

基于光谱的土壤剖面识别系统

用户手册

1 引言

1.1 编写目的

为了让本系统的用户正确安装、运行以及维护系统，特编写本用户手册。

1.2 背景

该软件全称“基于光谱的土壤剖面识别系统”，主要用于基于土壤光谱数据的土壤剖面识别应用。

本软件由解宪丽（中国科学院南京土壤研究所）、徐诗宇（南京师范大学地理科学学院）、李安波（南京师范大学地理科学学院）共同开发。

1.3 定义

1.3.1 术语

表 1-1 术语表

术语	解释
土壤剖面	由与地表大致平行的层次组成的从地表到母质的三维垂直断面。不同类型的土壤，具有不同形态的土壤剖面。土壤剖面可以表示土壤特征的外部表现，包括土壤的若干发生层次、颜色、质地、结构、新生体等。
土层	土壤剖面中与地面大致平行的一层土壤。由成土作用形成的，具有发生学特征的土壤层次，称为土壤发生层。非成土作用产生的土层，称为土壤层次。
反射光谱曲线	反射光谱曲线系指地物反射电磁辐射的能力，随所反射的电磁波波长而变化的特性。如以横坐标表示波长的变化，纵坐标表示其反射率（或反射亮度系数）可构成反映反射光谱特性的曲线，称为反射光谱（特性）曲线。不同性质的地物，或相同属

	性的地物在其成份、颜色、表面结构、含水性（率）等不同时，其反射光谱特性也不同，构成反射光谱曲线的差异。遥感探测即是根据获取和记录不同地物不同波段的反射电磁波信息，通过分析其差异性，来识别地物属性的。
光谱角分类法	将光谱数据视为多维空间的矢量，利用解析方法计算待测剖面光谱与光谱数据库中案例剖面光谱之间矢量的夹角，根据夹角的大小来确定光谱间的相似程度，以达到识别地物的目的。
归一化	归一化是一种简化计算的方式，即将有量纲的表达式，经过变换，化为无量纲的表达式，成为标量。
数据平滑	测试数据带有误差，这些误差是随机性的，但其分布是有规律性的。数据平滑是指为了产生更准确的概率来调整最大似然估计的技术，基本思想就是提高低概率（如零概率），降低高概率，尽量使误差的概率分布趋于平均。
K-近邻算法（KNN）	KNN 是通过测量不同特征值之间的距离进行分类。它的思路是：如果一个样本在特征空间中的 k 个最相似(即特征空间中最邻近)的样本中的大多数属于某一个类别，则该样本也属于这个类别。 K 通常是不大于 20 的整数。 KNN 算法中，所选择的邻居都是已经正确分类的对象。该方法在定类决策上只依据最邻近的一个或者几个样本的类别来决定待分样本所属的类别。

1.3.2 英文缩写

表 1-2 相关英文缩写表

缩写	全称
SA	Spectra Angle
KNN	k-Nearest Neighbor

1.4 参考资料

1. 中国科学院南京土壤研究所土壤系统分类课题组，中国土壤系统分类检索(第三版)，中国科学技术大学出版社，2001.
2. 许禄，邵学广. 化学计量学方法（第二版），科学出版社，2004.

3. 薛毅, 陈立萍. 统计建模与 R 软件, 清华大学出版社, 2007.
4. 刘万军等. 基于光谱信息散度与光谱角匹配的高光谱解混算法, 计算机应用, 2015.
5. 应义斌等. 基于光谱曲线波形相似度的光谱匹配方法, 中国发明专利, 2014.
6. Brett Lantz 著, 李洪成等译. 机器学习与 R 语言, 机械工业出版社, 2015.
7. 周健民、沈仁芳主编, 土壤学大辞典, 科学出版社, 2013.

2 软件概述

2.1 目标

本系统的目的是将已经采集好的土壤剖面的反射光谱数据，通过一定的算法计算与已知土壤样本剖面的反射光谱数据的相似度，从而判断未知土壤剖面的土壤类型。

2.2 功能

表 2-1 软件功能描述表

功能描述	输入	处理		输出
数据读取	CSV 格式土壤光谱数据	打开 CSV 数据并逐条读入		无
数据查看	CSV 格式样本类型数据	读入内存，并在主界面显示		无
数据预处理	土壤光谱数据	除噪平滑	11 点平滑法	将平滑后土壤光谱数据存为 CSV 文件
			5 点平滑法	
		归一化	Max-Min 标准法	将归一化后土壤光谱数据存为 CSV 文件
			Z-Score 标准法	
		一阶导		将一阶导后土壤光谱数据存为 CSV 文件
可视化	土壤光谱数据	绘制光谱曲线		JPG 格式图像
		绘制直方图		JPG 格式图像
光谱匹配	集合	光谱角匹配算法		报告
		波形相似度匹配算法		
		直方图匹配算法		
		峰谷特征值匹配算法		
样本训练	光谱样本集合	光谱角匹配算法		XLSX 格式训练报告
		波形相似度匹配算法		
		直方图匹配算法		
		峰谷特征值匹配算法		
帮助	无	使用帮助文档		帮助文档
关于	无	查看制作团队信息		关于文档

2.3 性能

2.3.1 数据精确度

表 2-2 软件数据精度要求表

数据	精度要求
土壤剖面采样深度	1cm
直方图组值	0.01
直方图组数	1
直方图距离值	0.00001
光谱角值	0.00001
波形相似度值	0.00001
峰谷特征值	0.00001
光谱匹配率精度	0.0001

2.3.2 灵活性

表 2-3 软件时间特性要求表

时间 特性 要求	系统开启速度由当前计算机状态与数据库连接状况决定
	数据导入数据库的时间有导入数据的数据量大小决定
	数据保存的时间由要保存的数据量大小决定
	匹配响应时间由匹配条件的复杂程度和当前计算机的性能决定
	匹配时间由测试集、训练集的数据量大小以及当前计算机的计算速度决定
	系统清空反应时间由当前系统内存中数据量决定
	邮件发送时间由当前服务器状态与网络传输速度决定

3 运行环境

3.1 硬件

3.1.1 处理器型号及内存容量

表 3-1 软件处理器及内存型号要求表

处理器	1.6GHz 或更高版本的处理器（推荐 2.3GHz 以上）
运行内存	2 GB 及以上（推荐 4GB 以上）

3.1.2 物理存储需求

外存容量、联机或脱机、媒体及其存储格式，设备的型号数量（见表 3-2）。

表 3-2 软件物理存储需求表

项目	规格及要求
外存容量	50 MB 以上 可用硬盘空间
联机或脱机	脱机（邮件功能需联机）
媒体及其存储格式	本软件无媒体设备支持
设备的型号数量	常用计算机*1

3.1.3 输入及输出设备

输入及输出设备的型号和数量，联机或脱机（见表 3-3）。

表 3-3 软件输入及输出设备表

设备	设备名称	型号	数量	工作状态
输入设备	键盘	任何键盘均可	1	脱机
输出设备	显示器	1024*768 分辨率及以上	1	脱机

3.2 软件环境

表 3-4 相关支持软件表

软件名称	版本要求
Net Framework	4.0 以上
R	3.3.3 及以上版本

4 使用说明

4.1 安装和初始化

4.1.1 软件安装

- 1) 双击基于光谱的土壤剖面识别系统.exe（图 4-1）文件，执行安装程序；



图 4-1 系统图标



图 4-2 安装界面图

- 2) 点击【下一步】，选择【典型安装】，点击【浏览】选择安装位置（图 4-3）；

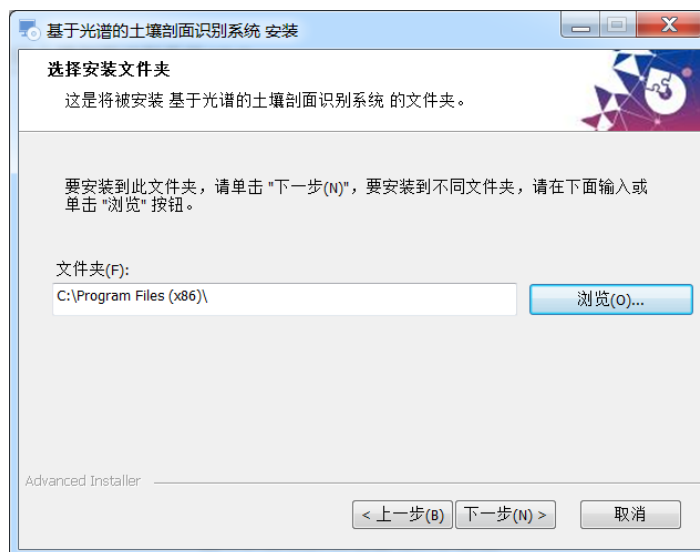


图 4-3 系统安装位置选择示意图

3) 点击【下一步】，点击【安装】（图 4-4）；

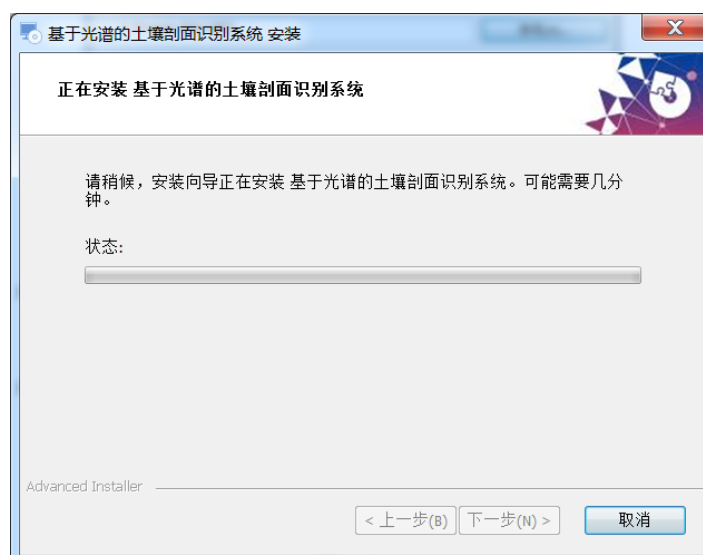


图 4-4 系统安装示意图

4) 点击【完成】按钮（图 4-5），完成安装；



图 4-5 系统安装完成界面图

4.1.2 初次运行准备


- 1) 安装 R 语言环境：安装教程参见链接：<http://www.pc0359.cn/downinfo/44174.html>；
- 2) 双击桌面图标运行软件；
- 3) 菜单栏【配置】->【配置界面】（图 4-7），更改数据路径以及光谱波段范围；



图 4-6 系统配置界面

- 4) 重启软件；

4.2 输入

4.2.1 测试集合

1) 数据格式

输入测试集合数据要求为 CSV 格式文本文件；

输入测试集数据前三个字段分别为 ProfileID（剖面代码）、UpperDepth（土层上表面深度）、LowerDepth（土层下表面深度），土壤库后续字段为土壤光谱反射率的波段重采样位置（在 350nm~2590nm 范围内，采样步长为 1nm）。

2) 输入举例（表 4-1）

表 4-1 土壤剖面光谱数据

Id	UpperDepth	LowerDepth	360nm	361nm	362nm	363nm	364nm	2488nm	2489nm	2490nm
100										
32-036	0	1	5.70E-02	5.89E-02	6.07E-02	6.24E-02	6.39E-02	#####	#####	#####
32-036	1	2	5.70E-02	5.89E-02	6.07E-02	6.24E-02	6.39E-02	#####	#####	#####
32-036	2	3	5.69E-02	5.88E-02	6.06E-02	6.23E-02	6.38E-02	#####	#####	#####
32-036	3	4	5.68E-02	5.87E-02	6.05E-02	6.22E-02	6.37E-02	#####	#####	#####
32-036	4	5	5.66E-02	5.86E-02	6.04E-02	6.21E-02	6.36E-02	#####	#####	#####
32-036	5	6	5.64E-02	5.84E-02	6.02E-02	6.19E-02	6.34E-02	#####	#####	#####
32-036	6	7	5.62E-02	5.81E-02	6.00E-02	6.17E-02	6.32E-02	#####	#####	#####
32-036	7	8	5.59E-02	5.78E-02	5.97E-02	6.14E-02	6.30E-02	#####	#####	#####
32-036	8	9	5.55E-02	5.75E-02	5.94E-02	6.11E-02	6.27E-02	#####	#####	#####
32-036	9	10	5.51E-02	5.71E-02	5.90E-02	6.08E-02	6.24E-02	#####	#####	#####
32-036	10	11	5.47E-02	5.67E-02	5.86E-02	6.04E-02	6.20E-02	#####	#####	#####
32-036	11	12	5.42E-02	5.62E-02	5.82E-02	6.00E-02	6.16E-02	#####	#####	#####
32-036	12	13	5.37E-02	5.57E-02	5.77E-02	5.96E-02	6.12E-02	#####	#####	#####
32-036	13	14	5.31E-02	5.52E-02	5.72E-02	5.91E-02	6.08E-02	#####	#####	#####
32-036	14	15	5.25E-02	5.46E-02	5.66E-02	5.85E-02	6.03E-02	#####	#####	#####
32-036	15	16	5.18E-02	5.40E-02	5.60E-02	5.80E-02	5.97E-02	#####	#####	#####
32-036	16	17	5.11E-02	5.33E-02	5.54E-02	5.73E-02	5.91E-02	#####	#####	#####
32-036	17	18	5.04E-02	5.26E-02	5.47E-02	5.67E-02	5.85E-02	#####	#####	#####
32-036	18	19	4.96E-02	5.18E-02	5.40E-02	5.60E-02	5.79E-02	#####	#####	#####
32-036	19	20	4.88E-02	5.11E-02	5.33E-02	5.54E-02	5.73E-02	#####	#####	#####
32-036	20	21	4.82E-02	5.05E-02	5.27E-02	5.48E-02	5.68E-02	#####	#####	#####
32-036	21	22	4.77E-02	5.00E-02	5.23E-02	5.44E-02	5.64E-02	#####	#####	#####
32-036	22	23	4.74E-02	4.97E-02	5.20E-02	5.41E-02	5.61E-02	#####	#####

4.2.2 训练集合

数据格式与要求与测试集合相同。

4.2.3 类型数据

1) 数据格式

类型数据文件格式要求为 CSV 文本文件；

类型数据六个字段分别为 ProfileID（剖面代码）、OrderCST（中国土壤系统分类土纲）、SuborderCST（中国土壤系统分类亚纲）、GroupCST（中国土壤系统分类土类）、SubgroupCST（中国土壤系统分类亚类）、Depth（剖面采样深度）。

