

DOI: 10.11766/trxb202302070050

杨凯, 慈恩, 刘彬, 陈洋洋, 谢宇. 基于 K-means 聚类 and 图像分割的紫色土发生层边界识别[J]. 土壤学报, 2024, 61 (4): 939–951.

YANG Kai, CI En, LIU Bin, CHEN Yangyang, XIE Yu. Boundary Identification of Purple Soil Horizon Based on K-means Clustering and Image Segmentation[J]. Acta Pedologica Sinica, 2024, 61 (4): 939–951.

基于 K-means 聚类 and 图像分割的紫色土发生层边界识别^{*}

杨 凯¹, 慈 恩^{1, 2†}, 刘 彬¹, 陈洋洋¹, 谢 宇¹

(1. 西南大学资源环境学院, 重庆 400715; 2. 三峡库区农业面源污染控制工程技术研究中心, 重庆 400715)

摘 要: 土壤学始于对土壤剖面及其形态特征的观察, 剖面发生层的划分与发生层边界特征的描述是土壤调查的基础。实地划分发生层需要丰富的土壤学实践经验, 存在主观和缺乏统一划分标准的问题。以紫色土剖面图像为研究对象, 采用 K-means 聚类 and 图像分割技术, 结合图像的颜色特征 (CIE Lab 色彩空间) 和纹理特征 (Entropy) 识别紫色土剖面发生层边界, 并与实地划分的结果进行比较。结果表明: (1) CIE Lab 色彩空间的 a、b 通道和 Entropy 纹理特征, 可以划分出供试剖面的主要发生层 (A、B、C) 和基岩 (R); (2) 聚类识别的发生层数量和发生层深度与实地识别的结果基本一致; 除 Z2 剖面的 C 层和 Z6 剖面的 Ap 层聚类识别与实地识别的发生层下边界深度相差较大 (分别为 13 cm 和 8 cm) 外, 其余发生层下边界深度相差均在 3 cm 以内; (3) 聚类识别的发生层边界形状更为不规则, 明显度更为模糊。K-means 聚类 and 图像分割技术实现了紫色土剖面发生层边界的客观识别, 可为土壤剖面智能辨识系统的开发提供科学参考。

关键词: 剖面图像; 发生层; K-means 聚类; 图像分割; 颜色; 纹理

中图分类号: S159 **文献标志码:** A

Boundary Identification of Purple Soil Horizon Based on K-means Clustering and Image Segmentation

YANG Kai¹, CI En^{1, 2†}, LIU Bin¹, CHEN Yangyang¹, XIE Yu¹

(1. College of Resource and Environment, Southwest University, Chongqing 400715, China; 2. Engineering Research Center of Agricultural Non-point Source Pollution Control in Three Gorges Reservoir Region, Chongqing 400715, China)

Abstract: **【Objective】** Pedology begins with the observation of soil profile and its morphological characteristics. The division of the soil profile horizon and description of the characteristics of the horizon boundary are the basis of soil investigation. The division of soil horizon in the field requires rich practical experience in pedology and is more subjective, which makes it difficult to form a set of unified division standards. **【Method】** In this paper, the purple soil profile image was taken as the research object, and using K-means clustering and image segmentation technology, combined with the color (CIE Lab color space) and texture characteristics (Entropy) of the image, we identified the horizon boundary of the purple soil profile, by comparing with the results of field division. **【Result】** The results show that (1) the a and b channels of CIE Lab color space and Entropy texture

^{*} 国家自然科学基金项目 (41977002) 和中央高校基本科研业务费专项资金项目 (XDJK2020B069) 资助 Supported by the National Natural Science Foundation of China (No. 41977002) and the Fundamental Research Funds for the Central Universities (No. XDJK2020B069)

[†] 通讯作者 Corresponding author, E-mail: cien777@163.com

作者简介: 杨 凯 (1997—), 男, 重庆市人, 硕士研究生, 主要研究方向为土壤发生与分类。E-mail: yangkkk1015@163.com

收稿日期: 2023-02-17; 收到修改稿日期: 2023-05-12; 网络首发日期 (www.cnki.net): 2023-11-21

characteristics can delineate the master horizon (A, B, and C) and bedrock(R) of the profile; the a channel values range from 7-22, the b channel values range from 7-19, and the Entropy values were 4 or 5; the Munsell colors converted by the CEL XYZ system had a certain deviation from the colors visually discerned in the field using colorimetric cards, with a hue range of 10R-2.5Y, a value range of 4-8, and a chroma range of 3-8. (2) The number of soil horizon and the depth of soil horizon identified by clustering were consistent with the results of field identification; the difference between the lower boundary depth of soil horizon identified by clustering identification and those identified in the field was within 3 cm, except for C in profile Z2 and the Ap in profile Z6, where the difference was larger (13 cm and 8 cm, respectively). (3) The topography of the soil horizon identified by clustering was more irregular and the distinctness was more blurred. The clustering algorithm can identify more subtle differences in the soil profile image and reflect the local variation of soil properties in more detail. 【Conclusion】 K-means clustering and image segmentation techniques achieved the identification of the horizon boundary of purple soil, and this study provides a scientific reference for the development of an intelligent identification system for soil profiles.

Key words: Profile image; Horizon; K-means clustering; Image segmentation; Color; Texture

土壤学始于对土壤剖面及其形态特征的观察^[1], 土壤剖面形态特征是土壤形成过程的直观表现, 也是野外调查获得的第一手资料^[2]。土壤在形成过程中所形成的具有特定性质和组成的、大致与地面相平行的, 并具有成土过程的层次称为发生层^[3-4]。发生层的划分与发生层边界特征的描述是土壤调查的基础^[5], 在过去几十年, 划分土壤剖面发生层的技术和工具并未发生实质性的变化^[6-7], 仍是基于传统的“手眼并用”法来观察土体中砾石、根系、土壤颜色、结持性、结构体、新生体等的差异。土壤发生层的准确识别和划分需要土壤调查者具有丰富的土壤学基础知识和实践经验, 但大量实践表明, 即使面对同一个土壤剖面, 不同的专业人员对土壤发生层的识别结果也会存在一定差异, 而非专业人员则更难准确地识别土壤发生层, 这限制了土壤学研究与应用快速发展^[6-7]。

为了对土壤属性进行更客观和定量地描述, Hartemink 和 Minasny^[6]提出了数字土壤形态计量学 (Digital Soil Morphometrics, DSMorph) 这一概念, 数字形态测量技术已应用于诊断土壤发生层, 包括探地雷达 (GPR)、电阻率 (ER)、X 射线荧光光谱仪 (XRF) 等, 但该方法通常检测费用高、耗时耗力, 不适用于野外的快速测定以及大规模非专业人员的使用。随着计算机视觉领域的高速发展, 数字图像处理技术同样在经历着革新与更替^[8]。近年来, K-means 聚类 and 图像分割技术在智能安防、无人驾驶、卫星遥感、医学影像处理、生物特征识别、农业生产等领域均得到了广泛应用, 如金文东等^[9]采用 K 均值聚类并结合空间分析和统计学方法实现了

肿瘤生物标记物的空间数据分析及可视化, Zhang 和 Peng^[10]基于 PSO 和 K-means 聚类对各种复杂环境下的农产品图像进行了有效、准确地分割, 但 these 技术在土壤学领域尤其是土壤剖面识别方面的应用和研究尚鲜有报道。

