

中图分类号：

UDC：

学校代码： 10055

密级： 公开

南开大学
硕士学位论文

便携式近红外光谱仪的软件设计及其在定量与判别分析中
的应用

Software Design for Portable NIR Spectrometer and Its
Application in Quantitative and Discriminant Analysis

论文作者 张伊挺

指导教师 蔡文生 教授

申请学位 工学硕士

培养单位 化学学院

学科专业

研究方向

答辩委员会主席 乔园园

评阅人 乔园园 董轶望

南开大学研究生院

二〇一六年五月

南开大学学位论文使用授权书

本人完全了解《南开大学关于研究生学位论文收藏和利用管理办法》关于南开大学(简称“学校”)研究生学位论文收藏和利用的管理规定,同意向南开大学提交本人的学位论文电子版及相应的纸质本,并委托印刷存档论文。

本人了解南开大学拥有在《中华人民共和国著作权法》规定范围内的学位论文使用权,同意在以下几方面向学校授权。即:

1. 学校将学位论文编入《南开大学博硕士学位论文全文数据库》,并作为资料在学校图书馆等场所提供阅览,在校园网上提供论文目录检索、文摘以及论文全文浏览、下载等信息服务;

2. 学校可以采用影印、缩印或其他复制手段保存学位论文;学校根据规定向教育部指定的收藏和存档单位提交学位论文;

3. 非公开学位论文在解密后的使用权同公开论文。

4. 同意学校将本人向有关电子出版单位授权的学位论文(含电子版和授权书)转交相关授权单位。

本人承诺:本人的学位论文是在南开大学学习期间创作完成的作品,并已通过论文答辩;提交的学位论文电子版与纸质本论文的内容一致,如因不同造成不良后果由本人自负。

本人签署本授权书一份,交图书馆留存。

学位论文作者暨授权人(亲笔)签字: 张伊挺

20 16 年 5 月 27 日

南开大学研究生学位论文作者信息

论 文 题 目	便携式近红外光谱仪的软件设计及其在定量与判别分析中的应用				
姓 名	张伊挺			答 辩 日 期	2016 年 5 月 27 日
论 文 类 别	博 士 ☐	学 历 硕 士 ☐	专 业 学 位 硕 士 ☒	同 等 学 力 硕 士 ☐	划 ☒ 选 择
学 院 (单 位)	化学院		学 科 / 专 业 (专 业 学 位) 名 称		材料工程
联 系 电 话			电 子 邮 箱	142857@mail.nankai.edu.cn	
通信地址(邮编): 天津市南开区南开大学中心实验室					
非 公 开 论 文 编 号			备 注		

注:本授权书适用我校授予的所有博士、硕士的学位论文。由作者填写一份并签字后交校图书馆,如已批准为非公开学位论文,须附批准通过的《南开大学研究生申请非公开学位论文审批表》和“非公开学位论文标注说明”页。

南开大学学位论文原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师指导下进行研究工作所取得的研究成果。除文中已经注明引用的内容外，本学位论文的研究成果不包含任何他人创作的、已公开发表或者没有公开发表的作品的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。本学位论文原创性声明的法律责任由本人承担。

学位论文作者签名：张伊挺 2016 年 5 月 27 日

非公开学位论文标注说明

(本页表中填写内容须打印)

根据南开大学有关规定，非公开学位论文须经指导教师同意、作者本人申请和相关部门批准方能标注。未经批准的均为公开学位论文，公开学位论文本说明为空白。

论文题目						
申请密级	☐ 限制(≤ 2 年)		☐ 秘密(≤ 10 年)		☐ 机密(≤ 20 年)	
保密期限	20	年	月	日至 20	年	月 日
审批表编号			批准日期	20	年	月 日

南开大学学位评定委员会办公室盖章(有效)

注：限制 2 年(可少于 2 年);秘密 10 年(可少于 10 年);机密 20 年(可少于 20 年)