2) 输入举例（表 4-2）

表 4-2 土壤剖面类型数据

ProfileID	OrderCST	SuborderCST	GroupCST	SubgroupCST	Depth
32-036	人为土	旱耕人为土	肥熟旱耕人为土	斑纹肥熟旱耕人为土	100
32-037	人为土	水耕人为土	铁渗水耕人为土	底潜铁渗水耕人为土	100
32-038	人为土	水耕人为土	铁聚水耕人为土	普通铁聚水耕人为土	100
32-043	人为土	水耕人为土	潜育水耕人为土	普通潜育水耕人为土	100
32-044	人为土	水耕人为土	潜育水耕人为土	铁聚潜育水耕人为土	100
32-045	人为土	水耕人为土	潜育水耕人为土	铁聚潜育水耕人为土	100
32-048	人为土	水耕人为土	简育水耕人为土	普通简育水耕人为土	100
32-050	雏形土	潮湿雏形土	淡色潮湿雏形土	水耕淡色潮湿雏形土	100
32-050a	人为土	水耕人为土	简育水耕人为土	普通简育水耕人为土	100
32-051	雏形土	潮湿雏形土	淡色潮湿雏形土	水耕淡色潮湿雏形土	100
32-053	淋溶土	湿润淋溶土	黏磐湿润淋溶土	表蚀黏磐湿润淋溶土	100
32-057	人为土	水耕人为土	简育水耕人为土	普通简育水耕人为土	100
32-058	人为土	水耕人为土	简育水耕人为土	普通简育水耕人为土	100
32-059	雏形土	湿润雏形土	简育湿润雏形土	普通简育湿润雏形土	65
32-060	雏形土	湿润雏形土	铁质湿润雏形土	普通铁质湿润雏形土	120
32-060a	淋溶土	湿润淋溶土	简育湿润淋溶土	斑纹简育湿润淋溶土	100
32-061	人为土	水耕人为土	简育水耕人为土	普通简育水耕人为土	100
32-065	人为土	水耕人为土	铁聚水耕人为土	普通铁聚水耕人为土	100
32-066	人为土	水耕人为土	铁渗水耕人为土	漂白铁渗水耕人为土	100
32-068	雏形土	湿润雏形土	简育湿润雏形土	普通简育湿润雏形土	65
32-069	人为土	水耕人为土	简育水耕人为土	普通简育水耕人为土	120
32-070	人为土	水耕人为土	铁渗水耕人为土	普通铁渗水耕人为土	100
32-071	人为土	水耕人为土	简育水耕人为土	底潜简育水耕人为土	100
32-072	人为土	水耕人为土	铁聚水耕人为土	普通铁聚水耕人为土	100
32-074	淋溶土	湿润淋溶土	黏磐湿润淋溶土	表蚀黏磐湿润淋溶土	100
32-074a	淋溶土	湿润淋溶土	铁质湿润淋溶土	红色铁质湿润淋溶土	100
32-075	雏形土	湿润雏形土	酸性湿润雏形土	普通酸性湿润雏形土	80
32-075a	淋溶土	湿润淋溶土	黏磐湿润淋溶土	普通黏磐湿润淋溶土	100
32-076	人为土	水耕人为土	铁聚水耕人为土	底潜铁聚水耕人为土	100
32-079	淋溶土	湿润淋溶土	酸性湿润淋溶土	红色酸性湿润淋溶土	100
32-084	人为土	水耕人为土	铁聚水耕人为土	普通铁聚水耕人为土	100
32-085	人为土	水耕人为土	简育水耕人为土	普通简育水耕人为土	100
32-086	人为土	水耕人为土	简育水耕人为土	普通简育水耕人为土	100

4.3 输出

4.3.1 预处理数据

- 1) 数据背景
预处理数据为测试集或训练集数据经过预处理后得到的数据。
- 2) 数据格式
预处理数据的数据格式与测试集、训练集数据相同。

4.3.2 匹配结果数据

- 1) 数据匹配
匹配结果数据为系统对测试集剖面的土壤类型判断结果。
- 2) 数据格式
匹配结果数据为 xls 格式的 EXCEL 表格文件；

文件共有三张 sheet 表格：

- (1) 表格一：匹配进行的时间与使用的匹配算法；
- (2) 表格二：匹配得到的测试集与训练集的相似度矩阵；
- (3) 表格三：系统判断的测试集各土壤剖面的土壤类型结果；

3) 输出举例

表格一：

匹配方法	直方图匹配
匹配时间	2017/6/28 星期三

表格二：

	34-002	34-003	34-006	34-013	34-014
33-001	0.142505147	0.199203934	0.12942923	0.156506135	0.18447898
33-002	0.252143521	0.201412964	0.265411437	0.288214784	0.155185296
33-003	0.14957993	0.186355184	0.143421664	0.172656129	0.132638751
33-004	0.255832835	0.352113635	0.226417945	0.253012552	0.330661818
33-005	0.368264299	0.457660151	0.355286947	0.388526009	0.436208334
33-006	0.175988798	0.264748179	0.155612893	0.173867194	0.243303004
33-007	0.266034071	0.220533639	0.27709205	0.319223617	0.166608554
33-008	0.148898851	0.115811251	0.143519293	0.16678909	0.064652653
33-009	0.24418244	0.210532311	0.239240663	0.254490934	0.15440692
33-010	0.22010416	0.284771756	0.200987913	0.161103363	0.280597618
33-011	0.126885037	0.186098692	0.115010228	0.108444909	0.178569038
33-012	0.132510903	0.106385181	0.143843949	0.179567643	0.119130637
33-013	0.118620269	0.1103287	0.113240711	0.143468452	0.091919935
33-014	0.239006702	0.203672566	0.250064682	0.280077854	0.143504538
33-016	0.289423192	0.24392276	0.300481172	0.340889509	0.189997675
33-017	0.100903234	0.092900312	0.111798499	0.136604791	0.116799495
33-018	0.266676629	0.362595471	0.237679374	0.199378584	0.344901375
33-019	0.165184521	0.236633017	0.145223152	0.139989152	0.220026123
33-020	0.199793784	0.16949027	0.194414226	0.202205176	0.110114565
33-021	0.119367515	0.08522116	0.116507826	0.128893538	0.068139957
33-022	0.253649096	0.24724971	0.248269539	0.210845085	0.184415148
33-023	0.333921653	0.355605366	0.335088331	0.356283899	0.335989905
33-024	0.297374643	0.388989839	0.26960406	0.245942972	0.376826061
33-025	0.100906555	0.155459255	0.09235018	0.156079675	0.13508335
33-026	0.141621837	0.217359036	0.140190609	0.157426778	0.206025437
33-027	0.128393217	0.22455115	0.119944212	0.111266078	0.20416805
33-028	0.101612207	0.143811848	0.096311793	0.070494344	0.150084678
33-030	0.302889575	0.212200859	0.330222377	0.396451263	0.223534458

表格三：

剖面编号	匹配类型
33-001	铝质湿润锥形土
33-002	酸性湿润淋溶土
33-003	酸性常湿锥形土
33-004	铁聚水耕人为土
33-005	湿润正常新成土
33-006	铝质湿润淋溶土
33-007	铝质湿润锥形土
33-008	铁聚水耕人为土
33-009	铝质湿润锥形土
33-010	铁聚水耕人为土
33-011	暗红湿润铁铝土
33-012	钙质湿润淋溶土
33-013	筒育湿润富铁土
33-014	铁质湿润淋溶土
33-016	铝质湿润淋溶土
33-017	淡色潮湿锥形土
33-018	淡色潮湿锥形土
33-019	铝质湿润淋溶土
33-020	黏化湿润富铁土
33-021	铁聚水耕人为土
33-022	筒育湿润富铁土
33-023	铁质湿润锥形土
33-024	筒育水耕人为土
33-025	筒育水耕人为土
33-026	铁聚水耕人为土

4.4 出错和恢复

4.4.1 配置文件出错

- 1) 启动程序或读取训练集数据时出现以下错误提示（图 4-8）；

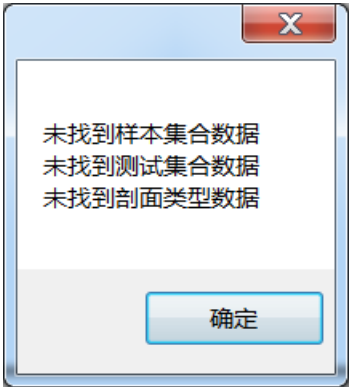


图 4-7 数据读取错误提示图

- 2) 错误原因
- a) 数据路径与配置文件不符：

解决方法：打开配置文件界面（图 4-9），检查并修改相应路径，然后重新启动程序；

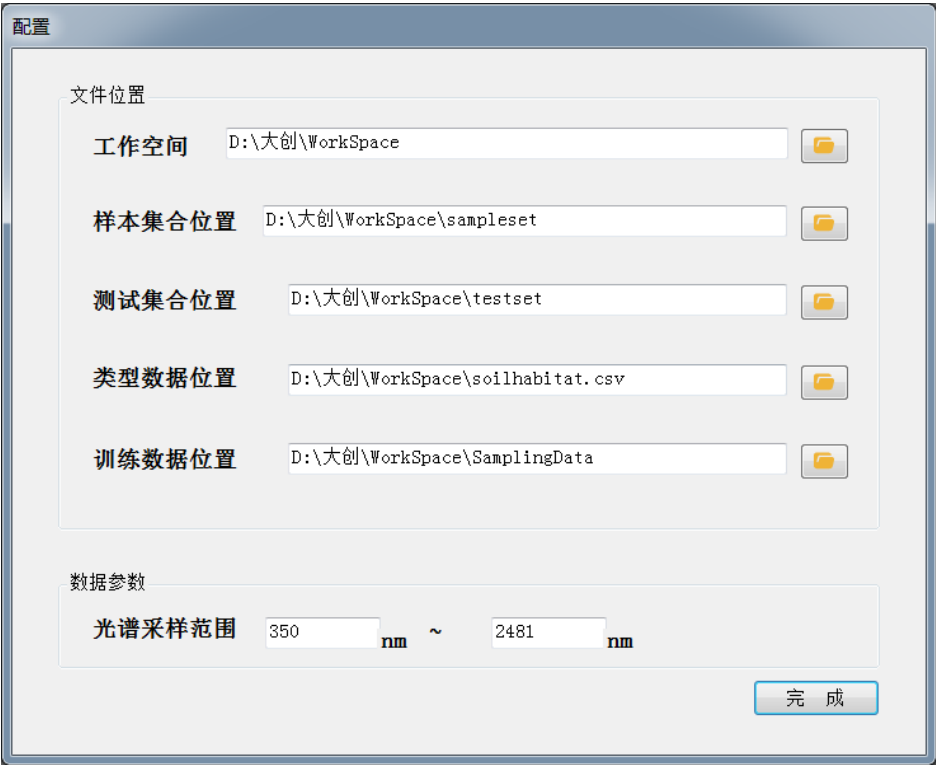


图 4-8 系统配置界面图

- b) 配置文件丢失
- 解决方法：检查程序安装路径下\bin\Debug 位置，按照以下格式重新创建组织 config.csv 文件（表 4-3）；

表 4-3 系统配置文件表示例

工作空间	D:\大创\Workspace			
样本集合	D:\大创\Workspace\sampleset			
测试集合	D:\大创\Workspace\testset			
类型数据	D:\大创\Workspace\soilhabitat.csv			
训练位置	D:\大创\Workspace\SamplingData			
波长最小	350			
波长最大	2481			

4.4.2 计算机环境问题

- 1) 计算机安装过程中出错
- 错误原因：计算机的系统环境、物理内存不符合本软件安装要求；
- 解决方法：参照本文档第 3 节运行环境修改计算机系统环境；

5 操作说明

5.1 数据配置

5.1.1 读取测试集数据

- 1) 运行目的：读取测试集数据；
- 2) 操作步骤：
 - a) 菜单栏【文件】->【数据读取】（图 5-1）；



图 5-1 菜单栏数据读取截图

- b) 选择一个测试集合光谱文件夹（图 5-2）；

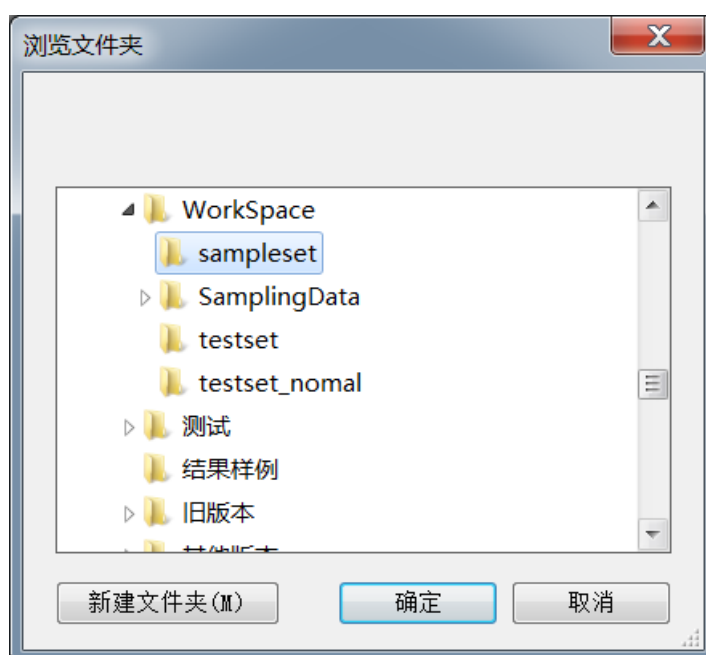


图 5-2 浏览文件夹界面图

5.1.2 更改配置文件

- 1) 操作目的：更改系统默认的；
- 2) 操作要求：新的训练集数据符合数据要求；训练集的 CSV 文件存入一个单独文件夹；
- 3) 操作步骤：
 - a) 菜单栏【配置】->【配置界面】（图 5-3）；

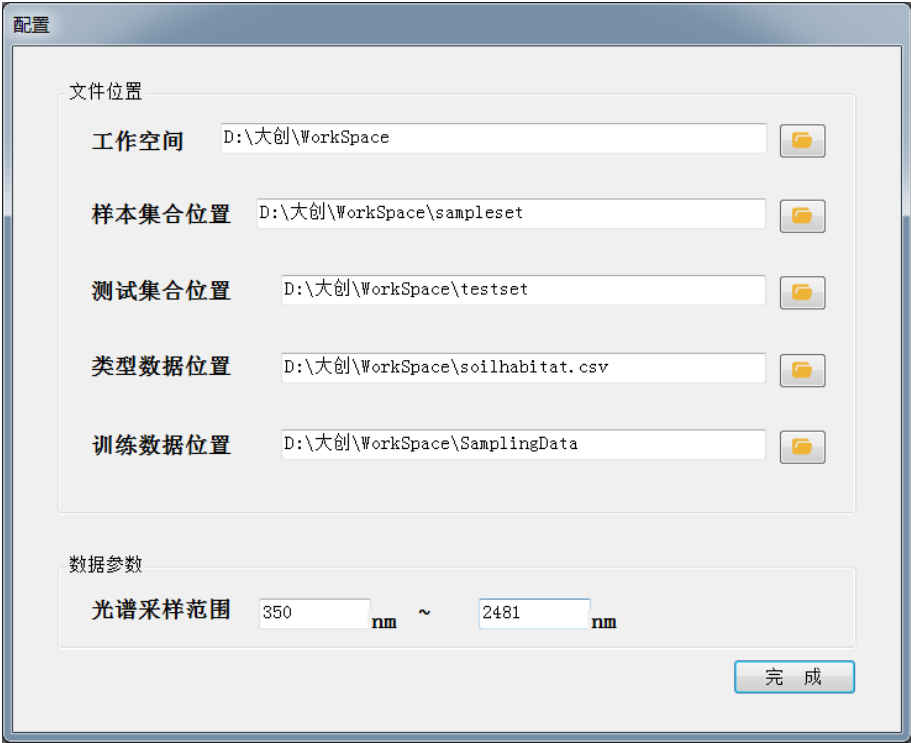


图 5-3 系统配置界面图

b) 将新的数据路径或者光谱采样范围输入文本框，点击【完成】按钮；

5.2 数据可视化

5.2.1 显示直方图

- 1) 操作目的：显示某个剖面的直方图，或多个剖面的直方图对比图；
- 2) 操作要求：系统已成功读取测试集或训练集数据；
- 3) 操作步骤：
 - a) 菜单栏【可视化】->【直方图显示】（图 5-4）；