紫色土是重庆市最为重要、分布面积最广的旱地土壤类型^[11], 由于其土体中物质淋溶和淀积较弱, 颜色较为均一, 剖面多无明显层次分异^[12], 导致发生层划分存在较大的难度。由此, 建立方便、客观、定量的识别土壤剖面发生层边界的方法, 对于通过研发手机 APP 终端, 辅助户外紫色土剖面发生层识别和划分具有积极的指导意义。为此, 本文拟以紫色土剖面图像为研究对象, 采用 K-means 聚类和图像分割技术, 结合图像的颜色特征 (CIE Lab 色彩空间) 和纹理特征 (Entropy), 实现紫色土剖面发生层边界的识别。研究结果可为土壤剖面智能辨识系统的开发提供科学参考。

1 材料与方法

1.1 紫色土剖面图像

供试的紫色土剖面位于重庆市境内典型紫色土分布区, 在重庆市紫色土中有着较好的代表性, 土壤剖面深度从 50 cm 到 110 cm 不等, 质地多为壤土。本研究选择了 6 个紫色土剖面样本 (图 1), 有无明显层次划分、通体颜色均一的剖面 (Z1、Z2、Z5、Z6), 也有层次较清晰、层间颜色差异明显的剖面 (Z3、Z4)。拍照前对剖面进行清洁和光滑处理, 以减少表面粗糙度和阴影对图像质量的影响, 在自然



图 1 供试紫色土剖面图像

Fig. 1 Profile images of the tested purple soils

光照条件下数码相机的镜头垂直于剖面拍摄, 所用相机型号为 SONY NEX-5N, 图像为高分辨 (2364×4921) 的 JPEG 格式。完成拍摄后, 按传统的“手眼并用”方法实地对土壤剖面发生层及其边界特征进行了识别和划分。

表 1 列出了关于供试紫色土剖面的详细信息, 包括剖面号、剖面地点、母岩、发生层名称、发生层深度、质地以及实地识别的发生层边界特征 (过渡形状和明显度)。

1.2 图像特征提取

在 Adobe Photoshop CC 2019 中将每个土壤剖面图像按照一定的大小进行裁剪, 只保留光滑面, 裁剪深度从表层 (Ap) 到基岩 (R)。

图像特征包括颜色特征、纹理特征、形状特征和空间关系特征^[13], 本文使用土壤剖面图像的颜色特征 (CIE Lab 色彩空间) 和纹理特征 (Entropy) 进行聚类 and 划分发生层。通常通过相机或者手机所获得的图像均为 RGB 图像, RGB 色彩空间需要先转换成 CIE XYZ 色彩空间再转换为 CIE Lab 色彩空间, 即 RGB→XYZ→Lab, RGB 转 XYZ 公式如下^[14]:

$$\begin{cases} X = R \times 0.4124 + G \times 0.3576 + B \times 0.1805 \\ Y = R \times 0.2126 + G \times 0.7152 + B \times 0.0722 \\ Z = R \times 0.0193 + G \times 0.1192 + B \times 0.9505 \end{cases} \quad (1)$$

式中, X 、 Y 、 Z 分别为 XYZ 色彩空间的三色通道的值。

XYZ 再转换为 Lab 色彩空间:

$$\begin{cases} L = 116f\left(\frac{Y}{Y_n}\right) - 16 \\ a = 500\left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right] \\ b = 200\left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right] \end{cases} \quad (2)$$

式中, X_n 、 Y_n 、 Z_n 分别为参照白点的 CIE XYZ 三色刺激值, 计算时分别取 96.422 1、100.000 0、82.522 1; f 为以色道值与刺激值之比为变量的函数。 f 的取值情况如下:

$$f(t) = \begin{cases} t^{1/3} & t > \left(\frac{6}{29}\right)^3 \\ \frac{1}{3}\left(\frac{29}{6}\right)^2 + \frac{16}{116} & \text{其他} \end{cases} \quad (3)$$

式中, t 为 XYZ 色彩空间色道值与刺激值之比。

Entropy 通过以下公式得到^[15]:

$$ENT = -\sum_i \sum_{jp} (i, j) \lg p(i, j) \quad (4)$$

式中, $p(i, j)$ 表示特定位置关系下的像素对的频数。

1.3 发生层边界识别

剖面图像的 K-means 聚类 and 分割均在 MATLAB R2021b 中实现, 具体的流程如图 2 所示。图像结果包括图像特征图、聚类图、发生层边界图和发生层

表 1 供试紫色土剖面信息

Table 1 Profile information of the tested purple soils

剖面号 Profile No.	地点 Location	母岩 Parent rock	发生层 Horizon	深度 Depth /cm	质地 Soil texture	过渡形状 Topography	明显度 Distinctness	土壤颜色 Soil color	
								干态 Dry	润态 Wet
Z1	重庆市合川区古楼镇	紫色泥岩、砂岩	Ap	0~17	壤土 ^①	平滑 ^⑥	渐变 ^⑩	7.5R 5/3	7.5R 4/3
			Bw	17~51	壤土 ^①	平滑 ^⑥	突变 ^⑧	7.5R 5/3	7.5R 4/3
			R	51~65	—	—	—	—	—
Z2	重庆市江津区柏林镇	紫色泥岩	Ap	0~20	粉质黏壤土 ^②	平滑 ^⑥	渐变 ^⑩	2.5YR 5/2	2.5YR 4/2
			Bt	20~58	黏土 ^⑤	平滑 ^⑥	渐变 ^⑩	2.5YR 5/2	2.5YR 4/2
			C	58~68	壤土 ^①	平滑 ^⑥	清晰 ^⑨	2.5YR 4/2	2.5YR 3/2
			R	68~110	—	—	—	—	—
Z3	重庆市忠县官坝镇	紫色泥岩、 粉砂岩	Ap	0~15	黏壤土 ^④	波状 ^⑦	渐变 ^⑩	5YR 5/3	5YR 3/3
			Bw	15~40	粉质黏壤土 ^②	波状 ^⑦	突变 ^⑧	5YR 5/4	5YR 4/4
			R	40~52	—	—	—	—	—
Z4	重庆市丰都县树人镇	紫色砂岩、泥岩	Ap	0~15	壤土 ^①	波状 ^⑦	清晰 ^⑨	10RP 6/3	10RP 5/3
			Br1	15~54	壤土 ^①	波状 ^⑦	清晰 ^⑨	10RP 6/2	10RP 5/2
			Br2	54~70	壤土 ^①	平滑 ^⑥	突变 ^⑧	10RP 5/3	10RP 4/3
			R	70~82	—	—	—	—	—
Z5	重庆市长寿区八颗街道	紫色泥岩	Ap	0~18	壤土 ^①	波状 ^⑦	渐变 ^⑩	10R 6/3	10R 5/3
			Bw	18~37	壤土 ^①	波状 ^⑦	突变 ^⑧	10R 6/3	10R 5/3
			R	37~50	—	—	—	—	—
Z6	重庆市合川区钓鱼城街道	紫色泥岩	Ap	0~20	壤土 ^①	平滑 ^⑥	模糊 ^⑪	7.5YR 4/2	7.5YR 3/2
			Br	20~63	粉壤土 ^③	平滑 ^⑥	渐变 ^⑩	7.5YR 4/2	7.5YR 3/2
			E	63~75	壤土 ^①	波状 ^⑦	突变 ^⑧	7.5YR 4/2	7.5YR 6/2
			R	75~92	—	—	—	—	—

注：Ap，耕作层；E，漂白层；Bw，雏形层；Br，氧化还原层；Bt，黏淀层；C，母质层；R，基岩。质地为土壤颗粒组成测定的结果。Note: Ap, cultivated horizon; E, albic horizon; Bw, carnbic horizon; Br, redox horizon; Bt, argillic horizon; C, parent material horizon; R, rock. The soil texture is the result of particle composition measurement.