摘 要

近红外光谱技术的不断成熟和发展，促进了近红外仪器的持续升级，性能、尺寸也在改进。近年来，近红外分析技术在现场分析和在线分析中的大量应用使近红外仪器开始向小型化、便携式发展，称为便携式近红外仪器。这些小型仪器在带来极大便利的同时也存在光谱分辨率低、波长点数少、信息量不足等缺点。此外，许多近红外仪器配套软件的数据处理方法陈旧，很难满足新型仪器在线检测的需要，这些因素在一定程度上阻碍了近红外光谱技术的应用。因此，为便携式仪器开发一款具有新型数据处理和分析功能的配套软件很有必要。本论文采用 MATLAB 软件开发近红外仪器配套软件，并结合相关化学计量学方法探索便携式近红外仪器用于准确定量定性分析的可行性。本论文主要在以下几个方面做了探究：

1、结合近红外光谱信息处理技术以及化学计量学方法，将软件要实现的功能集中于一个整体框架当中，明确各个模块之间的关系，画出草图。使用 MATLAB 软件逐一实现各个功能模块并整合到一起，构成完整的近红外光谱信息处理软件。该软件同时具有光谱采集、预处理、建立定性定量模型、预测以及生成报告等完整的功能，并将近年来应用较为广泛的新型光谱数据处理和建模方法嵌入到软件当中，为便携式仪器的现场检测和在线分析提供了便利。

2、使用便携式近红外光谱仪器结合富集技术定量检测稀溶液中的重金属离子，其中镍离子和铜离子为目标组分，砷离子、汞离子、铅离子、锌离子、铬离子和镉离子为干扰组分。使用偏最小二乘建立定量模型，得到镍离子和铜离子的预测值和参考值相关系数分别为 0.9894 和 0.9524。结果表明，结合化学计量学方法，便携式近红外光谱仪能够对重金属离子定量分析，为实际应用中使用的便携式近红外仪器进行定量分析提供了依据。

3、使用便携式近红外光谱仪器结合化学计量学方法对五种维库溴铵样品进行判别分析，使用不同的光谱预处理方法和距离判据组合建立判别模型，找出判别效果最好的模型，五种样品的类内判别准确率平均可达到 96.6%，类外判别准确率为 99.8%，结果表明便携式近红外仪器同样可以用于准确的判别分析。

关键词： 近红外光谱 MATLAB 便携式近红外光谱仪 化学计量学

ABSTRACT

With the development of NIR, NIR instruments have made considerable progress, including performance and physical size. In recent years, NIR instruments have been becoming smaller and smaller so that they can be applied in site and on-line analysing situations. These instruments are called "Portable NIR Instruments". Although they are quite convenient for users, there are also some shortcomings, such as low resolution, less information and so on. Besides, many algorithms built in softwares which are supported for NIR instruments are old fashioned and difficult to meet the needs of new NIR instruments, especially portable ones. All these factors, to some extent, limit the application of NIR. Thus, it will be quite necessary to develop one software with latest data processing algorithms for portable instruments. Here, MATLAB is used to develop NIR instruments supporting software, and explore the feasibility of portable NIR instruments for accurate quantitative and qualitative analysis. More details are as follows:

1. Build a frame that contains functions needed, understand the relationship between individual module, integrate kinds of functional modules to form a complete NIR information processing software by MATLAB. Compared with traditional softwares, this one has complete functions, such as spectral acquisition, preprocessing and predicting.

2. Using portable NIR instrument and enrichment technology to quantitative analysis heavy metal ions in dilute solution. Nickel and copper ions were treated as target components, while arsenic, mercury, lead, zinc, chromium, cadmium ions were interfering components. Partial Least Squares is used to build quantitative models and the coefficients of nickel and copper ions are 0.9894 and 0.9524, respectively. The results show that with the help of chemometrics methods portable NIR instruments can be used in quantitative analysis for heavy metal ions, which provide a basis for quantitative analysis.

3. Collected the near infrared spectroscopy of four kinds of vecuroniums.

ABSTRACT

Process these data(different pretreatment methods and distance compositions) and find the best model. The average accuracy rate of discrimination is 96.6% for within class and 99.8% for outside class. The results show that the portable NIR instruments can also be used for accurate discriminant analysis.