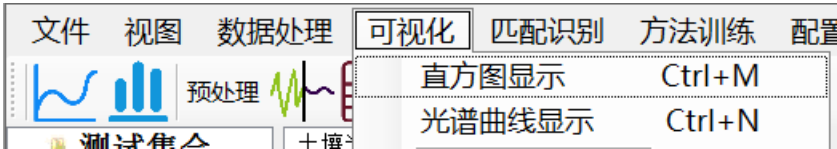


图 5-4 菜单栏直方图显示项截图

b) 直方图界面（图 5-5）选择【样本集合】->选择【剖面编号】->选择【组数】->选择【颜色】->【显示】；

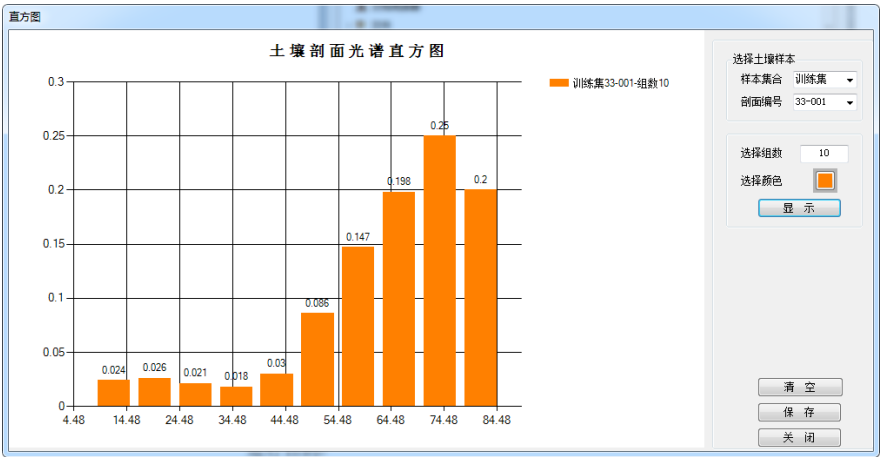


图 5-5 土壤光谱数据直方图界面

c) 【保存】->【文件名】->【保存图像】(图 5-6);

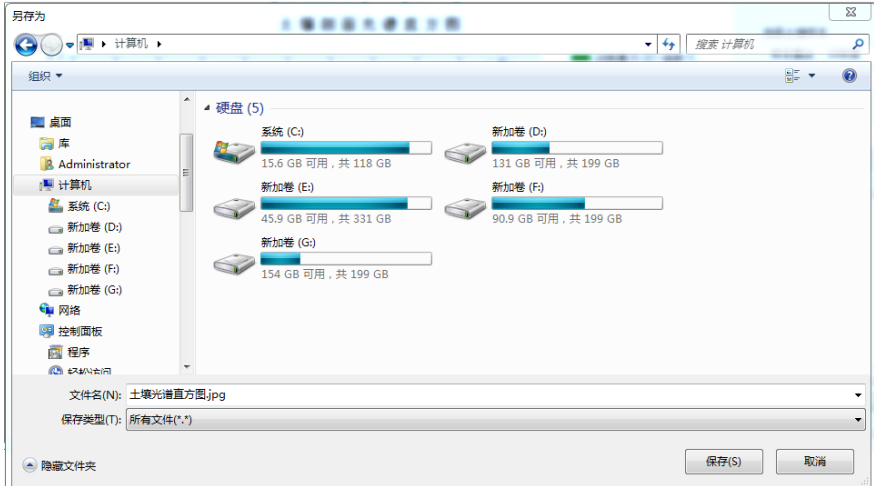


图 5-6 图像保存界面图

5.2.2 显示曲线图

- 1) 操作目的：显示某土壤剖面某个土层的光谱曲线图，或多个剖面多个土层的光谱曲线对比图；根据数据看出是 1cm 间隔土层的光谱数据，这是基于发生层光谱数据进行深度插值（二次样条）得来的；
- 2) 操作要求：与显示直方图要求相同；
- 3) 操作步骤：
 - a) 菜单栏【可视化】->【光谱曲线显示】(图 5-7)；

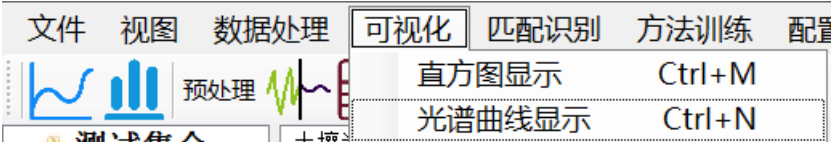


图 5-7 菜单栏光谱曲线项截图

- b) 曲线图界面 (图 5-8)【样本集合】->【剖面编号】->【采样深度】->选择【颜色】->【显示】；

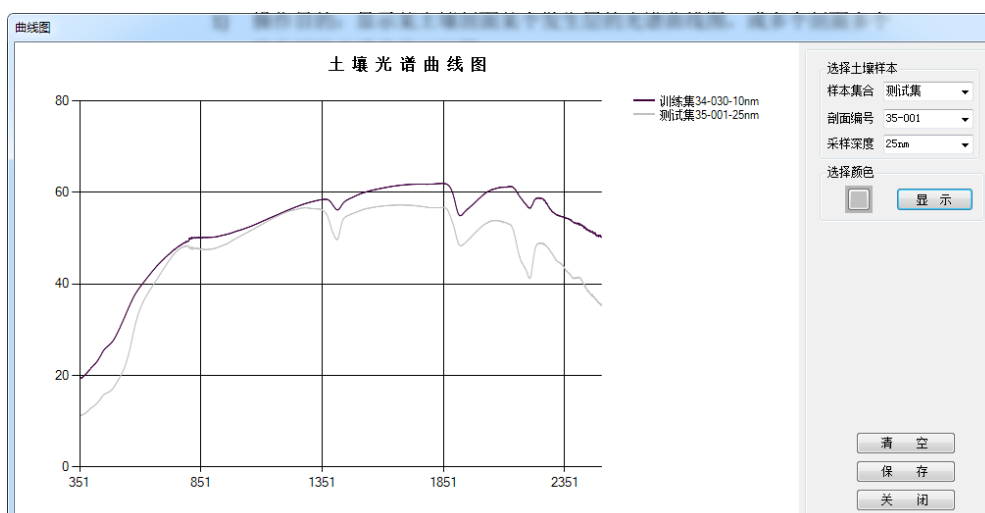


图 5-8 土壤光谱曲线界面图

c) 【保存】->【文件名】->【保存图像】(图 5-9);

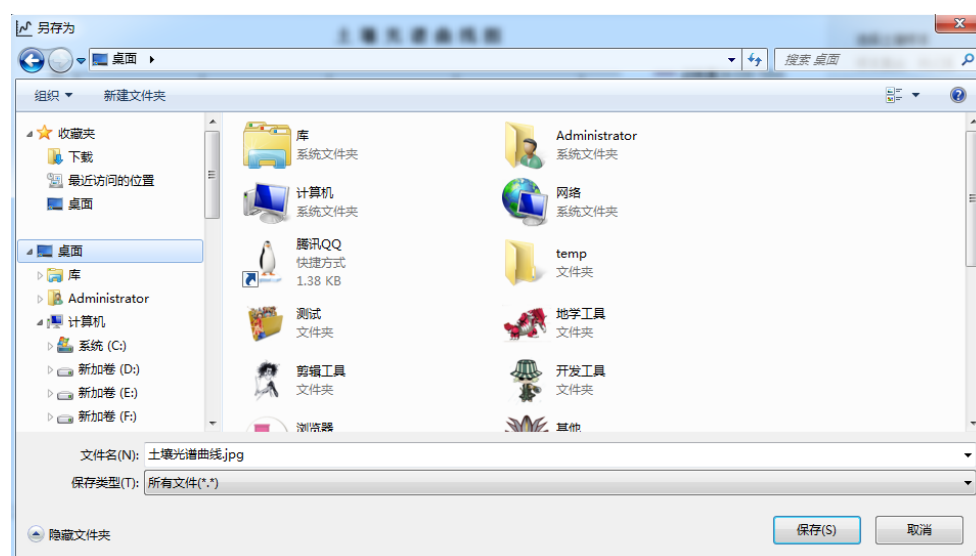


图 5-9 图像文件保存界面

5.3 数据预处理

5.3.1 归一化处理

- 1) 操作目的: 基于 Max-Min 标准化法和 Z-score 标准化法, 为测试集光谱数据作归一化处理;
- 2) 操作要求: 系统已导入测试集数据;
- 3) 操作步骤:
 - a) 菜单【预处理】->【归一化】(图 5-10);

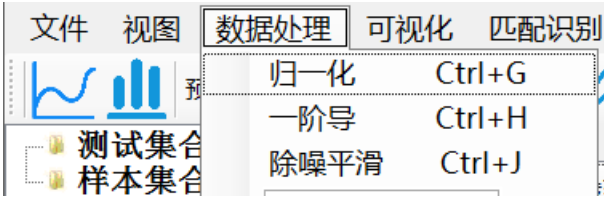


图 5-10 菜单栏归一化项截图

b) 弹出【归一化工具】界面（图 5-11）；

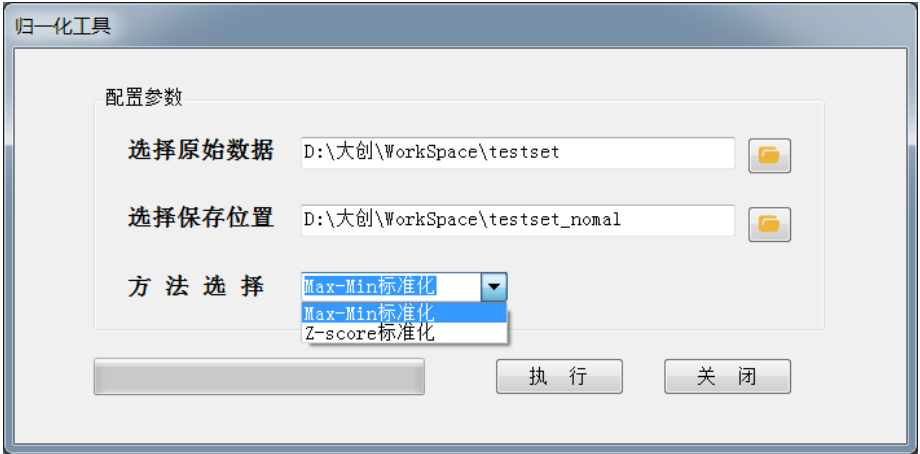



图 5-11 归一化界面图

- c) 点击按钮，浏览文件，选择原始光谱数据与归一化后数据保存位置，并设置归一化方法；
- d) 选择归一化方法
- e) 点击【执行】按钮（图 5-12）；

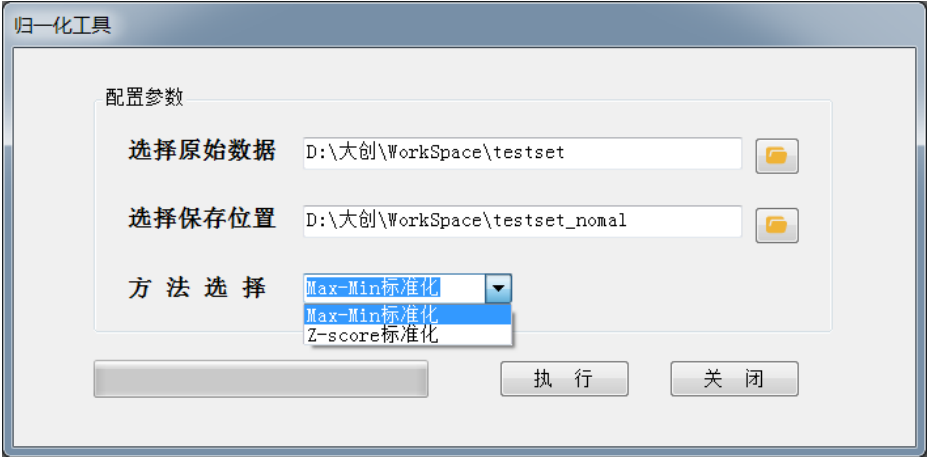


图 5-12 归一化执行状态图

● min-max 标准化

也称为离差标准化，是对原始数据的线性变换，使结果值映射到[0, 1]之间。转换函数如下：

$$x^* = \frac{x - \min}{\max - \min}$$

其中 max 为样本数据的最大值，min 为样本数据的最小值。

这种方法有个缺陷就是当有新数据加入时，可能导致 max 和 min 的变化，需要重新定义。使用这种方法的情况包括：1、对于方差非常小的属性可以增强其稳定性；2、维持稀疏矩阵中为 0 的条目。

- Z-score 标准化方法

也称为均值归一化(mean normalization)，给予原始数据的均值（mean）和标准差（standard deviation）进行数据的标准化。经过处理的数据符合标准正态分布，即均值为 0，标准差为 1。转化函数为：

$$x^* = \frac{x - \mu}{\sigma}$$

其中 μ 为所有样本数据的均值， σ 为所有样本数据的标准差。

5.3.2 平滑处理

- 1) 操作目的：基于 11 点平均值法或 5 点平均值法，为测试集光谱数据去除噪声影响；
- 2) 操作要求：与归一化处理要求相同；
- 3) 操作步骤：
 - a) 菜单【预处理】->【除噪平滑】（图 5-13）；

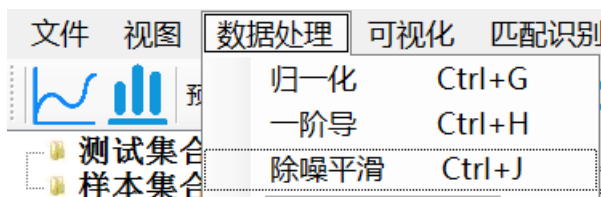


图 5-13 菜单栏除噪平滑项截图

- b) 弹出【除噪平滑工具】界面（图 5-14）；

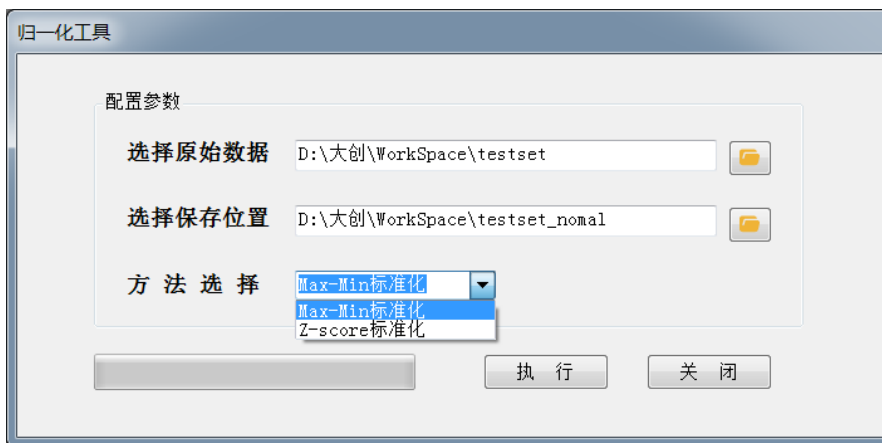



图 5-14 除噪平滑工具界面图

- c) 点击按钮 ，浏览文件，选择原始光谱数据与平滑处理后数据保存位置；
- d) 点击【执行】按钮（图 5-15）；

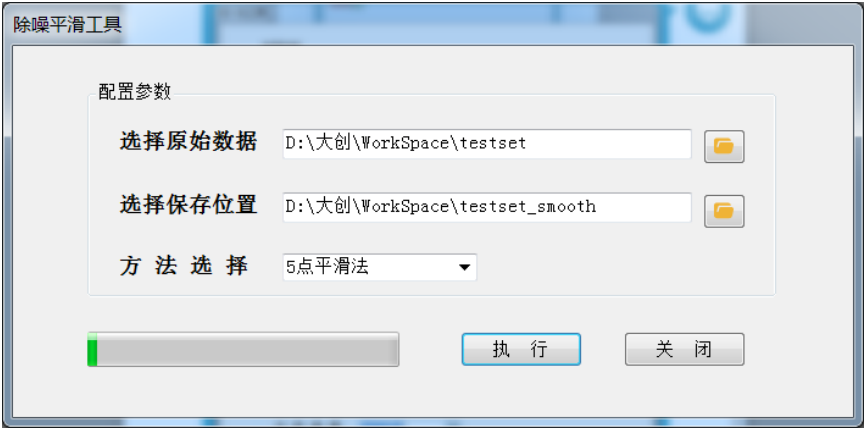


图 5-15 归一化执行状态图

5.3.3 一阶导处理

- 1) 操作目的：一阶导处理。离散值的一阶求导，其实就是求差分；
- 2) 操作要求：与归一化处理要求相同；
- 3) 操作步骤：
 - a) 菜单【预处理】->【一阶导】（图 5-17）；