① Silty clay loam; ② Silty clay loam; ③ Slit loam; ④ Clay loam; ⑤ Clay; ⑥ Smooth; ⑦ Wavy; ⑧ Abrupt; ⑨ Clear; ⑩ Gradual; ⑪ Diffuse

深度图。发生层深度为各发生层下边界深度的平均值。发生层边界特征包括过渡形状和明显度，过渡形状通过发生层边界图判别，包括平滑、波状、不规则和间断；明显度通过聚类图判别，包括突变、清晰、渐变和模糊^[16]。

2 结 果

2.1 图像特征

供试紫色土剖面图像的颜色特征和纹理特征如图 3 所示。从原图和裁剪图可以直观地看出，Z1、

Z2、Z3 和 Z5 剖面通体颜色偏红，Z4 剖面的 Br1 层颜色偏黄，Z6 剖面有颜色偏白的 E 层；L 通道与其所在剖面的灰度图十分接近，表现出与灰度图像相似的特征；a 通道能够识别出 Z3 剖面中的 Bw 层，从而划分出 Ap 层和 R 层，Z5 剖面 and Z6 剖面的 R 层也能在 a 通道中明显识别，对于 Z4 剖面的 Br1 层和 Br2 层以及 Z6 剖面的 Br 层和 E 层，由于其颜色偏黄和偏白，a 通道无法对其进行准确的识别；b 通道能够识别出 Z3 剖面的 Bw 层、Z4 剖面的 Br1 层和 Z5 剖面的 R 层；a、b 通道还能够较好地识别出 Z2、Z3 剖面中的孔洞，Z1、Z2 剖面通体颜色差

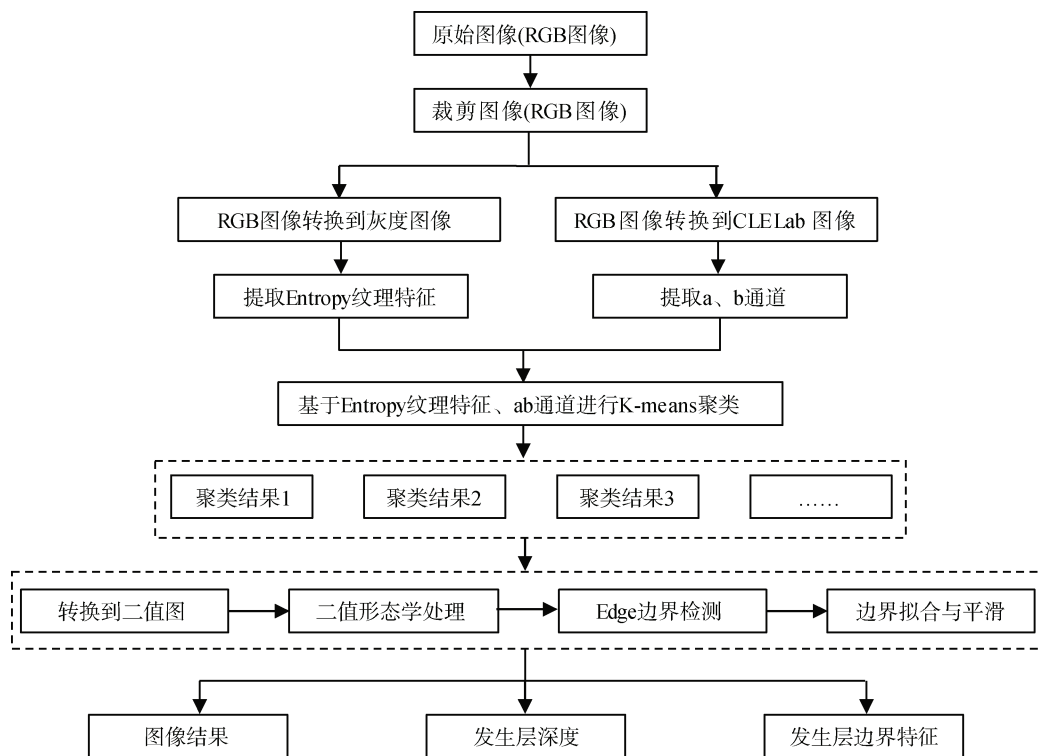


图 2 基于 K-means 聚类 and 图像分割的发生层边界识别流程

Fig. 2 Horizon boundary recognition process based on K-means clustering and image segmentation

异小, a、b 通道均无法明显地识别出发生层; 在 Entropy 纹理图中能够看出 Z1 剖面的 Bw 层、Z2 剖面的 Bt 层、Z4 剖面的 Ap 层和 R 层和 Z6 剖面的 R 层与其相邻的发生层纹理差异较大, 能够被准确识别。

由表 2 可知, 供试紫色土剖面图像的 CIE Lab 色彩空间 L 坐标值范围为 24~51, 多集中在 30~35 之间, ΔL 范围为 1~14, 多集中在 3~7 之间; a 坐标值范围为 7~22, 多集中在 12~18 之间, Δa 范围为 0~7, 多集中在 2~4 之间; b 坐标值范围为 7~19, 多集中在 11~14 之间, Δb 范围为 0~6, 多集中在 1~4 之间; Entropy 的值为 4 或者 5, 除 Z6 剖面的 Ap 层和 Z3 剖面的 R 层 Entropy 为 4, 其余剖面的 Ap 层和 R 层均为 5; Z1、Z2、Z3、Z5、Z6 剖面的 B 层 Entropy 为 4, Z4 剖面的 Br1 层为 5, Br2 层为 4。由表 3 可知, 由 CIE XYZ 转换而来的 Munsell 颜色, 色调范围为 10R~2.5Y, 多集中在 2.5YR~5YR 之间; 明度范围为 4~8, 多集中在 5~6 之间; 彩度范围为 3~8 多集中在 4~6 之间; 实地由土壤比色卡判读的土壤颜色色调 (干/润态) Z1~Z6 剖面依次为 7.5R、2.5YR、5YR、10RP、10R、7.5YR。

2.2 聚类识别

将实地划分的发生层与通过 K-means 聚类 and 图像分割得到的发生层和聚类图进行比较, 如图 4 所示。Z1、Z3、Z5 剖面在实地划分出 Ap、Bw 和 R 层, 通过聚类 and 图像分割技术, 这 3 个剖面的所有发生层均被准确识别; Z4 剖面在实地划分出的 Ap 层、Br1 层、Br2 层和 R 层同样被准确识别; Z2 剖面通过聚类识别的发生层数量与实地识别的发生层数量一致, 但 Ap 层和 Bt 层识别的结果与实地划分较为接近, C 层和 R 层均被错误识别; Z6 剖面通过聚类识别的发生层数量与实地识别的发生层数量一致, Ap 层与实地划分的结果相差较大。总的来看, 通过聚类识别的发生层数量与实地识别的发生层数量基本一致, Ap 层和 R 层被准确识别, 除 Z2 剖面 C 层、R 层和 Z6 剖面的 Ap 层以外, 其余发生层深度与实地划分的发生层深度也基本一致。