Key Words: NIR, MATLAB, Portable NIR Instrument, Chemometrics

目 录

摘 要.....	I
ABSTRACT.....	II
第一章 综述.....	1
第一节 近红外光谱仪器.....	1
1.1.1 近红外光谱仪器的种类.....	1
1.1.2 便携式光谱仪器.....	3
第二节 基于近红外漫反射光谱的分析方法.....	4
1.2.1 近红外分析方法原理.....	4
1.2.2 化学计量学方法.....	6
1.2.3 近红外光谱分析技术的应用.....	9
第三节 MATLAB 软件开发环境.....	10
1.4.1 MATLAB 语言及编程环境.....	11
1.4.2 图形用户界面设计.....	12
第四节 本论文选题.....	12
参考文献.....	14
第二章 基于 MATLAB 的近红外光谱软件开发.....	21
第一节 引言.....	21
第二节 软件设计.....	22
2.2.1 设计原则和模式.....	22
2.2.2 设计思路.....	23
2.2.3 窗口界面的设计.....	24
2.2.4 主要模块的设计.....	25
第三节 近红外光谱软件测试.....	28

2.3.1 光谱测量.....	29
2.3.2 建立定量模型.....	30
2.3.3 建立判别模型.....	32
2.3.4 生成建模报告.....	34
第四节 结论.....	35
参考文献.....	36
第三章 基于便携式近红外光谱仪的重金属离子定量分析研究.....	39
第一节 引言.....	39
第二节 样品准备与光谱采集.....	40
3.2.1 样品准备.....	40
3.2.2 近红外光谱的采集.....	40
第三节 结果与讨论.....	41
3.3.1 光谱预处理.....	41
3.3.2 校正模型建立.....	42
3.3.3 模型的预测效果.....	44
第四节 结论.....	45
参考文献.....	47
第四章 基于便携式近红外光谱仪的判别分析研究.....	48
第一节 引言.....	48
第二节 样品准备与光谱采集.....	49
4.2.1 实验样品及测试条件.....	49
4.2.2 近红外光谱的采集.....	49
第三节 结果与讨论.....	50
4.3.1 光谱预处理.....	50
4.3.2 判别模型建立及验证.....	51
4.3.3 模型的预测效果.....	53

目 录

第四节 结论.....	61
参考文献.....	62
致谢.....	63
个人简历、在校期间发表的学术论文及研究成果.....	64

第一章 综述

第一节 近红外光谱仪器

1.1.1 近红外光谱仪器的种类

近红外光谱仪器经过五十多年的发展，已经取得了很多成果，其应用领域也在不断的扩展。作为近红外光谱分析技术的基础和前提，近红外仪器在发达国家被普遍接受^[1]。早期的近红外光谱仪器是在紫外或可见光光谱仪的基础上配以近红外检测器而制成的，虽然降低了成本，但噪声大并且没有相应的数据处理系统，不能很好地满足近红外分析的要求。到了二十世纪七十年代，一些公司以 Norris 的研究成果为基础，生产出可以检测大豆中的水分和蛋白质的近红外光谱仪，同时微处理器的应用提高了近红外光谱仪的性能（可以使用微处理器来实现数据处理、存储和打印功能）。八十年代中后期，傅里叶变换近红外光谱仪的出现和光栅型近红外光谱仪器的提高标志着近红外仪器已经完全成熟。九十年代，光纤技术的应用使近红外仪器又一次实现了长足发展^[2,3]。现代计算机技术和计量学的发展使近红外光谱技术在工业和农业的在线分析领域得到了很好的应用，这也使得近红外仪器的发展趋于专用化、小型化、固态化和快速实时化^[4,5]。

1.1.1.1 滤光片型近红外仪器

滤光片是用来选取所需辐射波段的光学器件，一般应用于相对简单的近红外仪器中。将 6~12 片干涉滤光片（这些滤光片具有不同的透射波长）装在片盘上，不同的实验目的，选择对应的滤光片。滤光片型近红外仪器中有一类是便携式仪器，是为实现对某一单个成分进行检测分析而设计的，这种仪器不仅自身装有微型的类似芯片的处理装置，而且还带有串行接口，用以同计算机进行数据交互。这类仪器的优点是：采集样品光谱的速度快，可做成及时分析的便携式仪器；缺点是有移动部件，波长数目不高，分辨率太低，精确性难以保证。

