn leaves under different Cu^{2+} concentration stress; (b): Original spectral curve of corn leaves under different Pb^{2+} concentration stress

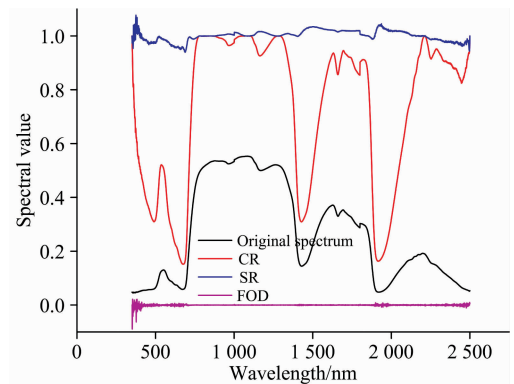


图 2 不同光谱变换处理的光谱

Fig. 2 Spectra processed by different spectral conversions

相关性可以反映数据间的关联程度。通过计算不同光谱变换处理后各波段光谱与叶片中重金属铜铅元素种类之间的相关系数, 反映光谱信息与铜铅元素种类的关联性。分析得

出：不作任何变形处理的原始光谱数据与叶片中铜铅元素种类之间相关系数绝对值的最大值为 0.21，平均值为 0.09；而 CR 处理后的光谱信息与叶片中铜铅元素种类之间相关系数绝对值的最大值为 0.51，平均值为 0.17；SR 是在 CR 基础上的变换，其相关性与 CR 相同；0.1~2.0 各阶次 FOD 处理后的光谱信息与叶片中铜铅元素种类之间的相关系数如表 2 所示。

通过观察不同光谱变换处理后各波段光谱数据与叶片中铜铅元素种类之间的相关系数，可以发现经过光谱变换处理后相关系数最大值、平均值与原始数据相比均有较为明显的增加，说明光谱变换处理增加了光谱数据与铜铅元素种类之间的关联程度；不同阶次 FOD 处理后相关系数各不相同，为了更好地区分铜铅元素种类，需挑选出相关性较高的阶次用于后续研究。

2.2 建立与挑选 CLI

训练集样本数据经过 CR、SR 以及各阶次 FOD 处理后代入式(5)计算铜铅元素识别指数 CLI 并计算 0.1~2.0 各阶次 FOD 对应的 CLI 与重金属铜铅元素种类之间的相关性，相关系数随波长变化情况如图 3 所示。

表 2 各阶次 FOD 与重金属铜铅元素种类相关系数

Table 2 Correlation coefficients between FOD of each order and heavy metal copper-lead element types

FOD 阶次	相关系数绝对值最大值	相关系数绝对值平均值
0.1	0.399 7	0.193 5
0.2	0.425 1	0.194 0
0.3	0.441 9	0.193 0
0.4	0.462 8	0.190 0
0.5	0.458 6	0.184 5
0.6	0.406 1	0.177 0
0.7	0.411 1	0.169 3
0.8	0.450 3	0.160 9
0.9	0.481 7	0.146 3
1.0	0.492 5	0.133 5
1.1	0.476 4	0.128 0
1.2	0.474 2	0.123 7
1.3	0.473 2	0.120 7
1.4	0.470 5	0.118 3
1.5	0.466 4	0.116 6
1.6	0.460 7	0.115 5
1.7	0.453 4	0.114 8
1.8	0.453 9	0.114 4
1.9	0.453 8	0.114 0
2.0	0.448 0	0.113 7