2.3 边界特征

将实地划分的发生层边界特征与通过 Lab 色空间聚类得到的发生层边界特征进行比较, 表 4 是对基于不同划分方式的紫色土发生层深度对比。在识别出的所有发生层中, Z2 剖面 C 层的发生层下边界

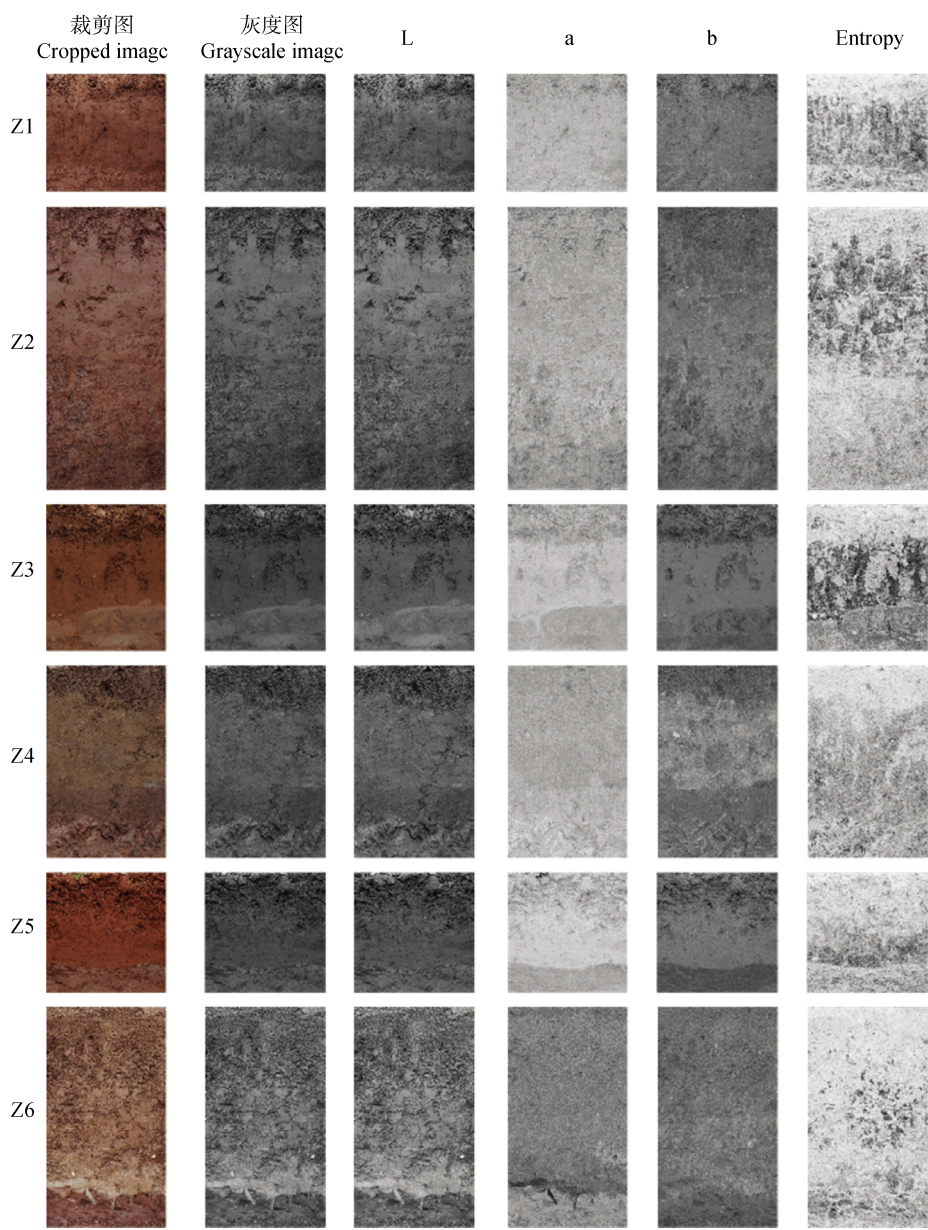


图 3 供试紫色土剖面图像的颜色和纹理特征

Fig. 3 Color and texture features of the tested purple soil profile images

表 2 供试紫色土剖面发生层的 Lab 值、Entropy 值及其层间差

Table 2 The Lab and Entropy and interlayer difference of the tested purple soil profiles

剖面号 Profile No.	发生层 Horizon	L	a	b	Entropy	ΔL	Δa	Δb	$\Delta Entropy$
Z1	Ap	35	15	13	5	3	3	1	1
	Bw	32	18	14	4	1	0	1	1
	R	33	18	13	5	—	—	—	—
Z2	Ap	32	12	7	5	7	2	3	1
	Bt	39	14	10	4	5	0	1	1
	C	34	14	11	5	6	2	3	0
	R	28	12	9	5	—	—	—	—

续表									
剖面号 Profile No.	发生层 Horizon	L	a	b	Entropy	ΔL	Δa	Δb	$\Delta Entropy$
Z3	Ap	24	12	13	5	5	4	4	1
	Bw	29	16	17	4	6	3	3	0
	R	35	13	14	4	—	—	—	—
Z4	Ap	28	7	7	5	7	1	4	0
	Br1	35	8	11	5	5	2	3	1
	Br2	30	10	8	4	3	1	1	1
	R	33	11	9	5	—	—	—	—
Z5	Ap	26	19	17	5	3	3	2	1
	Bw	29	22	19	4	3	7	6	0
	R	32	15	13	4	—	—	—	—
Z6	Ap	40	13	12	4	1	0	0	0
	Br	41	13	13	4	10	2	1	1
	E	51	11	14	5	14	4	5	0
	R	37	15	9	5	—	—	—	—

注：L、a、b 和 Entropy 为各发生层的平均值， ΔL 、 Δa 、 Δb 和 $\Delta Entropy$ 为相邻发生层之间 L、a、b 和 Entropy 差的绝对值。
Note: L, a, b and Entropy are the average values of each horizon, while ΔL , Δa , Δb and $\Delta Entropy$ are the absolute values of the differences between L, a, b and Entropy between adjacent horizon.

表 3 供试紫色土剖面颜色转换结果

Table 3 Color conversion results of the tested purple soil profiles

剖面号 Profile No.	发生层 Soil horizon	CIE XYZ			转换后的 Munsell 颜色
		X	Y	Z	Converted Munsell color
Z1	Ap	31	27	19	2.5YR 6/5
	Bw	23	19	12	10R 5/6
	R	—	—	—	—
Z2	Ap	24	21	18	7.5R 5/4
	Bt	32	28	22	10R 6/5
	C	25	22	16	2.5YR 5/5
	R	—	—	—	—
Z3	Ap	15	13	8	2.5YR 4/4
	Bw	18	1	8	5YR 5/6
	R	—	—	—	—
Z4	Ap	17	17	14	2.5Y 5/2
	Br1	24	23	16	7.5YR 5/3
	Br2	18	16	13	10R 5/3
	R	—	—	—	—
Z5	Ap	18	15	8	2.5YR 5/6
	Bw	20	15	7	10YR 5/8

续表

剖面号 Profile No.	发生层 Soil horizon	CIE XYZ			转换后的 Munsell 颜色 Converted Munsell color
		X	Y	Z	
Z5	R	—	—	—	—
Z6	Ap	37	34	25	5YR 6/4
	Br	37	34	25	5YR 6/4
	E	57	55	42	7.5YR 8/4
	R	—	—	—	—

注：转换后的 Munsell 颜色由 CIE XYZ 颜色空间转换而来。Note: The converted Munsell color is converted from the CIE XYZ color space.