1.1.1.2 发光二极管型近红外仪器

该种仪器的原理是用发光二极管(LED)作为光源,由不同的二极管产生不同的波长。LED 组件尺寸相对较小,因此一个小巧、便宜的过程控制器由几个LED 就可以组装完成。此类仪器常用于在线检测,优点是没有可移动部件,整体结构牢靠,而缺点在于光谱带宽变化大,波长点数少,精确度有限。

1.1.1.3 光栅单色器型近红外仪器

这类仪器是常见的近红外光谱仪,它和紫外-可见光谱仪具有通用的光学设计,只需要更换光栅、滤光片和检测器即可组成近红外光谱仪。该仪器的光源为能斯特灯(可在 360~3000nm 区域提供高能量的光源),检测器有硅检测器和硫化铅检测器(一般使用带制冷的 PbS 检测器以克服其温度漂移的缺点)。这种类型的仪器通常作为实验室近红外仪器使用,是一个通用型的近红外光谱仪。随着技术的发展,光谱仪不再限于实验室使用,借助于光导纤维探头实现现场检测功能。此类型仪器的优点是扫描速度快并且可扩展扫描范围,而缺点是有移动附件。

1.1.1.4 傅里叶变换型近红外仪器

傅里叶变换型近红外光谱仪以迈克尔逊干涉仪为核心,光源采用钨灯,分束器主要采用石英。与其它类型光谱仪相比,此类仪器的波长精度更高、分辨率好,扫描速度快(短时间内可进行多次扫描),光能利用率高、输出能量大,使得仪器的信噪比和灵敏度较高,可以对样品中的微量成分进行定量分析。这类仪器由于干涉仪中有移动部件,需要稳定的工作环境。

1.1.1.5 声光可调滤光片型近红外仪器

此类仪器是没有移动部件的单色器,波长精度高,光谱重现性好,结构坚固无移动部件,有良好的计算机接口与计算机进行通信。声光可调滤光片型近红外仪器主要的优点是速度快。

1.1.1.6 多通道检测型近红外仪器

近些年来,多通道检测器近红外光谱仪得到发展,其结构紧凑、光路固定,可高速采集样品光谱。在短波近红外区域使用的多通道检测器有二极管检测器 PAD 和电感耦合器件 CCD。多通道检测器对所有波长的单色光同时检测,仪器内无移动附件,光路固定,因此可以保证仪器的可靠性和波长精度,但不足之

处在于这种仪器对温度敏感。

1.1.2 便携式仪器

在近红外仪器的应用领域中很多时候需要体积小、方便移动的仪器用于现场分析，为满足这一需求，众多科研机构和企业基于不同的技术研发出了多种轻便、紧凑的近红外光谱仪。这些仪器被统称为便携式近红外光谱仪。

便携式近红外光谱仪器这个大家族里面包括了不同类型的仪器，比如比较早期的滤光片型、结构相对简单的光栅型以及高精密度的傅里叶变换型，这些不同类型的仪器是根据不同的光路原理结构来划分的。所有这些类型的仪器中，有一种基于微机电技术（microelectro-mechanical system，MEMS）的新型光谱仪近来发展势头尤为迅猛，本论文的软件开发工作的对象就是针对这一类型的便携式近红外光谱仪器——Axsun 公司的 IntegraSpec 系列的便携式近红外光谱仪，IntegraSpec™系列多功能近红外分析仪是目前美国近红外光谱领域无与伦比的尖端产品，其核心技术之一微光机电系统 MEMS，是对传统产品的一次革命，具有高可靠性，高稳定性，高一致性等特点，仪器设计寿命为 25 年，不需维修，无消耗品，是理想的在线过程控制和便携式近红外分析仪。IntegraSpec™系列的另一项核心技术是近红外波段独特的波长可调激光器，其亮度比传统仪器用的灯泡亮度要高好几个数量级，并且激光的波长和强度的短期和长期稳定性非常高。该仪器的主要特点有：