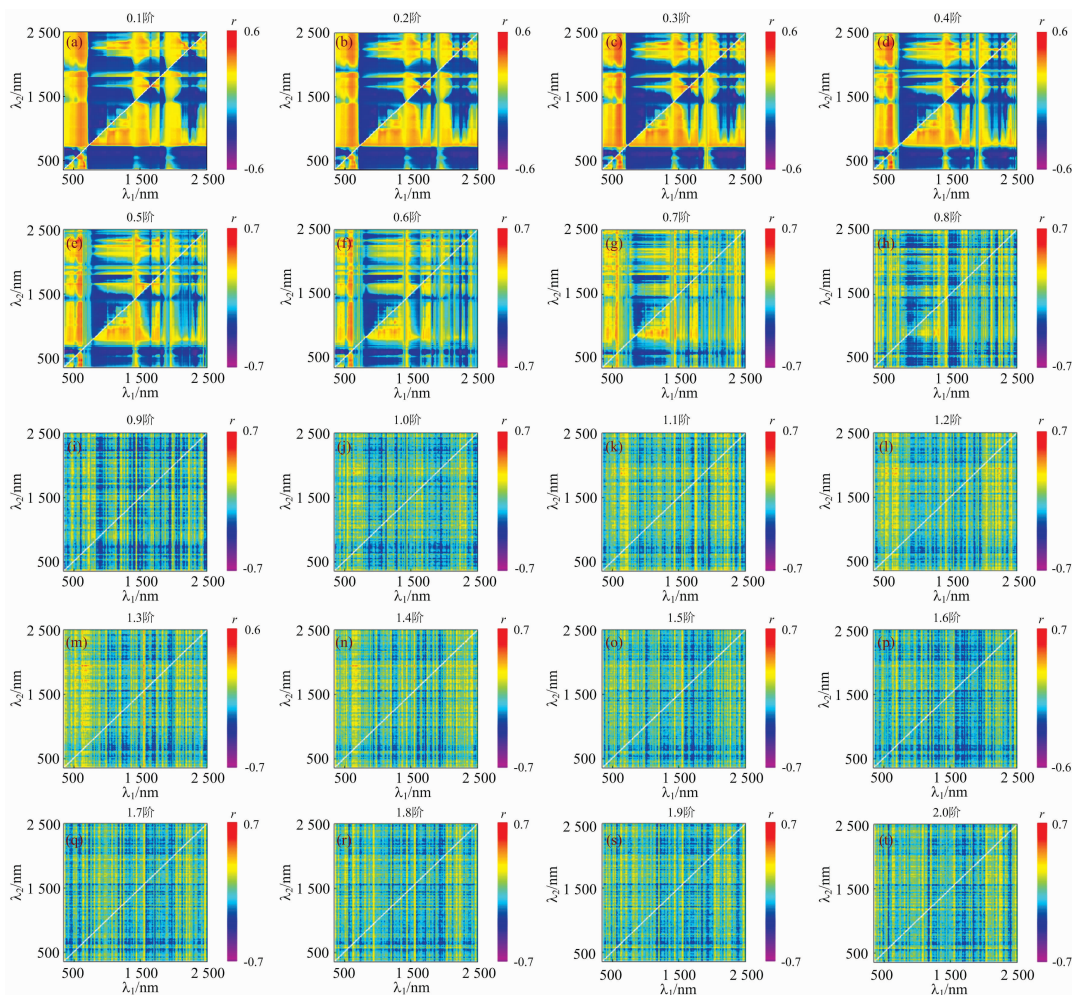


图 3 各阶次 CLI 与铜铅元素种类相关性

Fig. 3 Correlation between CLI and Cu-Pb element types of each order

其中, 0.1—0.4 阶 FOD 对应的 CLI 与铜铅元素种类相关系数值在 $-0.60 \sim 0.63$ 之间; 0.5—0.7 阶 FOD 对应的 CLI 与铜铅元素种类相关系数值在 $-0.70 \sim 0.69$ 之间; 0.8—0.9 阶 FOD 对应的 CLI 与铜铅元素种类相关系数值在 $-0.67 \sim 0.66$ 之间; 1.0—1.2 阶 FOD 对应的 CLI 与铜铅元素种类相关系数值在 $-0.72 \sim 0.70$ 之间; 1.3—1.5 阶 FOD 对应的 CLI 与铜铅元素种类相关系数值在 $-0.67 \sim 0.69$ 之间; 1.6—2.0 阶 FOD 对应的 CLI 与铜铅元素种类相关系数值在 $-0.66 \sim 0.67$ 之间。通过各阶次 FOD 对应的 CLI 与铜铅元素种类相关系数可以发现, 随着阶次的增加, 相关性呈现先递增, 后递减的趋势, 其中相关系数最高的三个阶次分别为 1.2 阶、0.7 阶、1.0 阶。