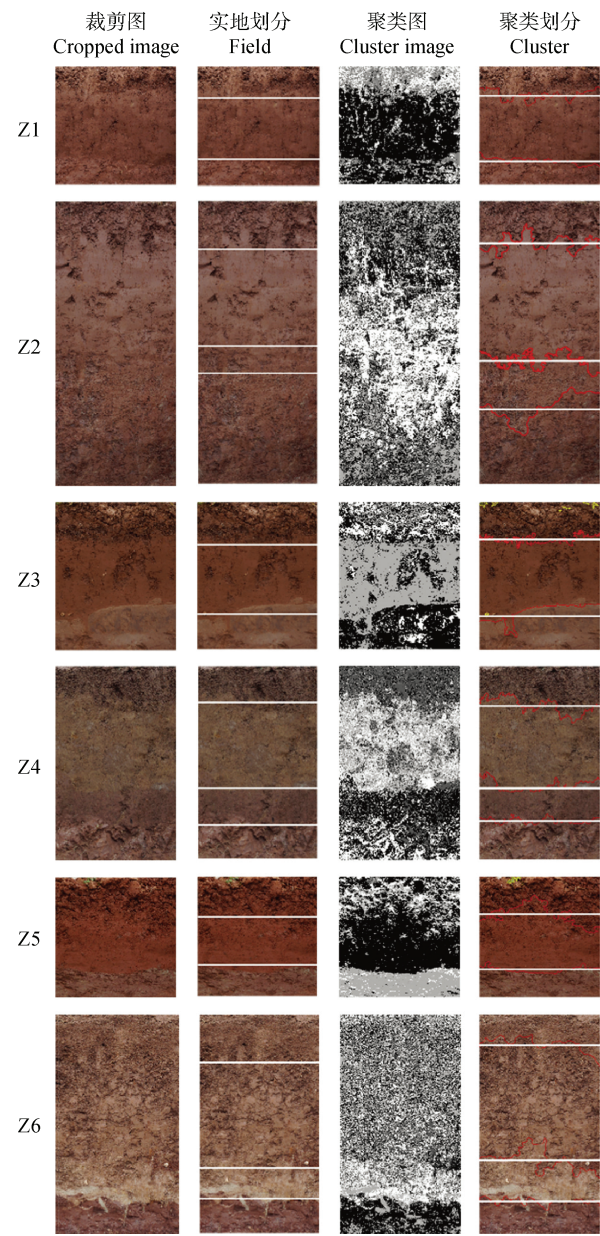


图 4 供试紫色土发生层实地划分与聚类划分对比
Fig. 4 Comparison of field division and cluster division of horizon of the tested purple soil profiles

深度与实地识别的发生层下边界深度相差最大，为 13 cm；其次是 Z6 剖面的 Ap 层，为 8 cm；其余剖面的发生层下边界深度与实地识别的发生层下边界深度差均在 3 cm 以内，其中 Z2 剖面的 Bt 层、Z4 剖面的 Br2 层、Z6 剖面的 Br 层为 3 cm，Z2、Z3、Z5 剖面的 Ap 层为 2 cm，Z3 剖面的 Bw 层下边界深度差为 0 cm，与实地划分的发生层完全一致，其余发生层下边界深度差为 1 cm。

由表 2 和表 5 可知，Z1 剖面的 Bw 层、Z2 剖面的 Bt 层和 Z4 剖面的 Br2 层聚类识别的过渡形状与实地划分的过渡形状一致，均为平滑；Z3 剖面的 Ap 层、Z4 剖面的 Ap 层和 Br1 层以及 Z6 剖面的 E 层聚类识别的过渡形状与实地划分的过渡形状一致，均为波状；其余发生层聚类识别的过渡形状与实地划分的过渡形状不一致。Z1 剖面的 Bw 层实地划分的明显度为突变，聚类识别为渐变；Z2 剖面的 C 层实地划分的明显度为清晰，聚类识别为模糊；Z4 剖面的 Ap 层实地划分的明显度为清晰，聚类识别为渐变；Z5 剖面的 Ap 层和 Z6 剖面的 Br 层实地划分的明显度均为渐变，聚类识别为模糊；其余发生层聚类识别的明显度与实地划分的明显度一致。

图 5 是实地划分与聚类划分的过渡形状和明显度数量对比，可以得到聚类划分的过渡形状更多的是波状，实地划分的过渡形状更多的是平滑；实地划分没有过渡形状为不规则的边界，而聚类划分识别出 3 个过渡形状为不规则的边界；实地划分和聚类划分的明显度渐变类型的数量均一样；实地划分中明显度为模糊的占少数，数量为 1，而聚类划分中明显度为模糊的边界增加为 4，清晰占少数，数量为 1。可见，聚类识别出的发生层边界更为精细和不规则。

表 4 基于不同划分方式的供试紫色土发生层深度对比

Table 4 Comparison of horizon depth of the tested purple soil profiles based on different classification methods				
剖面号	发生层	实地划分	聚类划分	下边界深度差
Profile No.	Horizon	Field/cm	Clustering /cm	Lower boundary depth difference/cm
Z1	Ap	0~17	0~16	1
	Bw	17~51	16~52	1
	R	51~65	52~65	—
Z2	Ap	0~20	0~18	2
	Bt	20~58	18~61	3
	C	58~68	61~81	13
	R	68~110	81~110	—
Z3	Ap	0~15	0~13	2
	Bw	15~40	13~40	0
	R	40~52	40~52	—
Z4	Ap	0~15	0~16	1
	Br1	15~54	16~53	1
	Br2	54~70	53~67	3
	R	70~82	67~82	—
Z5	Ap	0~18	0~16	2
	Bw	18~37	16~38	1
	R	37~50	38~50	—
Z6	Ap	0~20	0~12	8
	Br	20~63	12~60	3
	E	63~75	60~76	1
	R	75~92	76~92	—

3 讨 论

3.1 颜色和纹理特征的发生层识别原理及效果

颜色和纹理是图像分割中常用的图像特征^[17]。常见的颜色特征为 RGB 色彩空间，RGB 色彩空间将所有的颜色看作是三基色的组合，但三分量之间存在很强的相关性^[17-18]，因而不适于直接用于基于三个分量独立运算的图像分割。CLE Lab 色彩空间是基于人类视觉原理的一种色彩系统^[9]，不同于日常生活中广为人知的 RGB 色彩空间，CIE Lab 色彩空间通过数字化的方式描述人的视觉感应，因此，Lab 色彩空间表现能力强，色域宽，在进行图像分割时可以使图像之间的差异更显著^[19]。L 通道代表图像的亮度或明度，取值范围为（0~100），表示纯

黑到纯白，与灰度图有着相似的特性，因此不适于作为聚类和图像分割的变量，而 a、b 通道表示颜色的变化，a 通道取值范围为（+127~-128），表示了从红色到绿色的颜色变化，b 通道取值范围为（+127~-128），表示颜色从黄色到蓝色的变化，因此 a、b 通道可以作为聚类和图像分割的有效变量。在本文 6 个剖面中（图 3），利用 a、b 通道能够区分主要的发生层（A、B）和基岩层（R），但是仅通过颜色特征不能准确地识别出每一个发生层，如 Z1 和 Z2 剖面，这两个剖面通体颜色差异较小，各发生层之间 a、b 通道变化值在 3 以内，导致识别效果并不理想。