1. 坚固结实，分光系统为全密封芯片，电子制冷，恒温工作。因而仪器对使用环境非常不敏感，抗震动，耐冲击，不怕温度湿度变化，特别适用于在线监测和便携式使用；
2. 光源为波长可调激光器模块，波长和强度稳定性最佳，信号强度高；
3. 性能优越，各项指标优于或等同于大型而昂贵的实验室用傅立叶近红外光谱仪，而且仪器的一致性良好，使用灵活，可以用于气体、液体和固体的透射/反射测量，广泛应用于制药、石化、农业等各行业的过程控制和质量监测；
4. 采用成熟可靠的半导体集成芯片技术，仪器的设计寿命长，耐用；
5. 功耗低，电池供电时间长；
6. 独特的专利技术，内置式校正系统(WARM™ 波长/信号强度校正模块)，使这种仪器可以非常方便地同其它仪器进行模型转换；

7. 数据采集速度快，一条谱线的采集时间为毫秒级，适用于实时在线测量；
8. 操作使用简单，仪器的长期稳定性优异，使用成本低。

便携式近红外光谱仪的出现很大程度上提高了检测效率，节省了检测成本，是一类非常重要的快速检测仪器^[6]，同时也是近红外光谱仪器发展的一个重要方向。

第二节 基于近红外漫反射光谱的分析方法

近红外光谱 (near infrared spectroscopy, NIRS) 是波长在 780~2526nm 范围的电磁波，是人们最早发现的非可见光区域。它主要记录了分子中含氢基团 (C-H, O-H, N-H, S-H) 的倍频信息和不同形式组合的合频信息^[7]，使得光谱谱带比较复杂，从而使光谱的解析变得困难。化学计量学的引入解决了近红外光谱谱带重叠严重且不适合进行定性定量分析的问题。较之于中红外，近红外漫反射光谱几乎不需要样品处理，可直接对实际样品进行快速、无损和在线分析^[8-12]，广泛地应用于食品^[13]、烟草^[14]、石化^[15]、环境保护^[16]和医药^[17,18]等领域。

1.2.1 定量分析方法原理

按照测试方式的不同，近红外光谱分析技术可分为透射和反射光谱法两大类。对于均匀或者透明的样品一般采用透射法，理论依据在于透射光强度与样品中组分的浓度符合朗伯比尔定律 (Lambert-Beer Law)^[19]。本论文中采用的是近红外漫反射光谱技术。对于颗粒状和粉末状等固态样品的分析多采用漫反射方式。所谓漫反射是指，当一束光的入射光线照射到粗糙的表面时，由于表面的凹凸不平，入射光虽然相互平行，但由于各点的法线方向不一致，导致反射光线向不同的方向无规则地反射的现象^[3]。

使用漫反射光谱测量时，入射光是垂直于样品表面的，光谱检测器从侧面收集到的光包括镜面反射光和漫反射光。如果样品足够密实，不产生透射光，并且镜面反射光由于被样品表面直接反射回去，并没有与样品内部结构发生任何作用，故而不携带任何样品信息。反射光到达检测器时会产生一定的干扰信号，对分析产生不利影响，因此在测量时应尽量避免产生镜面反射光。入射光中一部分进入样品内部，并多次反射、散射、吸收、衍射、折射后进入检测器

的光是漫反射光。漫反射光携带有丰富的样品组成信息和结构信息，这是因为漫反射光在传播过程中与样品内部分子发生了相互作用。漫反射吸光度公式为：

$$A = \lg \frac{1}{R} \quad (1.1)$$

$$R = \frac{I}{I_0} \quad (1.2)$$

其中：A 为吸光度值；R 为反射率；I 是测得样品的漫反射光强度； I_0 是背景的漫反射光强度。

然而漫反射光的光程不规则，与样品散射系数和吸收系数有关，导致漫反射吸收光谱与样品中吸光成分的浓度之间的关系并不服从郎伯-比尔定律，意即漫反射吸光度与样品浓度之间并不能构成严格的线性关系。为解决这个问题，下面的公式被提出：

$$R = 1 + \frac{K}{S} - \left[\left(\frac{K}{S} \right)^2 + 2 \times \left(\frac{K}{S} \right) \right]^{0.5} \quad (1.3)$$