以 1.2 阶 CLI 作为 CLI1, 其构成波段为 639 nm(λ_a) 和 1 150 nm(λ_b), 与铜铅元素种类相关系数为 -0.719 ; 以 0.7 阶 CLI 作为 CLI2, 其构成波段为 567 nm(λ_a) 和 393 nm(λ_b), 与铜铅元素种类相关系数为 -0.701 ; 以 1.0 阶 CLI 作为 CLI3, 其构成波段为 2 009 nm(λ_a) 和 636 nm(λ_b), 与铜铅元素种类相关系数为 0.700 。

2.3 识别效果与验证

(1) 铜铅污染元素种类的判别规则线 (CLDRL)

分别将 CLI1 和 CLI2 对应的训练集样本值带入 2D 坐标系, 形成二维铜铅元素判别特征点 CLDFP_{2D}, 逐点寻找不同元素各样本点内最大两点的欧式距离 d_{2D1} 和 d_{2D2} 进行铜污染与铅污染的分类, 其中铜污染判别特征点间最大距离点为 (1.267, 0.701) 与 (-0.527, 0.329), 两点间欧式距离 d_{2D1} 为 1.833, 即以圆心为 (0.370, 0.515) (最大距离两点中点), 半径为 0.916 (最大距离的二分之一) 作铜污染判别特征类; 铅污染判别特征点间最大距离点为 (-0.424, 1.665) 与 (2.970, 1.457), 两点间欧式距离 d_{2D2} 为 3.400, 即以圆心为 (1.273, 1.561) (最大距离两点中点), 半径为 1.700 (最大距离的二分之一) 作铅污染判别特征类, 连接两圆圆心, 圆心连线的垂直平分线即为 CLDRL_{2D}, 其表达式如式 (10) 所示, 2D 坐标系下铜铅判别特征点与判别规则线分布情况如图 4 所示。

$$\text{CLDRL}_{2D}: y = -0.8630x + 1.7466 \quad (10)$$

通过图 4 可以明显看出 CLDFP_{2D} 根据两个判别特征类分成了不同的区域, 大部分 CLDFP_{2D} 可以由 CLDRL_{2D} 正确区分, 总训练集样本为 38 个, 判别正确样本为 30 个, 判别正确率达到了 78.95%。

将 CLI1 和 CLI2 对应的验证集样本值带入 2D 坐标系, 以相同的 CLDRL_{2D} 进行判别, 结果如图 5 所示, 总验证集样本为 16 个, 判别正确样本为 12 个, 判别正确率为 75.0%。

(2) 铜铅污染元素种类的判别规则面 (CLDRP)

为了增加样本的直观性, 对 2D 判别规则进一步推广, 将 CLI1, CLI2 和 CLI3 对应的训练集样本值带入 3D 坐标系, 构建三维铜铅元素判别特征点 CLDFP_{3D}, 逐点寻找不同元素各样本点内最大两点的欧式距离 d_{3D1} 和 d_{3D2} 进行铜污染与铅污染的分类, 其中铜污染判别特征点间最大距离点为 (0.169, 1.689, -0.831) 与 (0.447, 0.517, 2.266), 两点间欧式距离 d_{3D1} 为 3.323, 即以球心为 (0.308, 1.103, 0.717)

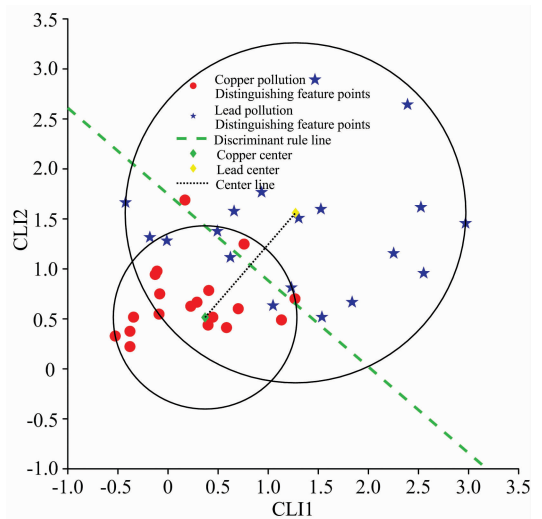


图 4 训练集中 2D 铜铅判别特征点与判别规则线分布
Fig. 4 The distribution of two-dimensional copper-lead discriminant feature points and discriminant rule lines in the training set