颜色是土壤最重要的特征之一^[20]，土壤颜色是

判断土壤形成和发育的依据和指标^[11,21],同时,土壤颜色也与土壤属性有直接或间接的关系,包括土壤有机碳、铁含量和母岩堆积方式^[22-24]。本研究所选用的紫色土是一类岩性土,受母岩影响深刻,其土壤形状的诸多方面均继承了母岩特性^[11, 25],由表 1、表 3 可知,供试土壤均是由紫色泥岩或紫色砂岩发育而来,其色调范围为 10PR~7.5YR,由 CIE XYZ 转换而来的 Munsell 颜色色调范围为 10R~2.5Y,其主色调均能达到 5YR 或更红,符合系统分类检索中铁质特性和亚类前缀为“红色”的色调要求^[26]。但是转换而来的色调与实地比色的色调出现了一定的偏差,可能是因为:(1)标准土壤比色卡和 CIE 系统转换为 Munsell 系统过程均存在一定的色差和误差;(2)实地判读土壤颜色存在较强的主观性,导致颜色的误判;(3)外部环境条件对人眼判别和照片质量的影响,如光照条件和拍摄角度等。

表 5 基于聚类划分的供试紫色土发生层过渡形状和明显度

Table 5 Topography and distinctness of the tested purple soil profiles horizon based on cluster division			
剖面号 Profile No.	发生层 Horizon	聚类划分 Clustering	
		过渡形状 Topography	明显度 Distinctness
Z1	Ap-Bw	波状 ^②	渐变 ^⑥
	Bw-R	平滑 ^①	渐变 ^⑥
Z2	Ap-Bt	波状 ^②	渐变 ^⑥
	Bt-C	平滑 ^①	渐变 ^⑥
	C-R	波状 ^②	模糊 ^⑦
Z3	Ap-Bw	波状 ^②	渐变 ^⑥
	Bw-R	不规则 ^③	突变 ^④
Z4	Ap-Br1	波状 ^②	渐变 ^⑥
	Br1-Br2	波状 ^②	清晰 ^⑤
	Br2-R	平滑 ^①	突变 ^④
Z5	Ap-Bw	不规则 ^③	模糊 ^⑦
	Bw-R	平滑 ^①	突变 ^④
Z6	Ap-Br	波状 ^②	模糊 ^⑦
	Br-E	不规则 ^③	模糊 ^⑦
	E-R	波状 ^②	突变 ^④

①Smooth; ②Wavy; ③Irregular; ④Abrupt; ⑤Clear;
⑥Gradual; ⑦Diffuse

纹理是一种重要的视觉线索,是图像中普遍存在而又难以描述的特征,纹理特征是一种全局特征,反映的是图像中同质现象的视觉特征,体现物体表面的具有缓慢变换或周期性变化的表面组织结构排列属性^[18, 27]。Haralick 等^[28]定义了 14 个用于纹理分析的灰度共生矩阵特征参数,常用的参数主要有均值 (Mean)、方差 (Variance)、相关性 (Correlation)、熵 (Entropy)。Entropy 反映出图像纹理的非均匀程度或者复杂程度,图像边缘处灰度变化剧烈且局部纹理熵具有较大值,而其他区域的局部纹理熵具有较小值,因此局部区域熵值的大小可作为边缘的判断依据,而土壤剖面各发生层之间由于孔隙、砾石含量、结构与质地的不同^[29-30],在图像中则反映为纹理熵值大小的差异。由表 2 可知,大多数发生层之间纹理均会发生变化,且 A 层纹理熵值高于 B 层,这可能是因为 A 层土壤含有较多的根系,并且由于耕作等原因导致孔隙较多,由此导致 A 层与 B 层之间纹理的差异。

3.2 K-means 聚类、图像分割与发生层边界识别

图像分割是指将图像划分成互不相交的、有意义的子区域,在同一个区域的像素点具有一定的相关性,不同区域的像素点存在一定的差异性,即是对图片中有相同性质的像素赋予相同标签的过程^[31-32]。随着分割场景的复杂化,图像分割技术也在不断地完善和发展,陆续出现了基于边缘、区域、图论和聚类等的分割方法^[32-34]。Mac 于 1967 年提出 K-means 聚类法^[35],该方法属于聚类分割,它首先确定 K 个类,每个类随机选择初始聚类中心,通过对初始聚类中心进行聚类运算得到 K 个簇,反复迭代聚类的结果至收敛,从而实现分割。K-means 聚类算法是目前颇受欢迎且应用较为广泛的聚类分析方法之一。该算法用于图像分割具有快速、直观、易于实现的优点,但其缺点是要事先确定聚类数目^[36]。本研究中,各剖面的 K 值是基于实地发生层划分的结果确定,虽然识别出的发生层数量与实地识别的发生层数量一致,但会存在识别错误的情况,如 Z2 和 Z6 剖面(图 4)。另一方面,利用聚类算法对图像进行分割,仅仅使用了图像的颜色特征和纹理特征,而未利用形状特征和空间关系特征,因而图像识别的结果受噪声影响很大。在未来的研究中,可以尝试使用结合更多的图像特征和分割方法(如区域生长^[37]和分水岭算法^[38])并结合滤波^[39]等技术噪声的影响,使分割效果更加理想。

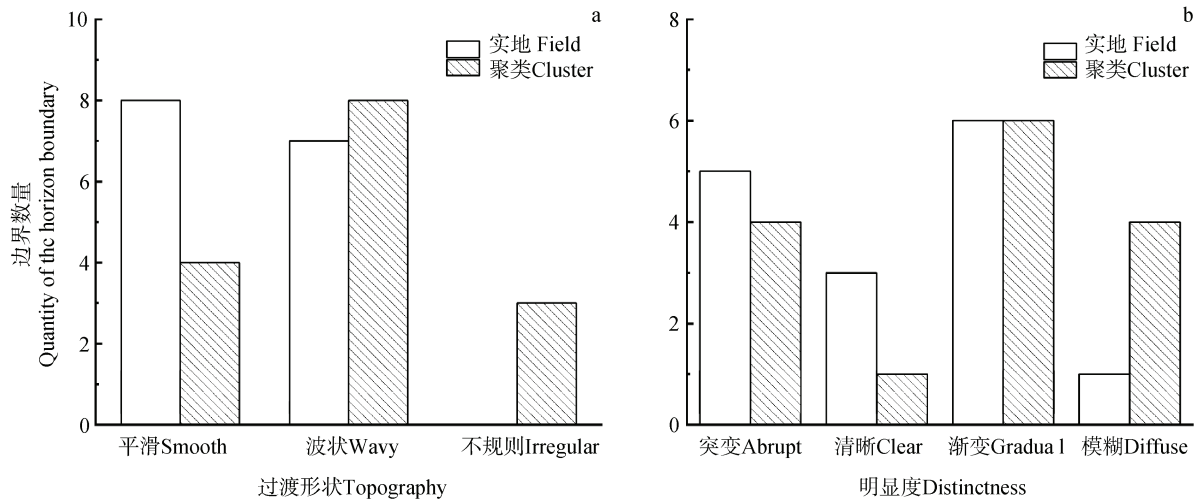


图 5 基于不同划分方式的供试紫色土发生层过渡形状和明显度数量对比

Fig. 5 Comparison of topography and distinctness quantity of the tested purple soil profiles horizon based on different division methods