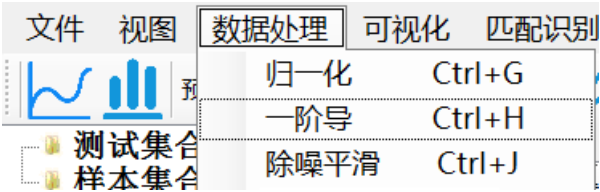


图 5-17 菜单栏一阶导项截图

- b) 弹出【一阶导工具】界面（图 5-18）；

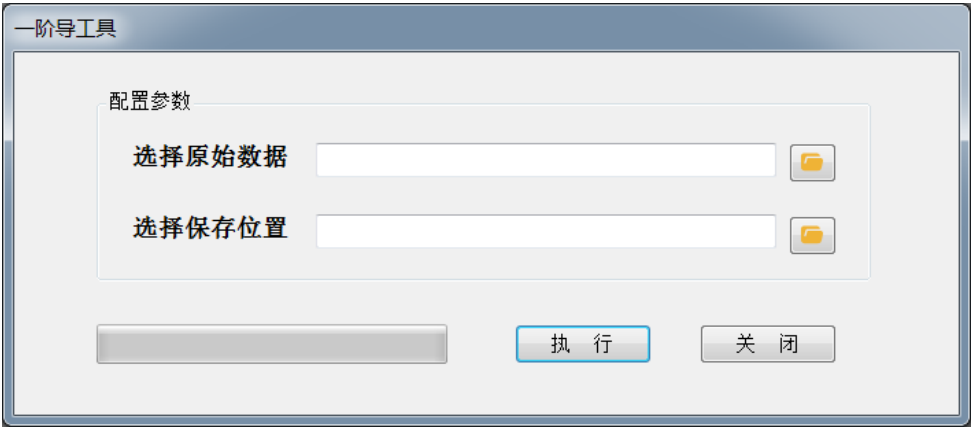



图 5-18 一阶导工具界面图

- c) 点击按钮，浏览文件，选择原始光谱数据与一阶导计算后数据保存位置；
 - d) 点击【执行】按钮（图 5-19）；

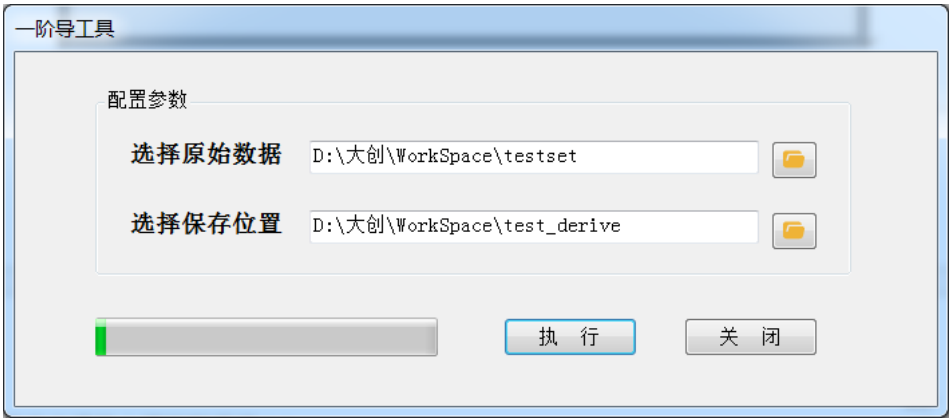


图 5-19 一阶导执行状态图

5.4 匹配与识别

5.4.1 直方图匹配

- 1) 操作目的：使用直方图匹配算法预测样本集土壤剖面类型；
- 2) 操作要求：系统已成功导入测试集、训练集以及类型数据；
- 3) 操作步骤：
 - a) 菜单栏【匹配】->【直方图匹配】（图 5-20）；



图 5-20 菜单栏直方图匹配项截图

- b) 【基于直方图的光谱匹配】界面选择【计算直方图】（图 5-21）->设置直方图组数->【计算】；

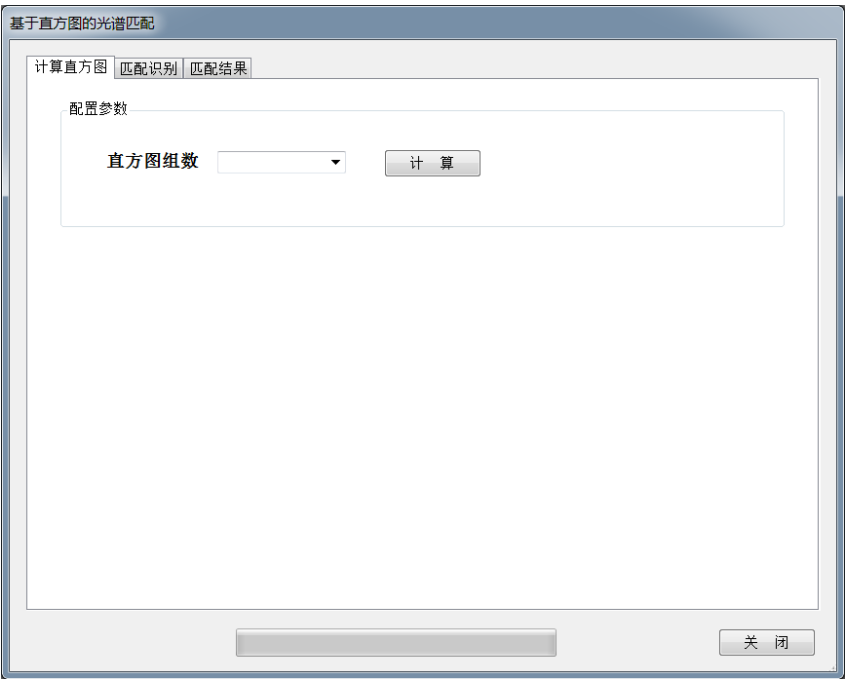


图 5-21 计算直方图界面图

- c) 【匹配识别】（图 5-22）->选择测试集合以及样本剖面；

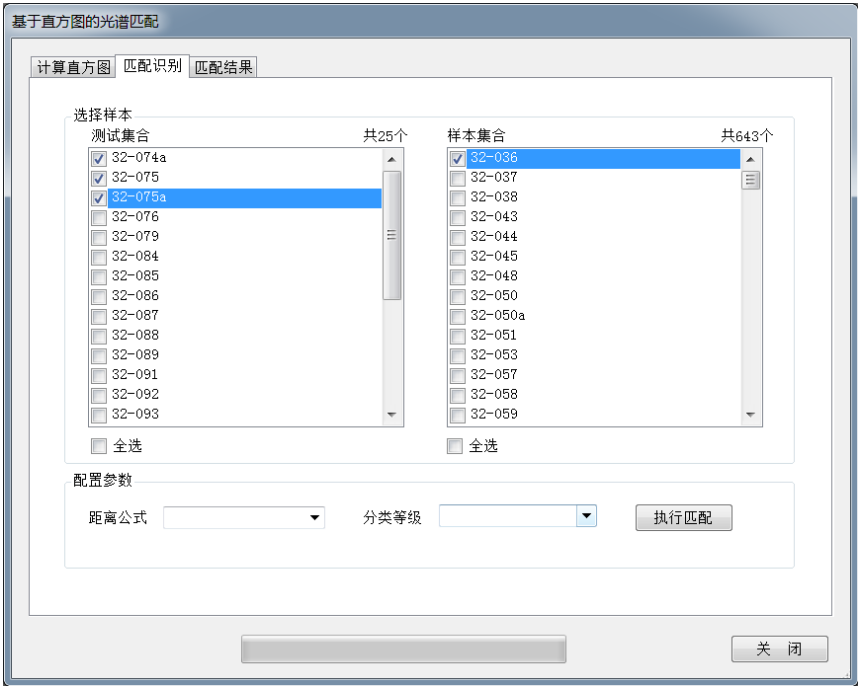


图 5-22 直方图匹配识别界面图

d) 选择距离公式->选择分类等级->【执行匹配】(图 5-23);

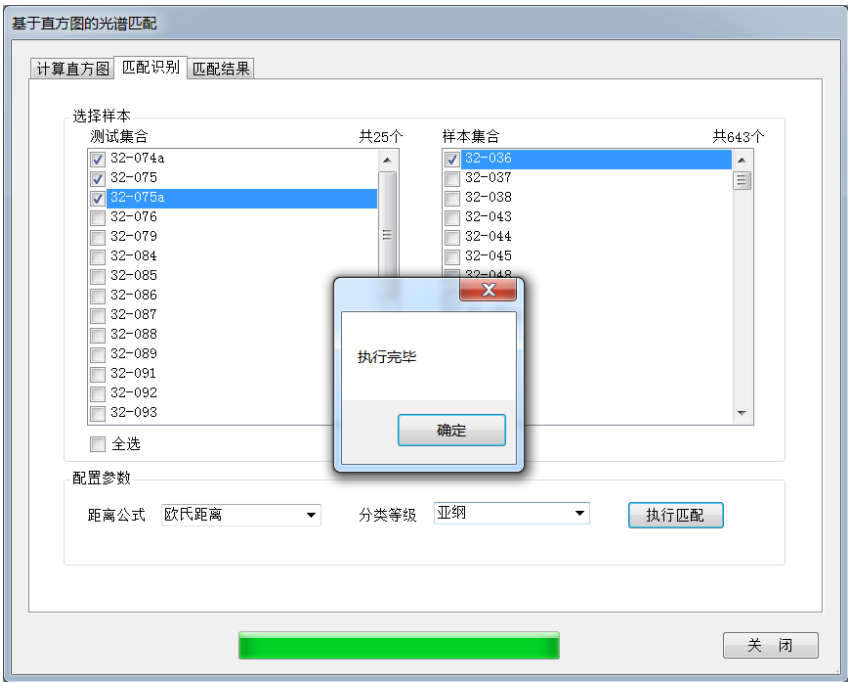


图 5-23 直方图匹配执行完成界面

e) 【匹配结果】->【导出结果】(图 5-24);

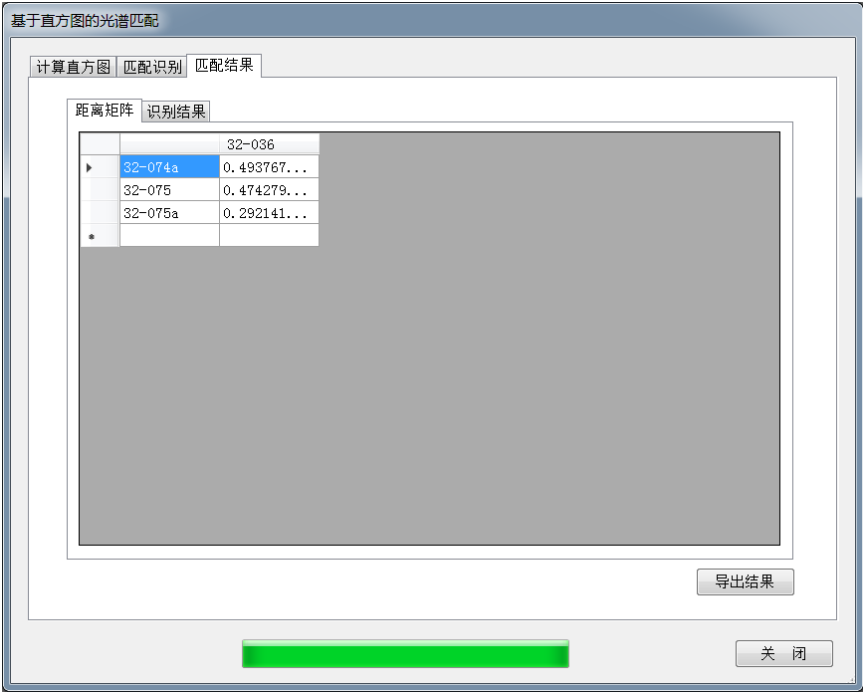


图 5-24 匹配结果界面图

- 4) 算法概述
- a) 算法基本思路：
- i. 计算测试集合光谱数据以及样本集合光谱数据的直方图；
 - ii. 根据测试集合直方图与样本集合直方图数据，选择合适的距离公式，计算测试集合与样本集合的土壤剖面的距离矩阵；
 - iii. 将与测试集合土壤剖面距离最小的样本集合剖面的土壤类型作为测试土壤剖面的类型；
- b) 算法流程图（图 5-25）：