当样品的厚度足够大时，样品的漫反射率决定于样品的吸光度系数 K 和散射系数 S 的比率 K/S。漫反射吸光度系数主要取决于漫反射体的化学性质，是光通过物体表面时，由于对光的吸收导致光衰减的能力。而散射系数主要取决于漫反射体的物理性质，是光通过物体时，物体对光的散射使光衰弱的能力，其值与物体的物理状态以及样品测定条件有关。由公式 (1.1) (1.3) 可得

$$A = \lg \left(\frac{1}{R} \right) = -\lg R = -\lg \left\{ 1 + \frac{K}{S} - \left[\left(\frac{K}{S} \right)^2 + 2 \times \left(\frac{K}{S} \right) \right]^{0.5} \right\} \quad (1.4)$$

漫反射吸光度 A 与比值 K/S 的关系是一条通过零点的曲线，曲线各点的切线随斜率 K/S 的增大而减小。当 K/S 变化范围较小时，可以导出吸光度 A 近似与 K/S 成线性关系的方程，即

$$A = a + b \cdot \left(\frac{K}{S} \right) \quad (1.5)$$

其中，a、b 为拟合方程的截距和斜率。使用漫反射光谱分析样品化学成分参数如某个成分的浓度值时，样品的物理参数保持不变，即 S 固定，此时，漫反射吸光度 A 与样品浓度 c 呈线性关系：

$$A = a + b \cdot c \quad (1.6)$$

式 (1.6) 说明，如果样品无透射光谱、样品浓度范围变化不大并且散射系数不变时，近红外漫反射光谱可用于定量分析。当样品的化学组分较为复杂时，由

于各组分对光吸收的独立性，吸光系数具有累加性。但是截距 a 、斜率 b 随浓度范围变化故而吸光度并不具有加和性。必须运用化学计量学方法通过建立光谱与待测组分间的多元线性关系来达到定量分析的目的。

1.2.2 化学计量学方法

近红外光谱在包含丰富的样品结构信息的同时，也具有一系列的缺点：光谱之间相似度高^[20-24]、重叠峰严重^[25,26]，谱带相对较宽^[27-30]，容易受背景、噪声的干扰^[31-33]并且检测灵敏度较低^[34-38]。因此，需要借助于化学计量学方法从采集的光谱数据中提取有用信息，建立相关模型进而实现分析样品的目的^[39-41]。

化学计量学方法在近红外光谱分析中的应用主要包括光谱的预处理、谱区的选择、模型建立过程中奇异样本的识别以及建模过程中多元校正方法的应用。

1.2.2.1 光谱预处理方法

近红外光谱主要由含氢基团的倍频和组合频吸收峰组成，包含了有用的样品结构信息，同时也包含了背景噪声等无关信息，例如杂散光、背景和随机噪声。因此，在建立定性定量模型之前对原始光谱数据进行预处理十分必要。计量学中经常用到的光谱预处理方法有数据增强（Data Enhancement）算法、多元散射校正（Multiplicative Scatter Correction, MSC）算法、平滑（Smoothing）算法^[42,43]、导数（Derivative）算法、标准正态变量变换（Standard Normal Variate Transformation, SNV）和小波变换（Wavelet Transform, WT）算法等。下面是一些常用的光谱预处理方法。

1. 数据增强算法

为增强所建模型的稳健性进而提高模型的预测能力，通常使用数据增强算法来消除近红外光谱中所固有的冗余信息，提高样品光谱间的差异度。常用的数据增强算法有中心化（Mean Centering）、归一化（Normalization）和标准化（Autoscaling）。

中心化是用样品光谱减去平均光谱以消除光谱的“共性”，突出样品间的差异。中心化后的光谱矩阵的列平均值为零，将每条光谱的变动与待测组分的含量变动进行关联。运用中心化处理光谱数据的同时，习惯上也会将相应待测组分的浓度等数据进行相同处理。