发生层深度由边界点的纵坐标取平均值得到，由于实地判读发生层深度存在误差以及拍照时不能保证完全垂直于剖面拍摄，因此，本研究认为当聚类识别的发生层深度与实地划分的发生层下边界深度相差不超过 3 cm 时，该识别结果是准确的。由表 4 可知，除被错误识别的 Z2 剖面的 C 层和 Z6 剖面的 Ap 层下边界深度与实地划分的深度相差为 13 cm 和 8 cm 外，其余各剖面所识别的发生层下边界深度与实地划分的发生层深度相差均小于 3 cm，因此对于发生层深度的判别认为是准确可靠的，而造成部分发生层被错误识别（下边界深度相差大于 3 cm）的原因则是发生层之间颜色和纹理差异较小或者没有差异，如 Z2 剖面 C 层的 Δa 、 Δb 、 $\Delta \text{Entropy}$ 分别为 2、3、0，Z6 剖面 Ap 层的 Δa 、 Δb 、 $\Delta \text{Entropy}$ 均为 0（表 3），在进行聚类时则不能准确地识别出发生层边界，在今后的研究中，可以结合砾石含量和孔隙等因素综合考虑，提高划分精度。通常情况下，计算机能够较人眼观察到更多的细节，聚类算法能识别出图像中更多细微的差异^[31]。由表 5、图 4 和图 5 可知，聚类使得发生层之间的边界更加清晰和直观，同时，聚类识别出的发生层轮廓更加的精细和不规则，更详细地反映了土壤属性的局部变异。此外，虽然图像颜色与实地判别的土壤颜色出现了一定的偏差，但 K-means 聚类 and 图像分割技术对紫色土发生层边界识别的结果与专家实地识别的结果基本一致，由此可见，这两种方法有着较好的适用性和稳定性，不易因拍摄器材、拍摄人员的不同导

致图像颜色差异而引起识别结果不同。

由于发生层边界形状和明显度是由图像识别的结果再次进行人为判读的，也存在一定的主观性，因此有待开发出一种新的算法实现发生层过渡形状和明显度的定量化判别。K-means 聚类 and 图像分割可以识别出主要的发生层，但进一步识别（如区分出 Br1 和 Br2）则很难实现，应加入更多的图像特征和土壤剖面特征并结合更多的分割方法对发生层进行更进一步的识别。在未来土壤剖面智能辨识系统的开发中，建议将相关系统与传统“手眼并用”方法相结合，以保证更为准确地识别和划分土壤发生层。

4 结 论

K-means 聚类 and 图像分割技术可以实现紫色土发生层边界的识别。CIE Lab 色彩空间的 ab 通道和 Entropy 纹理特征能够识别出紫色土的主要发生层和基岩，聚类识别的发生层数量和深度与实地识别的发生层数量和深度基本一致，聚类识别的发生层过渡形状更为不规则、精细程度更高；此外，K-means 聚类 and 图像分割技术还有着较好的适用性和稳定性。

参考文献 (References)

- [1] Gallegos Á, García-Oliva F, Pereira-Corona A, et al. Digital soil morphometrics of coarse fragments and

- horizon delineation in soil profiles from central Mexico[J]. *Geoderma Regional*, 2021, 26: e00403.
- [2] Ci E. Soil Series of China-Chongqing[M]. Beijing: Science Press, 2020. [慈恩. 中国土系志·重庆卷[M]. 北京: 科学出版社, 2020.]
- [3] Huang C Y, Xu J M. *Agrology*[M]. 3rd ed. Beijing: China Agriculture Press, 2010. [黄昌勇, 徐建明. 土壤学[M]. 3 版. 北京: 中国农业出版社, 2010.]
- [4] Li T J. *Pedogeography*[M]. 3rd ed. Beijing: Higher Education Press, 2004. [李天杰. 土壤地理学[M]. 3 版. 北京: 高等教育出版社, 2004.]
- [5] Zheng G H, Jiao C X, Shanguan C X, et al. Horizon classification in soil profile using imaging spectroscopy[J]. *Spectroscopy and Spectral Analysis*, 2019, 39 (3): 882-885. [郑光辉, 焦彩霞, 上官晨曦, 等. 基于成像光谱技术的土壤剖面发生层划分初探[J]. 光谱学与光谱分析, 2019, 39 (3): 882—885.]
- [6] Hartemink A E, Minasny B. Towards digital soil morphometrics[J]. *Geoderma*, 2014, 230/231: 305—317.
- [7] Viscarra Rossel R A, Brus D J, Lobsey C, et al. Baseline estimates of soil organic carbon by proximal sensing: Comparing design-based, model-assisted and model-based inference[J]. *Geoderma*, 2016, 265: 152—163.
- [8] Zeng L. Classification of purple soil images based on multi-feature fusion[D]. Chongqing: Chongqing Normal University, 2020. [曾莉. 多特征融合的紫色土图像分类[D]. 重庆: 重庆师范大学, 2020.]
- [9] Jin W D, Yin H J, Li Y X. Application of full-slice image segmentation based on K-means clustering in lab color space in spatial analysis of tumor biological effects[J]. *Biomedical Engineering and Clinical Medicine*, 2022, 26 (1): 1—8. [金文东, 阴慧娟, 李迎新. 基于 Lab 色彩空间 K 均值聚类的全切片图像分割在肿瘤生物学效应空间分析中应用[J]. 生物医学工程与临床, 2022, 26 (1): 1—8.]
- [10] Zhang H X, Peng Q X. PSO and K-means-based semantic segmentation toward agricultural products[J]. *Future Generation Computer Systems*, 2022, 126: 82—87.
- [11] Ci E, Tang J, Lian M S, et al. Higher category partition for purple soils in Chongqing in Chinese soil taxonomy[J]. *Acta Pedologica Sinica*, 2018, 55 (3): 569—584. [慈恩, 唐江, 连茂山, 等. 重庆市紫色土系统分类高级单元划分研究[J]. 土壤学报, 2018, 55 (3): 569—584.]
- [12] National Soil Census Office. *Soil of China* [M]. Beijing: China Agriculture Press, 1998. [全国土壤普查办公室. 中国土壤[M]. 北京: 中国农业出版社, 1998.]
- [13] Liu Y, Liu G D, Liu C Y, et al. A novel color-texture descriptor based on local histograms for image segmentation[J]. *IEEE Access*, 2019, 7: 160683—160695.
- [14] Liu J F, Li D B. The self-adaptive K-means method of image segmentation based on Lab color space[J]. *Machine Design and Manufacturing Engineering*, 2018, 47 (12): 23—27. [刘家丰, 李东波. 基于 Lab 色彩空间的自适应 K-means 图像分割方法[J]. 机械设计与制造工程, 2018, 47 (12): 23—27.]
- [15] Liu G Y, Huang Y, Cao Y, et al. Research on extraction of Image texture Feature based on Gray co-occurrence matrix[J]. *Technology Wind*, 2021 (12): 61—64. [刘光宇, 黄懿, 曹禹, 等. 基于灰度共生矩阵的图像纹理特征提取研究[J]. 科技风, 2021 (12): 61—64.]
- [16] Zhang G L, Li D C. *Manual of soil description and sampling*[M]. Beijing: Science Press, 2022. [张甘霖, 李德成. 野外土壤描述与采样手册[M]. 北京: 科学出版社, 2022.]
- [17] Liu M F, Fu X Y, Shang Y Y, et al. Pedestrian tracking based on HSV color feature and contribution reconstruction[J]. *Laser & Optoelectronics Progress*, 2017, 54 (9): 091004. [刘梦飞, 付小雁, 尚媛园, 等. 基于 HSV 颜色特征和贡献度重构的行人跟踪[J]. 激光与光电子学进展, 2017, 54 (9): 091004.]
- [18] Qu L, Wang K R, Chen L L, et al. Fast road detection based on RGBD image and convolutional neural network[J]. *Acta Optica Sinica*, 2017, 37 (10): 1010003. [曲磊, 王康如, 陈利利, 等. 基于 RGBD 图像和卷积神经网络的快速道路检测[J]. 光学学报, 2017, 37 (10): 1010003.]
- [19] Huang Y, De Bortoli V, Zhou F G, et al. Review of wavelet-based unsupervised texture segmentation, advantage of adaptive wavelets[J]. *IET Image Processing*, 2018, 12 (9): 1626—1638.
- [20] Fan Z S, Herrick J E, Saltzman R, et al. Measurement of soil color: A comparison between smartphone camera and the munsell color charts[J]. *Soil Science Society of America Journal*, 2017, 81 (5): 1139—1146.
- [21] Liedgren L, Hörnberg G, Magnusson T, et al. Heat impact and soil colors beneath hearths in northern Sweden[J]. *Journal of Archaeological Science*, 2017, 79: 62—72.
- [22] Yu X X, Yuan D G, Chen J K, et al. Taxonomy of soils developed from rocks of shaximiao formation (J2s) strata in Sichuan Basin[J]. *Acta Pedologica Sinica*, 2021, 58 (6): 1448—1459. [余星兴, 袁大刚, 陈剑科, 等. 四川盆地沙溪庙组地层 (J2s) 岩石发育土壤的系统分类研究[J]. 土壤学报, 2021, 58 (6): 1448—1459.]
- [23] Yu X X, Yuan D G, Chen J K, et al. Prediction of soil free iron oxide content based on soils munsell color[J]. *Acta Pedologica Sinica*, 2021, 58 (5): 1322—1329. [余星兴, 袁大刚, 陈剑科, 等. 基于 Munsell 颜色的土壤游离铁预测研究[J]. 土壤学报, 2021, 58 (5): 1322—1329.]
- [24] Zhang L, Zhang F R, Jin D S, et al. The correlation of soil color relative to clay and iron oxides in alfisol of the loess plateau and its genetic explanation—A case study of