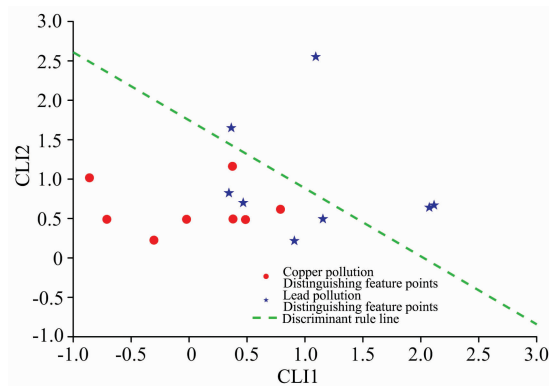


图 5 验证集中 2D 铜铅判别特征点与判别规则线分布
Fig. 5 The distribution of two-dimensional copper-lead discriminant feature points and discriminant rule lines in the verification set

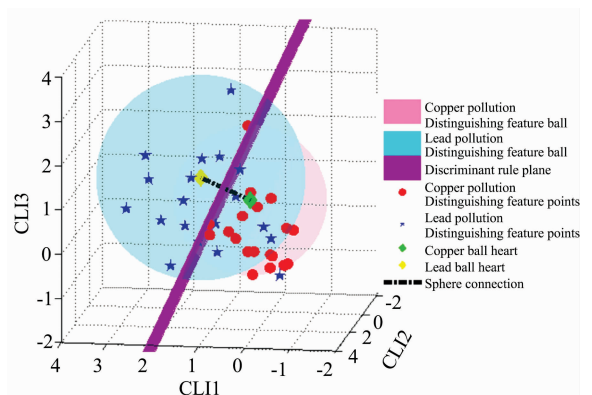


图 6 训练集中 3D 铜铅判别特征点与判别规则面分布
Fig. 6 The distribution of three-dimensional copper-lead discriminant feature points and discriminant rule planes in the training set

(最大距离两点中点), 半径为 1.662(最大距离的二分之一) 作铜污染判别特征类; 铅污染判别特征点间最大距离点为 (0.659, 1.578, 3.290) 与 (2.114, 0.670, -0.904), 两点间欧式距离 d_{3D2} 为 4.531, 即以球心为 (1.387, 1.124, 1.193) (最大距离两点中点), 半径为 2.266(最大距离的二分之一) 作铅污染判别特征类, 连接两球球心, 球心连线的垂直平分面即为 $CLDRP_{3D}$, 其表达式如式(11)所示, 3D 坐标系下铜铅判别特征点与判别规则面分布情况如图 6 所示。

$$CLDRP_{3D}: z = -2.2680x - 0.0430y + 2.9246 \quad (11)$$

通过图 6 可以发现 $CLDFP_{3D}$ 同样根据两个判别特征类形成了不同的区域, 判别效果更为直观, 由于新增加了一个维度, 误差也在相应累加, 总训练集样本为 38 个, 判别正确样本为 29 个, 判别正确率为 76.32%。

将 $CLI1$, $CLI2$ 和 $CLI3$ 对应的验证集样本值带入 3D 坐标系, 以相同的 $CLDRP_{3D}$ 进行判别, 结果如图 7 所示, 总验证集样本为 16 个, 判别正确样本为 12 个, 判别正确率为 75.0%。