标准化又称均值方差化，它将经过中心化处理后的光谱除以校正集光谱的

标准偏差，经标准化处理后的光谱阵列均值为零，方差为一。因为该方法给光谱中所有波长变量相同的权重，所以它特别适合于对低浓度成分建立定量模型。

归一化也是先对光谱数据中心化，然后将中心化后的光谱除以光谱的平方和。这种方法可以消除因微小光程差异带来的光谱变动。

2. 散射校正方法

样品颗粒表面不均匀、颗粒大小不同都会影响近红外光谱的吸收，散射校正方法提供了解决这一问题的途径。常用的散射校正方法有标准正态变量变换和多元散射校正两种方法。

标准正态变量变换主要用来消除固体颗粒大小、表面散射以及光程变化对近红外漫反射光谱的影响。标准正态变量（Standard Normal Variable, SNV）与标准化算法的计算公式相同，不同之处在于标准化算法对一组光谱处理（基于光谱阵的列），而 SNV 算法是对一条光谱进行处理（基于光谱阵的行）。其计算公式为：

$$x_{SNV} = \frac{x - \bar{x}}{\sqrt{\frac{\sum_{k=1}^m (x_k - \bar{x})^2}{(m-1)}}} \quad (1.6)$$

其中， $\bar{x} = \frac{\sum_{k=1}^m x_k}{m}$ ， m 为波长点数， $k=1, 2, \dots, m$

多元散射校正于 1985 年被提出，目的与 SNV 基本相同，主要是消除颗粒分布不均匀及颗粒大小产生的散射影响，在固体漫反射和浆状物透（反）射光谱中应用较为广泛。多元散射校正（Multiplicative Scatter Correction, MSC）算法的属性与标准化相同，是基于一组样品的光谱阵进行运算的。对一光谱 x ($1 \times m$)，MSC 的具体算法如下：

计算校正集样品的平均光谱 $\bar{x}$ （即“理想光谱”）；

将 x 与 $\bar{x}$ 进行线性回归： $x = b_0 + \bar{x}b$ ，用最小二乘法求取 b_0 和 b ；

$$\bar{x}_{MSC} = (x - b_0) / b$$

对于校正集外的光谱进行 MSC 处理时则需要用到校正集样品的平均光谱 $\bar{x}$ ，即首先求取该光谱的 b_0 和 b ，再进行 MSC 变换。对组分性质变化较大的样

品光谱进行处理时 MSC 算法的效果可能不理想，原因在于该算法假定散射与波长及样品的浓度变化无关。

3. 主成分分析 (Principal Component Analysis, PCA)

主成分分析在化学计量学中的地位举足轻重，是一种古老的多元统计分析技术，由 Hotelling 于 1933 年提出。PCA 的中心目的是将数据降维，转换原始变量，使得少数几个新变量等于原变量的线性组合，而这些变量需要尽可能多地表达原变量的数据特征并且不丢失信息。PCA 把数据变换到一个新的系统中，使得任何数据投影的最大方差为第一主成分，第二大方差为第二主成分，以此类推。经 PCA 转换得到的新变量相互正交，从而消除了众多共存信息中重叠的部分，即多重共线性。

4. 偏最小二乘法 (Partial Least Squares, PLS)

目前应用最为广泛的建模方法之一的 PLS 算法是在主成分回归的基础上发展起来的。主成分回归的不足之处在于仅对光谱矩阵进行了主成分分析而实际上，样品的浓度矩阵同样包含有无用信息。PLS 就是对光谱矩阵和浓度矩阵同时进行分解，在光谱数据分解的过程中引入浓度信息，将光谱主成分直接与待测组分含量相关联，进一步提高了模型的可靠性。实现过程为：

对光谱矩阵 X 和成分矩阵 Y 进行主成分分解：

$$X = TV^T + E_X \quad (1.7)$$

$$Y = RQ^T + E_Y \quad (1.8)$$

式中 T 、 R 、 V 、 Q 分别为 X 和 Y 的得分矩阵和载荷； E_X 和 E_Y 分别为 PLS 拟合残差矩阵。

对得分矩阵进行线性回归：

$$R = TB \quad (1.9)$$

$$B = (T^T T)^{-1} T^T Y \quad (1.10)$$

用未知样品的光谱阵乘以载荷得到未知样品光谱的得分，然后根据下面公式求得预测浓度值：

$$Y_{\text{未知}} = T_{\text{未知}} B Q^T \quad (1.11)$$

依据预测残差平方和 (Prediction Residual Error of Square, PRESS) 评价模型的质量——PRESS 的值与模型的预测能力负相关，PRESS 值越大模型预测能力越差，反之则模型的预测能力越好。PRESS 的计算公式如下：

$$PRESS(r) = \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \hat{y}_{ij})^2 \quad (1.12)$$