- the 31 argillic horizons covered in the soil series survey of Shanxi Province[J]. *Acta Pedologica Sinica*, 2021, 58 (4): 876—886. [张蕾, 张凤荣, 靳东升, 等. 黄土高原淋溶土黏粒、氧化铁与颜色的关系及发生学解释——以山西土系调查的 31 个黏化层为例[J]. *土壤学报*, 2021, 58 (4): 876—886.]
- [25] Yan Z M, Yuan D G, Yu X X, et al. Proposal on revising color definition for “L.C. of purplish sandstones and shales”—Based on color characteristics of parent rock of purplish soil in Sichuan[J]. *Acta Pedologica Sinica*, 2020, 57 (1): 60—70. [晏昭敏, 袁大刚, 余星兴, 等. “紫色砂、页岩岩性特征” 颜色修订建议——基于四川紫色土母岩颜色特征[J]. *土壤学报*, 2020, 57 (1): 60—70.]
- [26] Chinese Soil Taxonomy Research Group, Institute of Soil Science, Chinese Academy of Sciences, Cooperative Research Group on Chinese Soil Taxonomy. Keys to Chinese Soil Taxonomy[M]. 3rd ed. Hefei: University of Science and Technology of China Press, 2001. [中国科学院南京土壤研究所土壤系统分类课题组, 中国土壤系统分类课题研究协作组. 中国土壤系统分类检索[M]. 第 3 版. 合肥: 中国科技大学出版社, 2001.]
- [27] Yang C, Liu B Y. Image foreground and background separation based on texture features of Lab color space[J]. *Laser & Optoelectronics Progress*, 2019, 56 (12): 121001. [杨超, 刘本永. 基于 Lab 颜色空间纹理特征的图像前后景分离[J]. *激光与光电子学进展*, 2019, 56 (12): 121001.]
- [28] Haralick R, Shanmugam K, Dinstein I. Textural feature for image classification[J]. *IEEE Transactions on Systems*, 1973, 3 (6): 610—621.
- [29] Hu J, Ci E, Wang S F, et al. Genetic characteristics and classification of soil derived from Pleistocene sediments under cultivation in Chongqing[J]. *Acta Pedologica Sinica*, 2020, 57 (2): 273—283. [胡瑾, 慈恩, 王术芳, 等. 重庆市更新统沉积物发育土壤的特性及分类[J]. *土壤学报*, 2020, 57 (2): 273—283.]
- [30] Weng H L, Ci E, Li S, et al. Pedogenetic process and taxonomy of yellow soil in Chongqing, China[J]. *Acta Pedologica Sinica*, 2020, 57 (3): 579—589. [翁昊璐, 慈恩, 李松, 等. 重庆市黄壤的成土特点与系统分类研究[J]. *土壤学报*, 2020, 57 (3): 579—589.]
- [31] Zhang Y K, Hartemink A E. A method for automated soil horizon delineation using digital images[J]. *Geoderma*, 2019, 343: 97—115.
- [32] Huang P, Zheng Q, Liang C. Overview of image segmentation methods[J]. *Journal of Wuhan University (Natural Science Edition)*, 2020, 66 (6): 519—531. [黄鹏, 郑淇, 梁超. 图像分割方法综述[J]. *武汉大学学报 (理学版)*, 2020, 66 (6): 519—531.]
- [33] Achanta R, Shaji A, Smith K, et al. SLIC superpixels compared to state-of-the-art superpixel methods[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34 (11): 2274—2282.
- [34] Khan J F, Bhuiyan S M A, Adhami R R. Image segmentation and shape analysis for road-sign detection[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2011, 12 (1): 83—96.
- [35] MAC Q J. Some methods for classification and analysis of multivariate observations[M]. Berkeley: Berkeley Symposium on Mathematical Statistics and Probability, 1967.
- [36] Ahmed M, Seraj R, Islam S M S. The k-means algorithm: A comprehensive survey and performance evaluation[J]. *Electronics*, 2020, 9 (8): 1295.
- [37] Lin G C, Tang Y C, Zou X J, et al. Color-, depth-, and shape-based 3D fruit detection[J]. *Precision Agriculture*, 2020, 21 (1): 1—17.
- [38] Navarro A, Young M, Allan B, et al. The application of Unmanned Aerial Vehicles (UAVs) to estimate above-ground biomass of mangrove ecosystems[J]. *Remote Sensing of Environment*, 2020, 242: 111747.
- [39] Lei T, Jia X H, Zhang Y N, et al. Significantly fast and robust fuzzy C-means clustering algorithm based on morphological reconstruction and membership filtering[J]. *IEEE Transactions on Fuzzy Systems*, 2018, 26 (5): 3027—3041.

(责任编辑: 檀满枝)