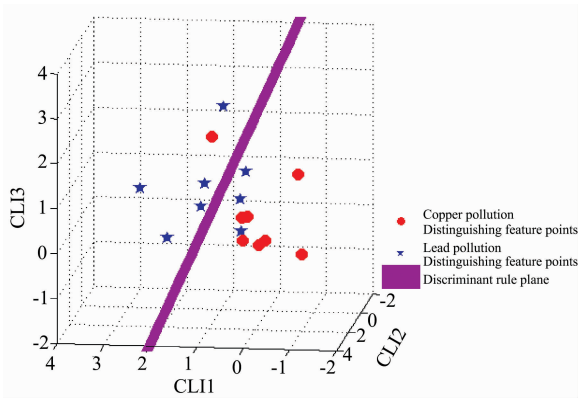


图 7 验证集中 3D 铜铅判别特征点与判别规则面分布
Fig. 7 The distribution of three-dimensional copper-lead discriminant feature points and discriminant rule plans in the verification set

References

- [1] Chakraborty S, Man T, Paulette L, et al. *Geoderma*, 2017, 306: 108.
- [2] Chakraborty S, Weindorf D C, Deb S, et al. *Geoderma*, 2017, 289: 72.
- [3] Gujre Nihal, Mitra Sudip, Soni Ankit, et al. *Chemosphere*, 2021, 262(5): 128013.
- [4] Fu Z S, Xi S H. *Toxicol. Mech. Methods*, 2020, 30(3): 167.
- [5] Fei X, Christakos G, Xiao R, et al. *Science of the Total Environment*, 2019, 661: 168.
- [6] Liu Y X, Sun X Y, Li S, et al. *Environ. Sci. Pollut. R*, 2020, 27(7): 7693.
- [7] Jin Z M, Deng S Q, Wen Y C, et al. *Sci. Total Environ.*, 2019, 697: 134148.
- [8] Fu P, Zhang W, Yang K M, et al. *Ecotoxicology and Environmental Safety*, 2020, 206: 111211.
- [9] Hou L, Li X, Li F. *Journal of Environment Quality*, 2019, 48(1): 57.
- [10] Jin M, Liu X N, Zhang B Y. *IEEE J-Stars*, 2017, 10(7): 3232.
- [11] YANG Ke-ming, GAO Wei, CHEN Gai-ying, et al(杨可明, 高伟, 陈改英, 等). *Transactions of the Chinese Society for Agricultural Machinery(农业机械学报)*, 2021, 52(6): 215.
- [12] HUANG Zhong-ting, YI Sheng-wei, CHEN Bei-bei, et al(黄钟霆, 易盛炜, 陈贝贝, 等). *Environmental Science(环境科学)*, 2022, 42(2): 975.
- [13] Huang Z, Turner B J, Dury S J, et al. *Remote Sensing of Environment*, 2004, 93(1-2): 18.

经 2D、3D 坐标系下训练集与验证集的判别结果可知, 通过光谱变换与 CLI 形成的 $CLDFP$ 包含了其所受铜铅元素种类相关的信息, 使得 EC-PB 法建立的 $CLDRL_{2D}$ 与 $CLDRP_{3D}$ 能够较好的区分铜铅污染下的玉米叶片光谱数据, 其中 2D 数据判别精度略高于 3D 数据; 3D 数据可从多个角度查看, 判别效果更为直观、易懂。

3 结 论

(1) 玉米叶片光谱数据中的铜铅元素污染信息难以通过原始光谱曲线进行准确判别, 原始光谱数据与铜铅元素种类之间的相关性较低; CR, SR 以及 FOD 光谱变换处理均增加了玉米叶片光谱数据与铜铅元素种类之间的相关性; 不同阶次 FOD 处理后的光谱数据与铜铅元素种类之间的相关性各不相同。

(2) 各阶次 FOD 对应的 CLI 与重金属铜铅元素种类相关系数各不相同, 随着阶次的增加, 相关性呈现先递增, 后递减的趋势, 其中相关系数最高的三个阶次分别为 1.2 阶, 0.7 阶, 1.0 阶, 其对应的相关系数为 -0.719, -0.701, 0.700。

(3) 基于 EC-PB 建立的 2D、3D 坐标系下的铜铅元素判别规则线 $CLDRL_{2D}$ 与铜铅元素判别规则面 $CLDRP_{3D}$ 能够较好地区分玉米叶片光谱数据中的铜铅污染元素种类, 在 2D 坐标系下训练集判别正确率为 78.95%, 验证集判别正确率为 75.0%; 在 3D 坐标系下训练集判别正确率为 76.32%, 验证集判别正确率为 75.0%。其中 2D 数据判别精度略高于 3D 数据; 3D 数据可从多个角度查看, 判别效果更为直观、易懂。