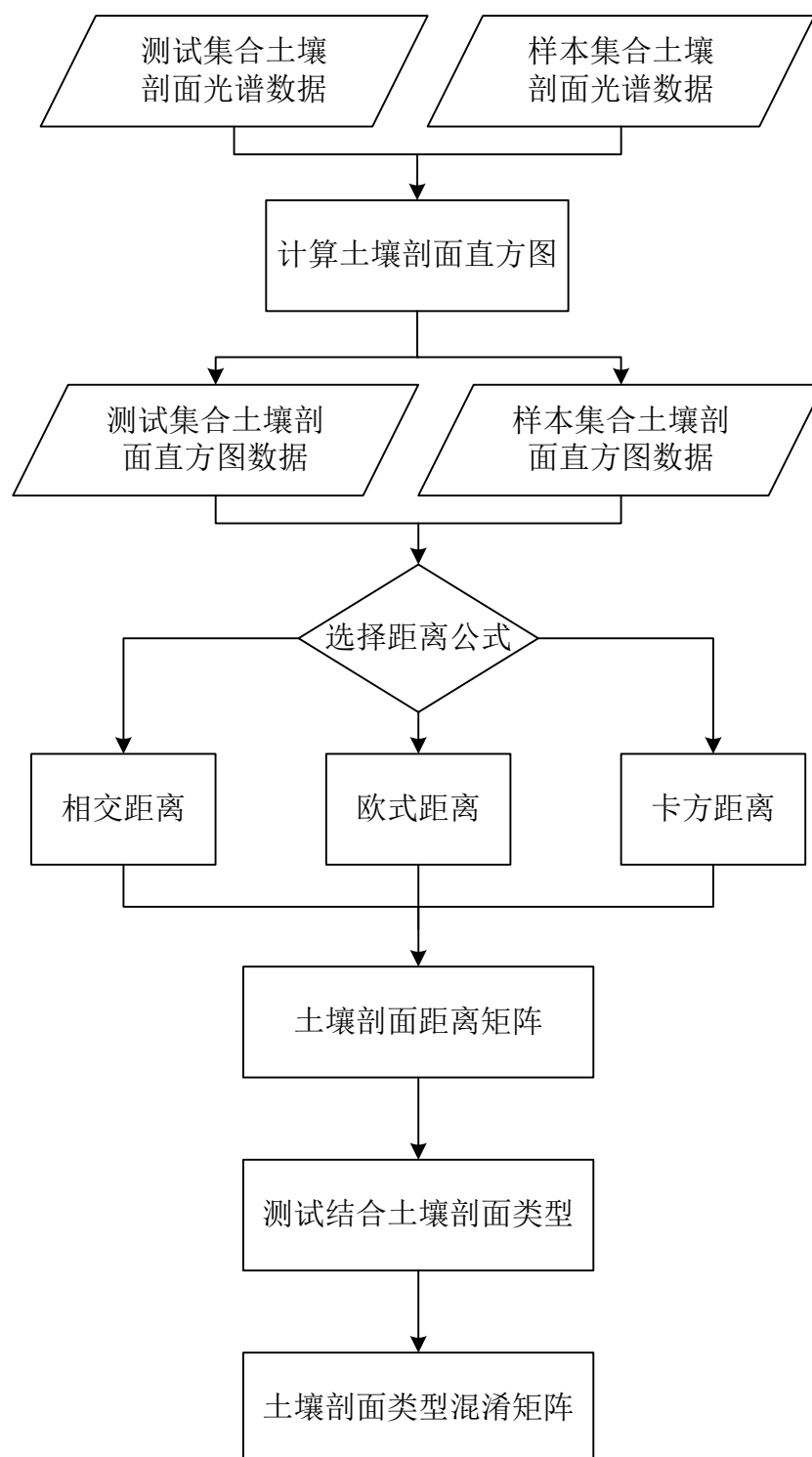


图 5-25 直方图匹配算法流程图

c) 算法文献:

朱磊. 一种基于直方图统计特征的直方图匹配算法的研究[J]. 计算技术与自动化, 2004, 23(2):48-51.

5.4.2 光谱角匹配

- 1) 操作目的：使用光谱角匹配算法获取测试集土壤剖面的土壤类型；
- 2) 操作要求：系统已成功导入测试集、训练集以及类型数据；
- 3) 操作步骤：
 - a) 菜单栏【匹配】->【光谱角匹配】（图 5-26）



图 5-26 菜单栏光谱角匹配项截图

- b) 【匹配识别】->选择测试集合以及样本剖面；
- c) 选择分类等级->【执行匹配】（图 5-27）；

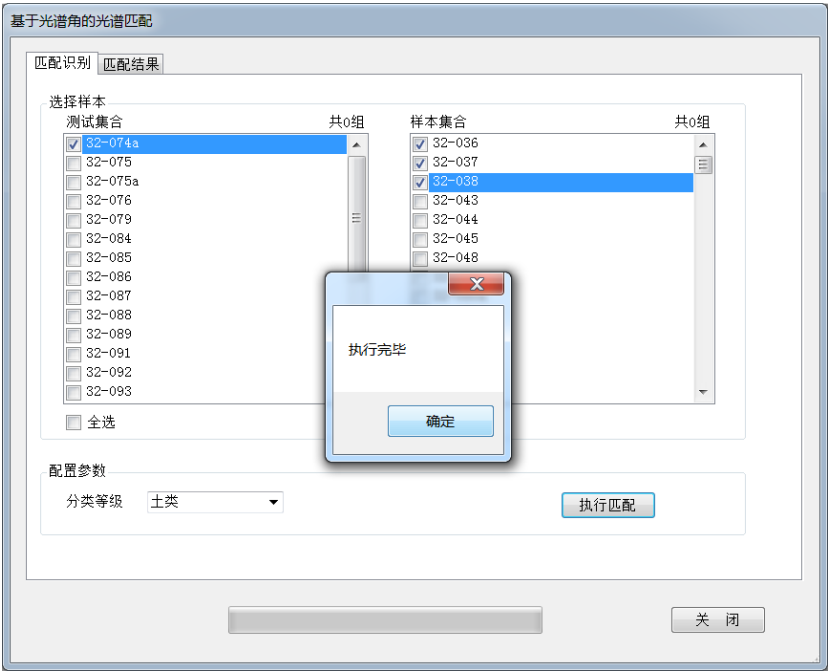


图 5-27 光谱角匹配识别界面图

- d) 【匹配结果】（图 5-28）->【导出结果】；

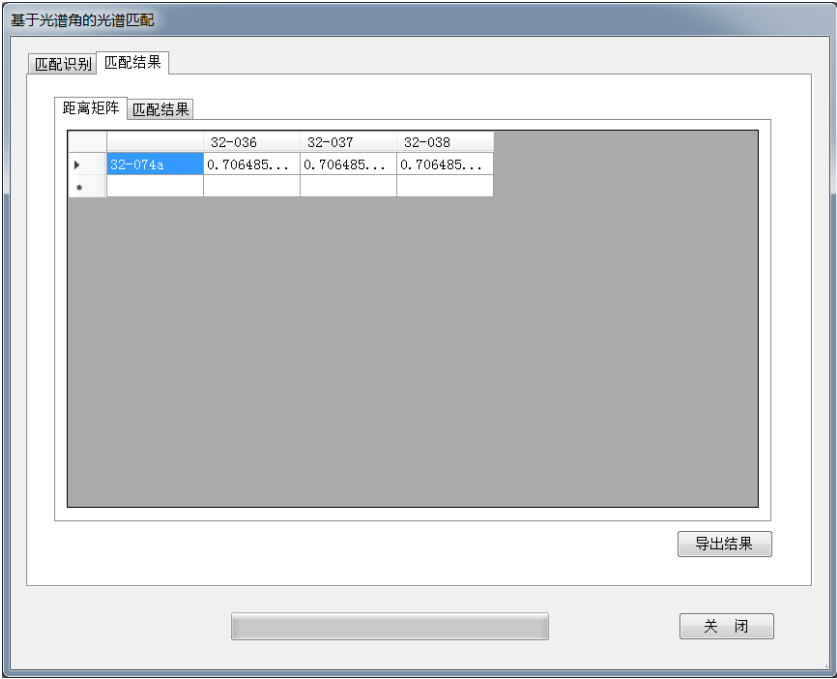


图 5-28 匹配结果界面图

- 1) 算法概述
- a) 算法基本思路:
- i. 计算测试集合以及样本集合土壤剖面各发生层间的光谱角;
 - ii. 依据光谱角值, 识别发生层的土壤类型;
 - iii. 汇总剖面各发生层的识别类型, 得到剖面的识别类型;
 - iv. 生成测试集合土壤剖面类型的混淆矩阵;
- b) 算法流程图 (图 5-29):

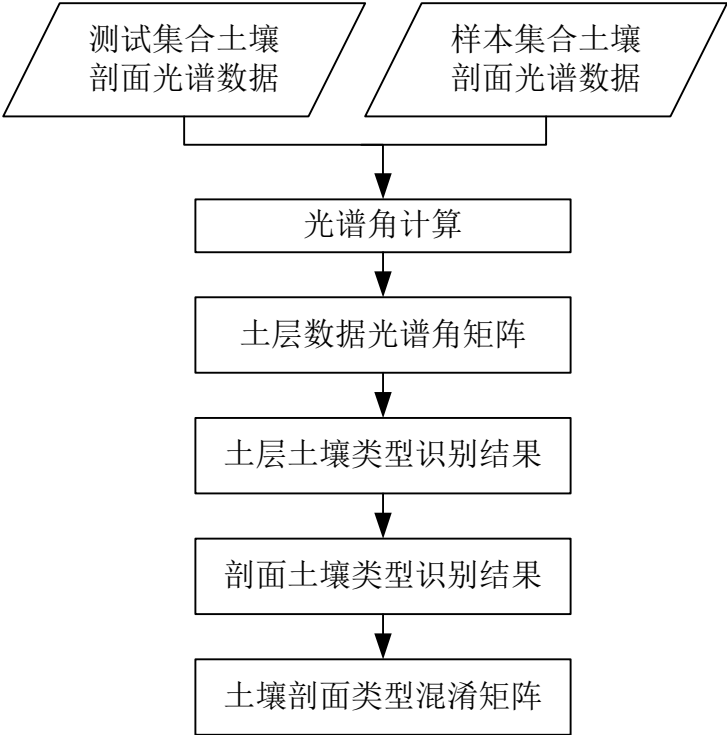


图 5-29 光谱角识别算法流程图

- c) 算法文献：
刘万军, 杨秀红, 曲海成. 基于光谱信息散度与光谱角匹配的高光谱解混算法[J]. 计算机应用, 2015, 35(3):844-848.

5.4.3 峰谷特征值匹配

- 1) 操作目的：提取测试集合基于样本集合光谱数据的峰谷特征值，并采用 KNN 分类算法获取测试集土壤剖面类型；
- 2) 操作要求：系统已成功导入测试集、训练集以及类型数据；
- 3) 操作步骤：
 - a) 菜单栏【匹配】->【峰谷特征值匹配】

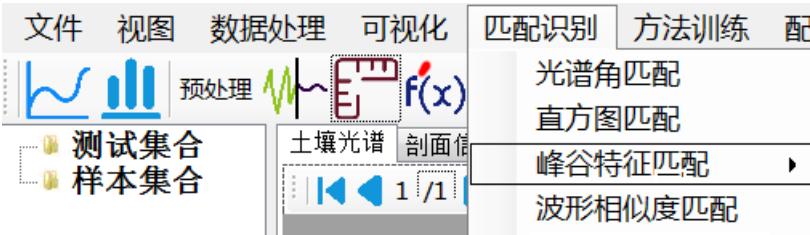



图 5-30 菜单栏峰谷特征匹配项截图

- b) 【基于峰谷特征的匹配算法】界面->【计算训练集特征】->【计算测试集特征】；
- c) 点击按钮，浏览文件，选择特征文件的保存位置；

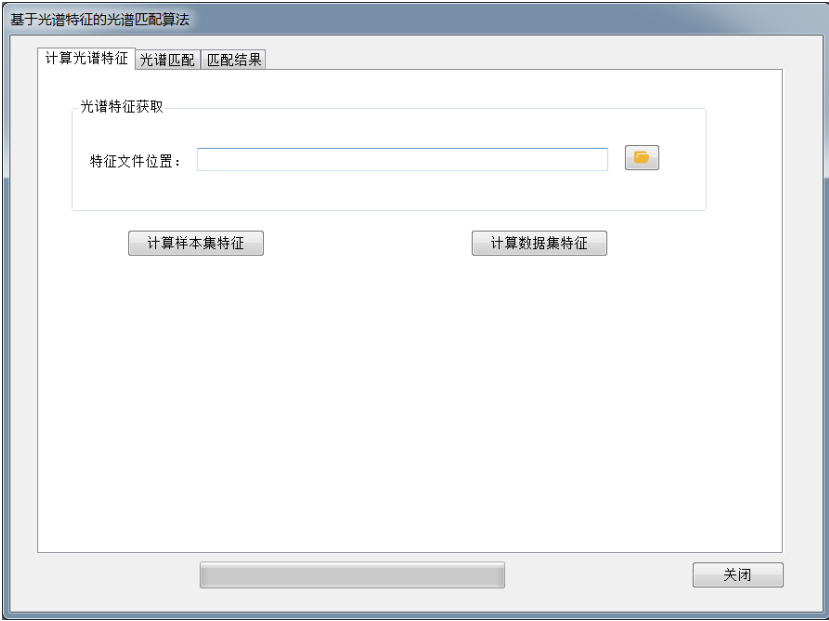


图 5-31 计算光谱特征界面图

- d) 【光谱匹配】->选择测试集合剖面以及样本集合剖面；
- e) 设置【匹配参数】->选择【匹配等级】->【执行匹配】；

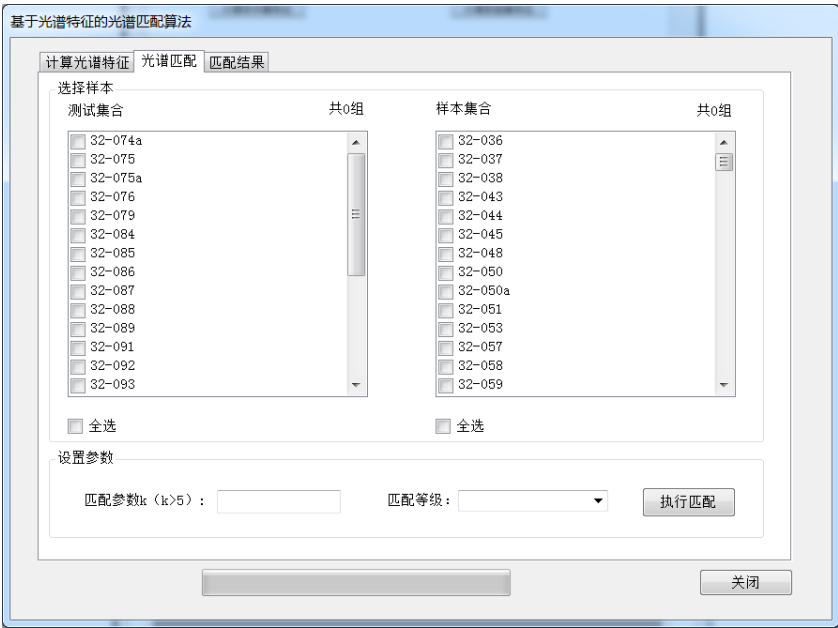


图 5-32 光谱特征匹配识别界面图

f) 【匹配结果】->【导出结果】:

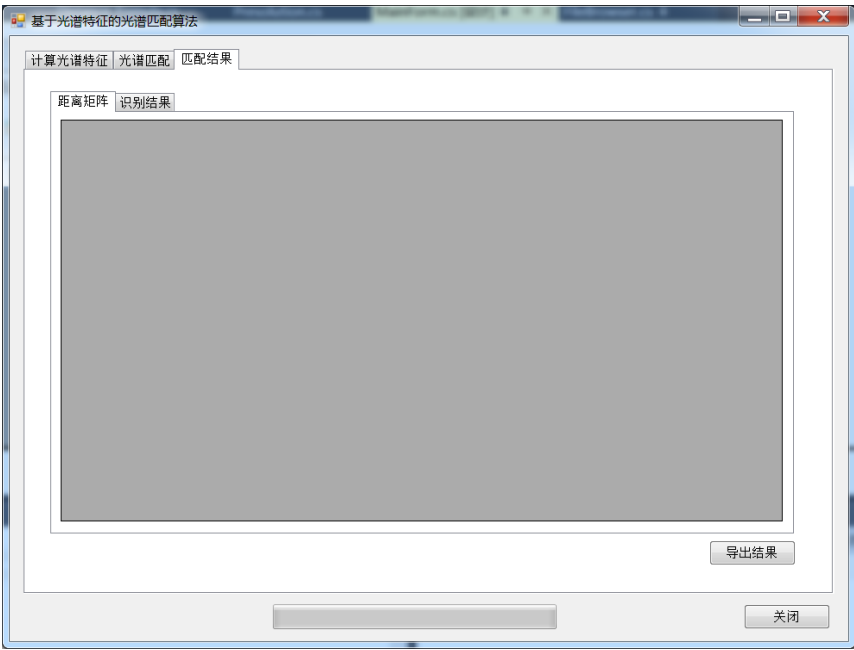


图 5-33 匹配结果界面图

4) 算法概述

a) 算法基本思路:

- i. 提取测试集合以及样本集合土壤剖面的最大波峰光谱值、波段值，最大波谷光谱值、波段值共四个特征值；
- ii. 依据四个特征，采用 KNN 分类法，得到测试集合土壤剖面的土壤类型；
- iii. 根据测试集合土壤剖面的识别类型生成土壤类型混淆矩阵；

b) 算法流程图（图 5-34）:

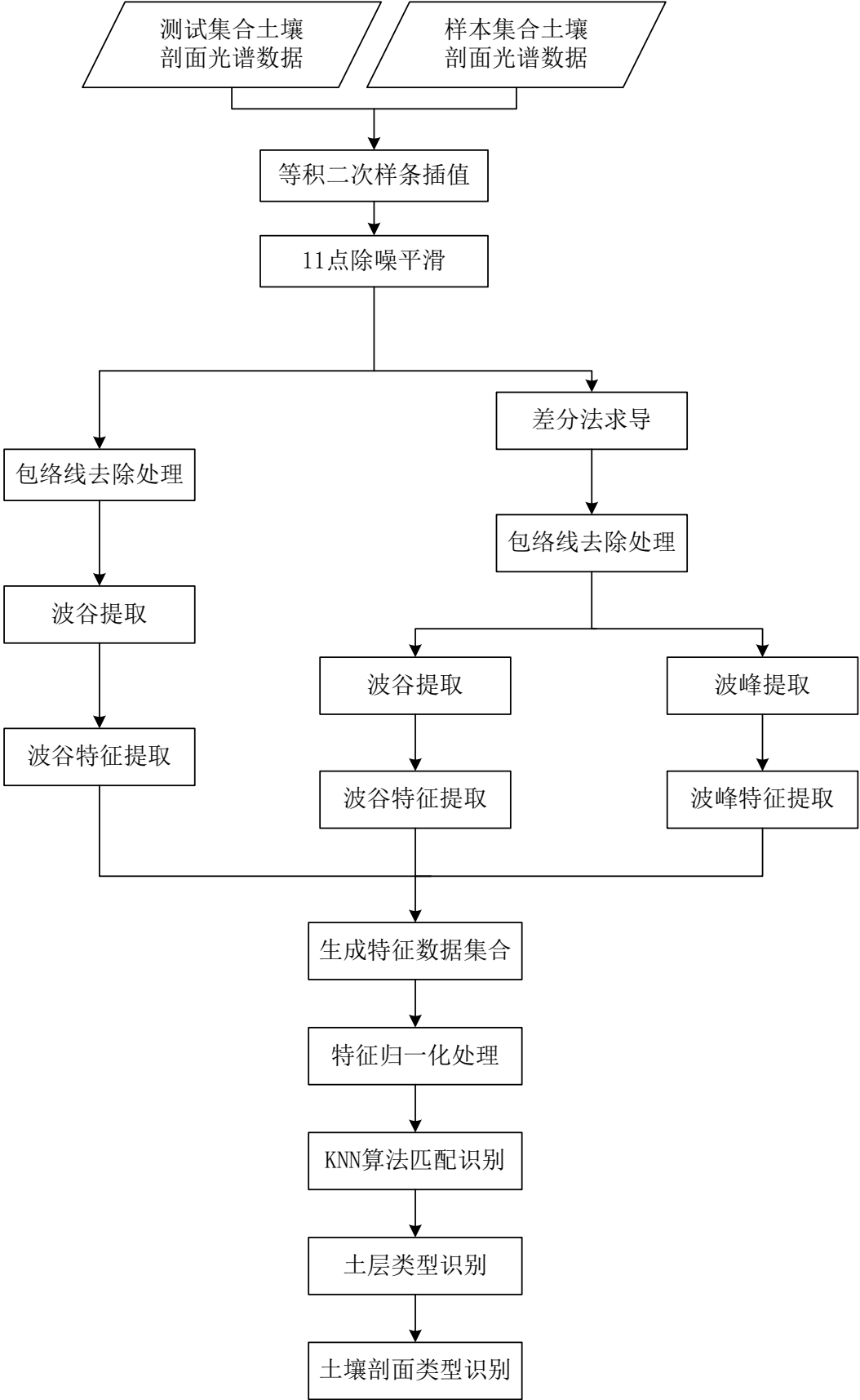


图 5-34 峰谷特征值识别算法流程图

c) 算法文献：

张文涵, 解宪丽, 李安波, 徐诗宇. 一种基于 KNN 分类进行土壤剖面类型识别方法: CN, CN107657264A [P]. 2017.

5.4.4 波形相似度匹配

- 1) 操作目的：使用波形相似度匹配算法获取测试集土壤类型；
- 2) 操作要求：系统已成功导入测试集、训练集以及类型数据；
- 3) 操作步骤：
 - a) 菜单栏【匹配】->【波形相似度匹配】（图 5-35）；

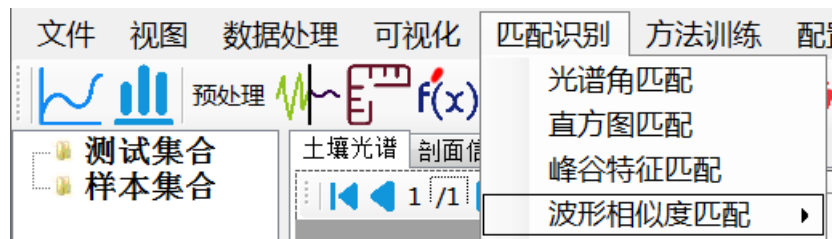


图 5-35 菜单栏波形相似度匹配项截图

- b) 【基于波形相似度的光谱匹配算法】界面->【发生层匹配结果】->选择测试集合剖面以及样本集合剖面；
 - c) 选择【分类等级】->【执行匹配】（图 5-36）；

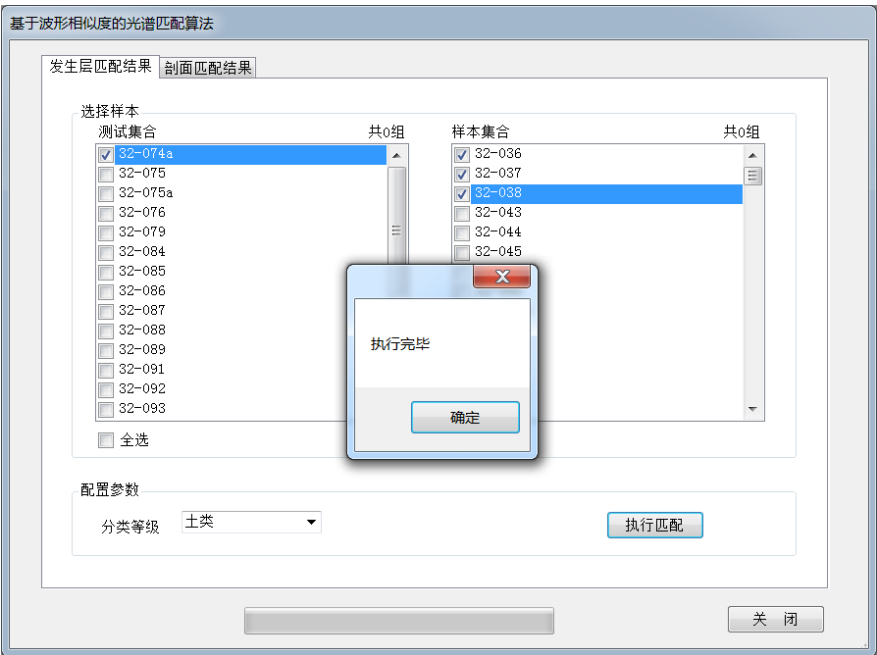


图 5-36 土层光谱匹配结果界面图

- d) 【剖面匹配结果】->【导出结果】（图 5-37）；

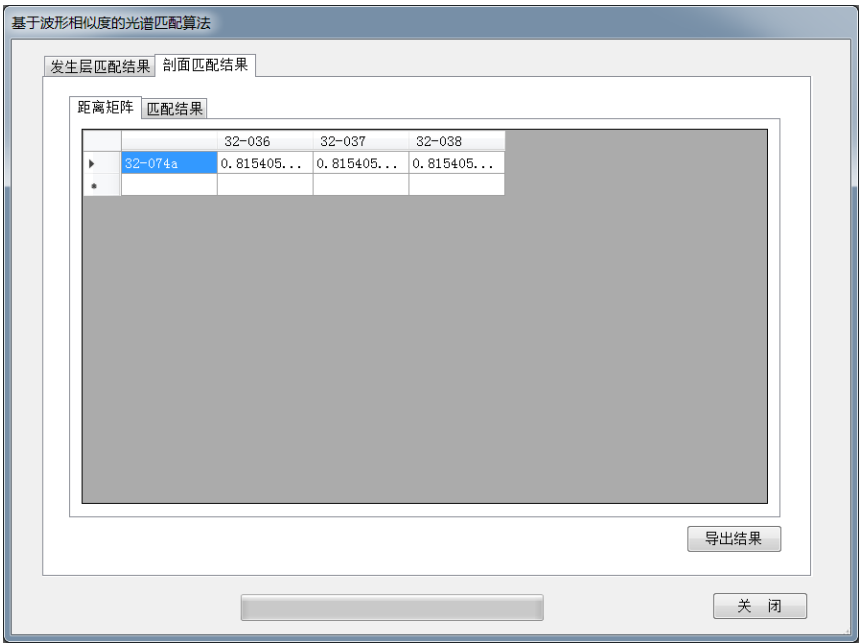


图 5-37 剖面匹配结果界面图

- 4) 算法概述
- a) 算法基本思路：
- i. 将测试集合以及样本集合土壤剖面光谱数据作归一化处理；
 - ii. 将测试集合与样本集合发生层归一化后的数据作比，并将同一发生层得到的比值加和，将结合结果作为波形相似度；
 - iii. 依据波形相似度，识别发生层的土壤类型；
 - iv. 汇总剖面各发生层的识别类型，得到剖面的识别类型；
 - v. 生成测试集合土壤剖面类型的混淆矩阵；
- b) 算法流程图：

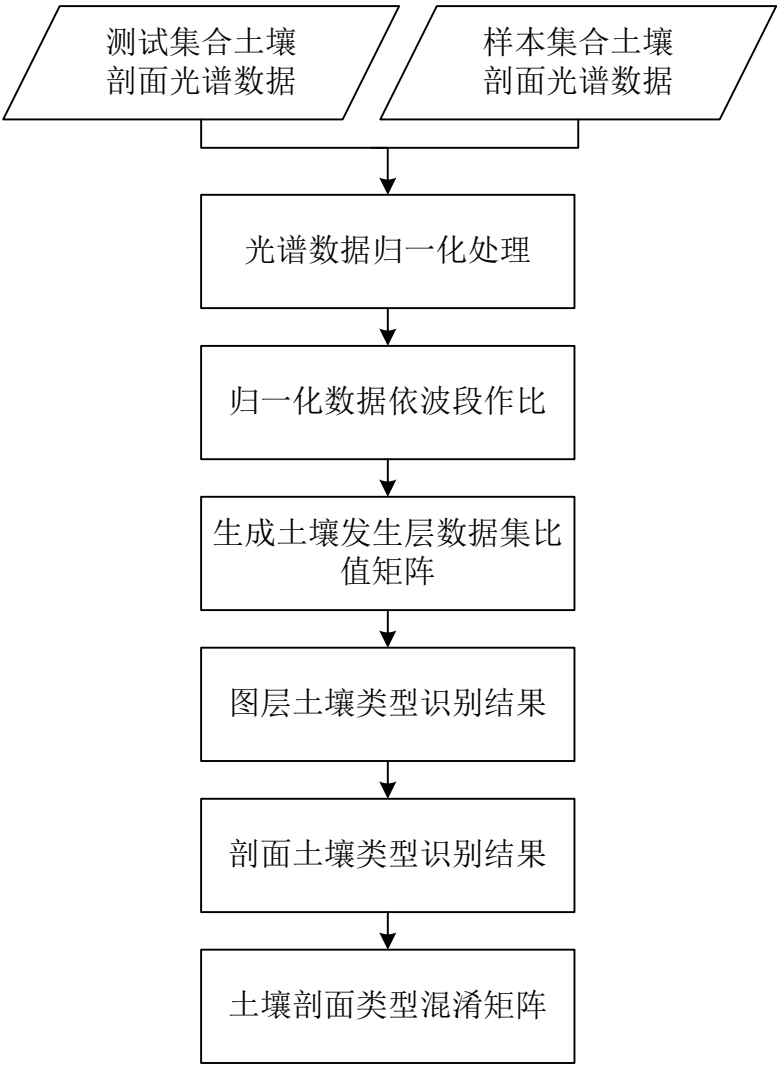


图 5-38 波形相似度识别算法流程图

c) 算法文献：
应义斌, 周万怀, 谢丽娟. 基于光谱曲线波形相似度的光谱匹配方法: CN, CN103488751A[P]. 2014.

5.5 方法训练

5.5.1 光谱角匹配方法

- 1) 操作目的：根据已有土壤样本训练光谱角匹配方法，测试其匹配精度；
- 2) 操作要求：系统已读取土壤样本；
- 3) 操作步骤：
 - a) 菜单栏【方法训练】->【光谱角匹配】（图 5-39）；



图 5-39 菜单栏光谱角匹配项截图

b) 【基于光谱角的匹配训练】界面（图 5-36）->【分配训练集】->设置训练集合比例->设置训练次数->【执行】;

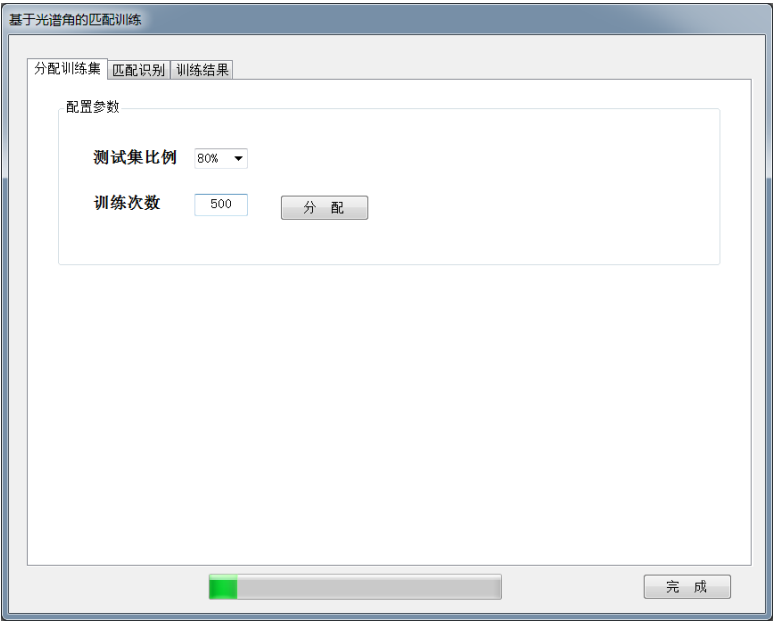


图 5-40 分配训练集界面图

c) 【匹配识别】（图 5-41）->选择训练结果文件 ->选择【分类等级】->【开始训练】;

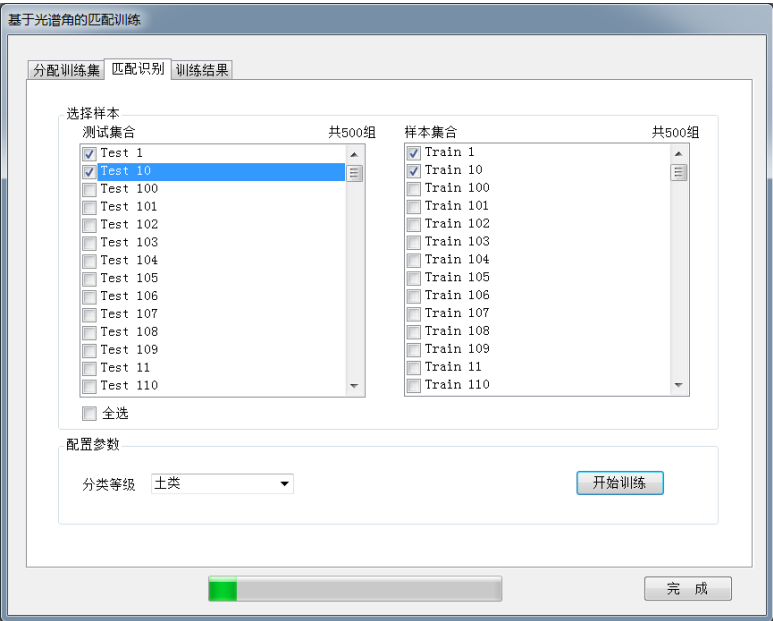


图 5-41 匹配识别界面图

d) 【训练结果】->【查看结果】;

5.5.2 直方图匹配

- 1) 操作目的：根据已有土壤样本训练直方图匹配方法，测试其匹配精度；
- 2) 操作要求：系统已读取土壤样本；
- 3) 操作步骤：
 - a) 菜单栏【方法训练】->【直方图匹配】（图 5-42）；



图 5-42 直方图匹配项截图

b) 【基于直方图的匹配训练】（图 5-43）界面->【分配训练集】->设置训练集合比例->设置训练次数->设置直方图组数->【执行】;

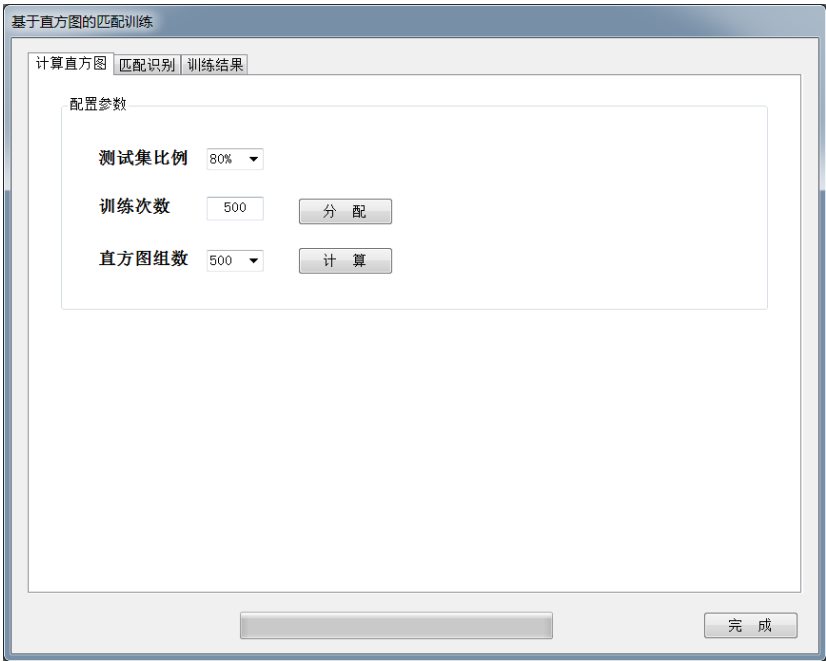


图 5-43 计算直方图界面

c) 【匹配识别】（图 5-44）->选择训练结果文件 ->选择【分类等级】->【开始训练】;

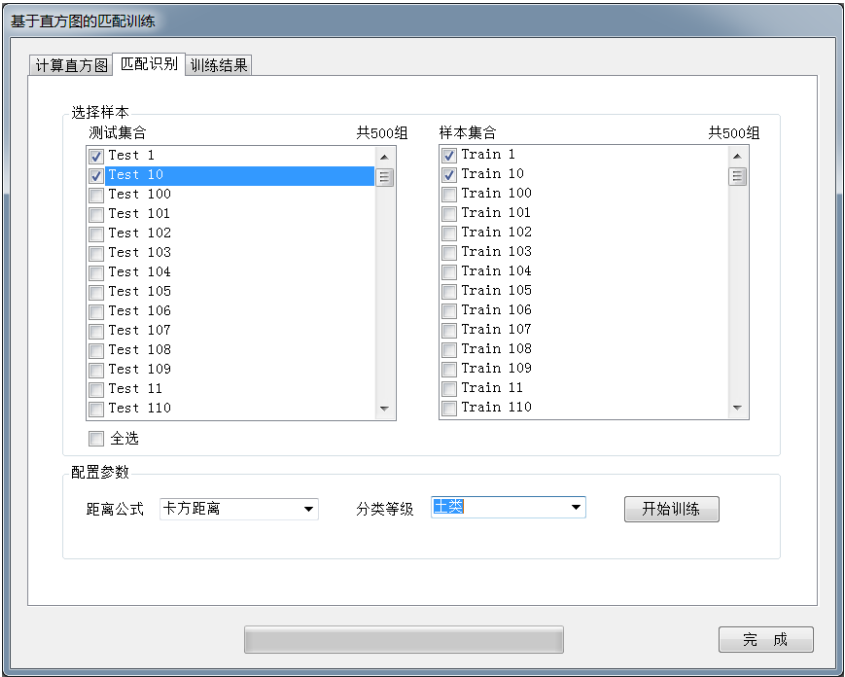


图 5-44 匹配识别界面

d) 【训练结果】->【查看结果】;

5.5.3 波形相似度匹配

- 1) 操作目的：根据已有土壤样本训练波形相似度匹配方法，测试其匹配精度；
- 2) 操作要求：系统已读取土壤样本；
- 3) 操作步骤：
 - a) 菜单栏【方法训练】->【波形相似度匹配】（图 5-45）；



图 5-45 峰谷特征值匹配项截图

b) 【基于波形相似度的匹配训练】（图 5-46）界面->【分配训练集】->设置训练集合比例->设置训练次数->【执行】;

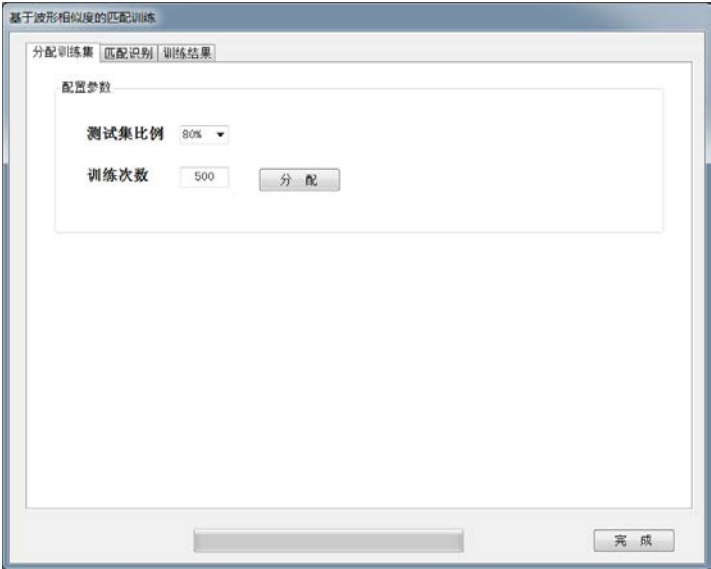


图 5-46 分配训练集界面

c) 【匹配识别】（图 5-47）->选择训练结果文件 ->选择【分类等级】->【开始训练】:

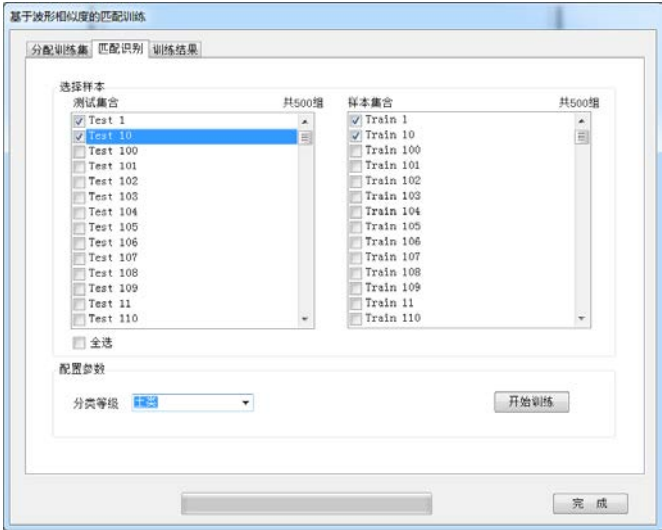


图 5-47 匹配识别界面

d) 【训练结果】->【查看结果】:

5.5.4 峰谷特征匹配

- 1) 操作目的：根据已有土壤样本训练峰谷特征匹配方法，测试其匹配精度；
- 2) 操作要求：系统已读取土壤样本；
- 3) 操作步骤：
 - a) 菜单栏【方法训练】->【峰谷特征匹配】（图 5-48）；



图 5-48 波形相似度匹配项截图

b) 【基于特征的匹配训练】(图 5-49) 界面->【分配训练集】->设置训练集合比例->设置训练次数->【执行】;

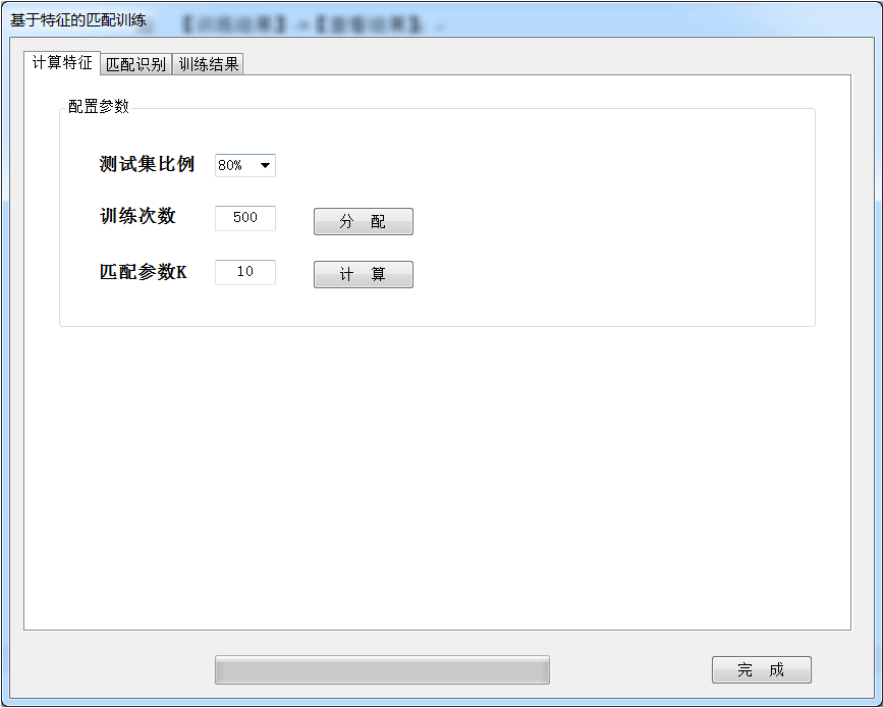


图 5-49 计算特征界面图

c) 【匹配识别】(图 5-50) ->选择训练结果文件 ->选择【分类等级】->【开始训练】;

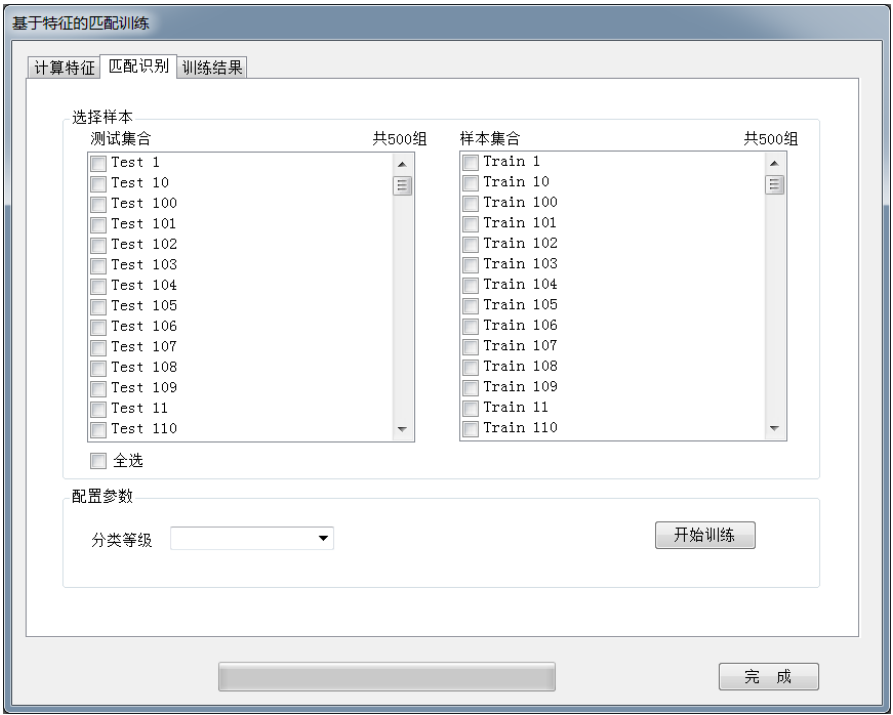


图 5-50 匹配识别界面

d) 【训练结果】(图 5-51) -> 【查看结果】:

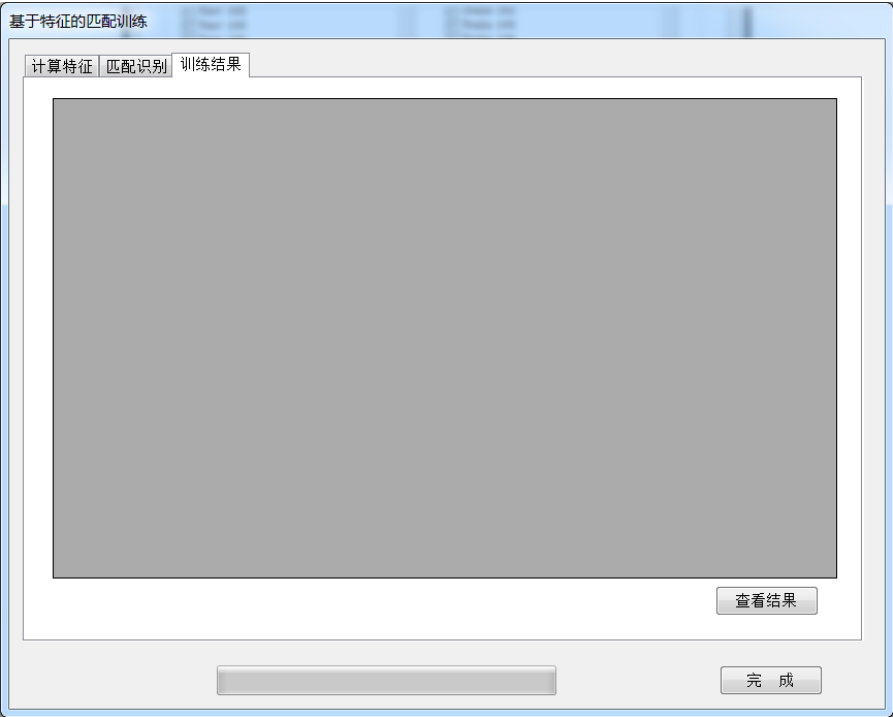


图 5-51 训练结果界面

5.6 主界面视图调整

5.6.1 控件可视性调整

- 1) 操作目的：更改系统界面上工具栏、状态栏、数据目录树的可视性；
- 2) 操作要求：电脑分辨率符合要求；
- 3) 操作步骤：
 - a) 菜单栏【视图】->【工具栏】（图 5-52）；

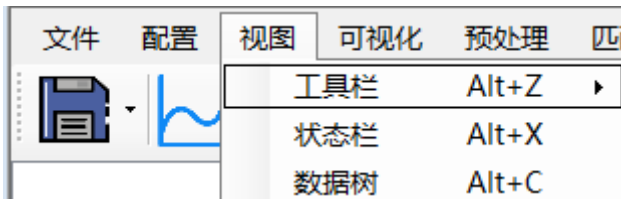


图 5-52 菜单栏工具栏项截图

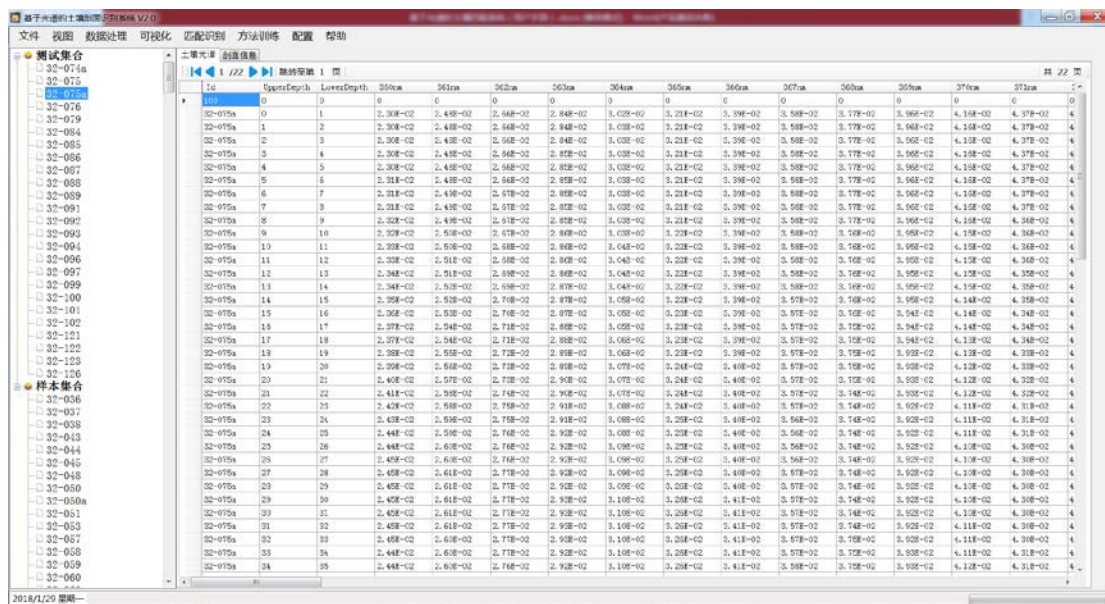


图 5-53 主界面工具栏隐藏效果图

- b) 菜单栏【视图】->【状态栏】（图 5-54）；

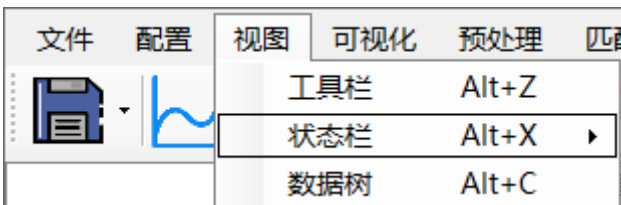


图 5-54 菜单栏状态栏项截图

基于光谱的土壤剖面识别系统 V1.0

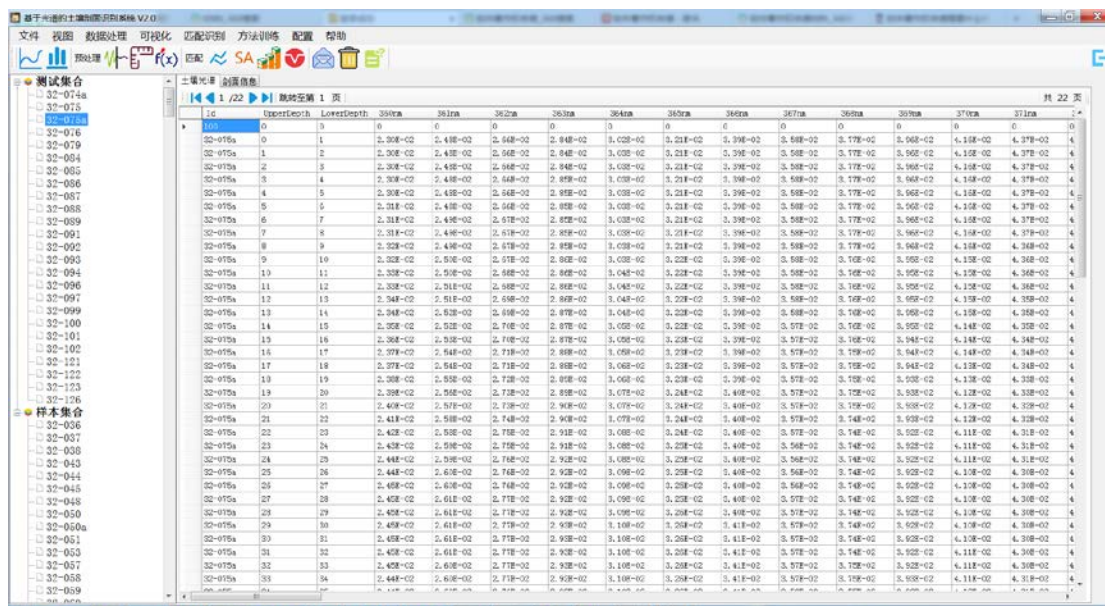


图 5-55 主界面状态栏隐藏效果图

c) 菜单栏【视图】->【数据树】(图 5-56);

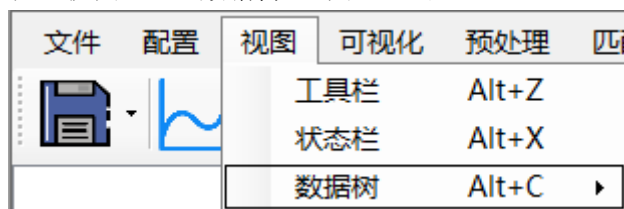


图 5-56 菜单栏数据树项截图

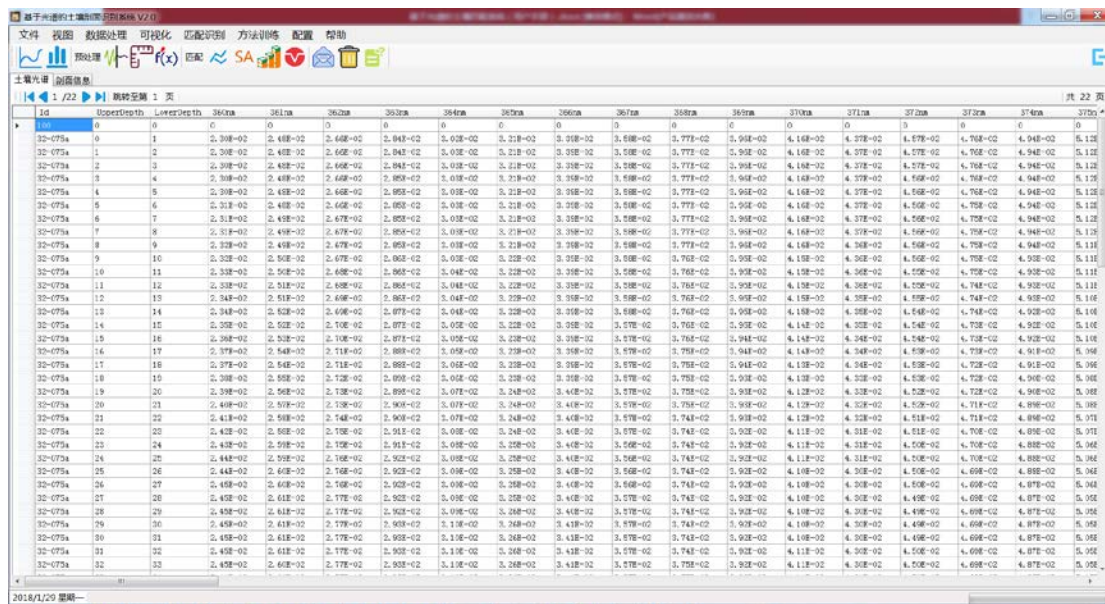


图 5-57 主界面数据树隐藏效果图

5.6.2 表格区域显示调整

- 1) 操作目的：调整表格区域显示的数据；
- 2) 操作要求：对应表格区域显示的数据成功导入系统；

3) 按钮功能（表 5-1）:



图 5-58 表格区域工具条截图

表 5-1 表格区域工具条按钮功能表

编号	功能
1	跳转至首页
2	跳转至下一页
3	跳转至上一页
4	跳转至末页
5	跳转至制页

5.7 主界面功能栏

- 1) 操作目的：快捷地实现功能；
- 2) 操作要求：无；
- 3) 功能说明（表 5-2）:



图 5-59 主界面工具条按钮分布图

表 5-2 主界面工具条按钮功能表

编号	功能
1	显示曲线图
2	显示直方图
3	除噪平滑
4	归一化
5	一阶导
6	波形相似度匹配
7	光谱角匹配
8	直方图匹配
9	峰谷特征值匹配
10	发送邮件
11	清空系统数据
12	打开帮助文档
13	退出系统

5.8 托盘

5.8.1 最小化到托盘

- 1) 操作目的：将界面最小化至托盘图标；
- 2) 操作要求：无；
- 3) 操作步骤：
 - a) 右击【托盘图标】->弹出【托盘菜单】(图 5-60)；

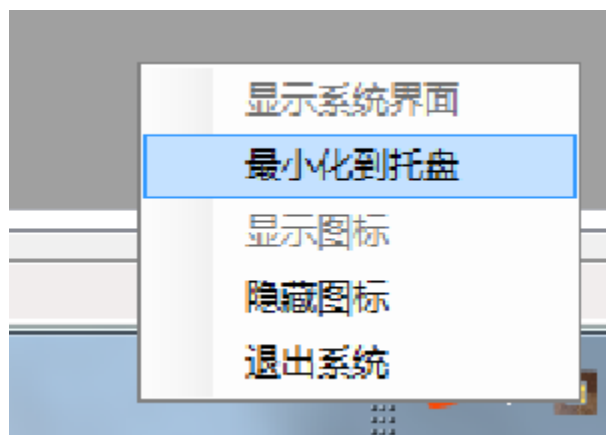


图 5-60 托盘左键菜单最小化按钮截图

- b) 点击【最小化到托盘】；

5.8.2 显示系统界面

- 1) 操作目的：显示最小化至托盘的系统界面；
- 2) 操作要求：界面已最小化至托盘图标；
- 3) 操作步骤：右击【托盘图标】->弹出【托盘菜单】->【显示系统界面】；

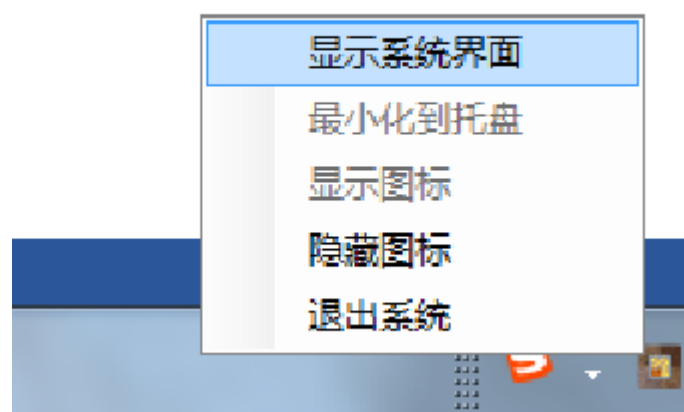


图 5-61 托盘左键菜单显示系统界面按钮截图

5.8.3 隐藏图标

- 1) 操作目的：隐藏系统在任务栏上的图标；
- 2) 操作要求：无；
- 3) 操作步骤：右击【托盘图标】->弹出【托盘菜单】->【隐藏图标】；

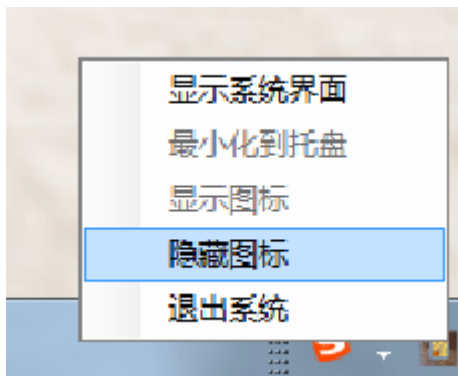


图 5-62 托盘右键菜单隐藏图标按钮截图

5.8.4 显示图标

- 1) 操作目的：显示已被隐藏的系统任务栏图标；
- 2) 操作要求：无；
- 3) 操作步骤：右击【托盘图标】->弹出【托盘菜单】->【显示图标】；

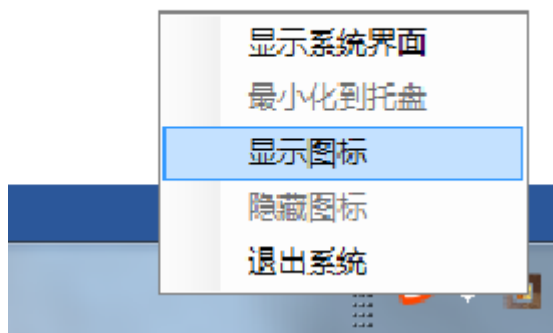


图 5-63 托盘右键菜单显示图标按钮截图

5.9 快捷键表

表 5-3 系统操作快捷键表

操作	快捷键
训练集读取	Ctrl+O
工具栏可视性	Alt+Z
状态栏可视性	Alt+X
数据树可视性	Alt+C
显示直方图	Ctrl+M
显示曲线图	Ctrl+N
归一化	Ctrl+G
一阶导	Ctrl+H
除噪平滑	Ctrl+J
训练集筛选	Ctrl+K
训练集统计	Ctrl+P
关于	Ctrl+U
帮助	Ctrl+I
关闭系统	ESC