利用 EC-PB 能较为理想地取得玉米叶片污染的铜铅元素判别与识别效果, 同时也存在一定的局限性, 如模型的普适性、除铜铅外其他重金属元素的污染识别以及野外大面积农作物重金属污染元素种类识别还需要做进一步的研究。

- [14] Wang J Z, Tiyip T, Ding J L, et al. PLOS ONE, 2017, 12(9): e0184836.
- [15] Chen J M. Canadian Journal of Remote Sensing, 1996, 22(3): 229.
- [16] FAN Jing-jing, WANG Li, CHU Wen-bo, et al(范晶晶, 王力, 褚文博, 等). Automotive Engineering(汽车工程), 2019, (12): 1410.

EC-PB Rules for Spectral Discrimination of Copper and Lead Pollution Elements in Corn Leaves

WU Bing, YANG Ke-ming*, GAO Wei, LI Yan-ru, HAN Qian-qian, ZHANG Jian-hong

College of Geoscience and Surveying Engineering, China University of Mining & Technology (Beijing), Beijing 100083, China

Abstract Heavy metal pollution of agricultural products has attracted much attention along with the improved human quality of life. The heavy metal elements in crops will harm human health through the food chain, and different heavy metal elements have a large difference in toxicity to the human body. Therefore, it is crucial to distinguish the types of heavy metal elements in crops. There are many shortcomings in the traditional methods of detecting heavy metals such as many links, long time, and high cost. However, hyperspectral remote sensing technology has the advantages of abundant information usage, strong physical and chemical inversion capabilities, fast analysis speed, non-destructive monitoring and so on. It has gradually become one of the important methods for analysing heavy metal pollution in crops. Taking the leaf spectra of a typical corn crop growing under soil stressed by different $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ and $\text{Pb}(\text{NO}_3)_2$ concentration gradients as the research object, the copper (Cu) and lead (Pb) identification index (CLI) was built based on spectral processing results of continuum removal (CR), spectral ratio (SR) and fractional-order derivative (FOD) combining with modified red edge simple ratio index (MSR). Then the Cu and Pb element discrimination feature points (CLDFP) were established by selecting the three CLI values of fractional differential orders that have the strongest correlation with the types of Cu and Pb elements. And then, the Cu and Pb elements discriminant rule line (CLDRL) under the two-dimensional coordinate system (2D) and the discriminant rule plane (CLDRP) under the three-dimensional coordinate system (3D) were structured to identify the types of Cu and Pb elements. Based on the Euclidean cluster (EC)- the perpendicular bisector (PB) by using the EC to divide the training samples into two sets of Cu pollution and Pb pollution and combining with the PB to connect the circle enters the sets so that the types could be accurately identified on the heavy metal Cu and Pb elements in the spectral information of corn leaves. The results showed that the correlation between the spectral information of corn leaves and the types of Cu and Pb elements was enhanced because of the CR-SR-FOD spectral transformation processing. The correlation coefficients of the CLI corresponding to each order of FOD and the types of Cu and Pb elements were different. With the increase of orders, the correlation showed a trend of increasing first and then decreasing. Among them, the three values of orders of the highest correlation coefficients were 1.2, 0.7, and 1.0 respectively. The accuracy rate of the training set samples was 78.95% and the accuracy rate of the verification set samples was 75.0% when discriminated under the 2D, and the accuracy rate of the training set samples was 76.32% and the accuracy rate of the verification set samples was 75.0% when discriminated under the 3D, it is proved that the spectral discriminant rules of 2D CLDRL and 3D CLDRP based on EC-PB could effectively identify the types of Cu and Pb pollution elements when they polluted the corn leaves.

Keywords Spectral analysis; Corn leaf; Spectral transformation; Identification of heavy metal elements; Euclidean cluster; Perpendicular bisector

(Received Aug. 22, 2021; accepted Jan. 21, 2022)

* Corresponding author