式中 m 是样本数, n 是组分数, y_{ij} 表示第 j 个组分在第 i 个样本子集中的成分值, $\hat{y}_{ij}$ 表示主成分数为 r 时所对应的预测值。

1.2.3 近红外光谱分析技术的应用

与传统分析技术相比, 近红外光谱分析技术有诸多优点——无损、快捷、方便、易于在线分析, 它能在短短几分钟的时间里完成对被测样品的采集测量同时测定多项性能指标 (可多达十余项指标), 光谱测量时不需要对分析样品进行过于复杂的前处理, 不消耗其它材料, 不需要破坏样品, 对于在线监控等实时分析领域是十分经济且快速的。这些优势使其具有一系列不可替代的作用, 满足了现代分析的要求。

与此同时, 现代计算机技术日新月异地发展和化学计量学方法的不断推进, 以及近红外光谱仪器性能的不不断提升, 现代近红外光谱技术的应用已经由传统的只简单用于农副产品的分析, 扩展到众多的应用领域, 诸如石油化工、生命科学、冶金、医学临床、食品质检、饮料、烟草、纺织、造纸、化妆品、质量监督、环境保护与制药等方面^[44-50]。

1、农业领域的应用

近红外光谱技术最早发挥作用的行业就是农业, 也是近红外技术应用最成熟的领域之一^[51-57]。在二十世纪六十年代, 就有多元线性分析方法结合近红外漫反射光谱技术测定小麦固体样品的水分含量的应用先例。二十一世纪以来, 近红外分析技术在农业领域的研究继续深入, 分析精度越来越高, 应用范围越来越广泛。孙君明等^[58]将气相色谱技术与近红外光谱结合, 借助 PLS 和交叉验证算法, 探索了利用近红外光谱分析技术测定豆类脂肪酸含量的可行性, 结果表明这种方法能够较准确地预测目标分析物; 2015 年在巴西举办的国际近红外会议上, Williams 等^[59]描述了如何使用近红外评估农业食品质量和安全性及在实验室水平上、行业和工厂应用水平上控制食品加工过程。Mignani 等^[60]运用近红外光谱技术对橄榄油中掺杂油的种类进行判别分析, 同时探究优质橄榄油的成分值。Oliveri 等^[61]使用近红外光谱鉴别蜂蜜的质量等级。

2、石化行业的应用

石油化工产品是石化行业的主产品，这些产品大都以石油为原料精馏制备而成，碳氢化合物是其主要成分^[62,63]。由于碳氢化学键在红外、近红外波段能够被充分地吸收，因此使用近红外光谱可以轻易地实现对石化产品的组分进行分析^[64-66]。目前近红外光谱技术在石油化工领域中的应用，集中在石化产品的组分及其物理化学性质的检测。Kelly 等人^[67]首次将近红外光谱运用到石油化工分析领域，并成功测出了汽油样品中的辛烷值。Balabin 等人^[68]利用近红外光谱结合拉曼光谱技术研究石油中的沥青吸附在铁表面的微观过程，并对其表面层型结构进行分析。Santos 等人^[69]使用近红外光谱技术分析油水乳状液中的水珠平均粒径和水分含量，同时实现了在线预测。

3、生命科学与制药分析方面的应用

近红外光谱可以实现对生物体非介入式分析和检测已经成为临床医学方面最有发展前景的分析手段之一，目前近红外技术在此行业的应用集中在血氧和血糖的检测方面。近红外光谱在药物分析中的可追溯到二十世纪六十年代后期，近红外漫反射光谱以其无损、快速、易于在线分析等特征，逐步开始在中药成分分析、体内药物和生物分子分析、制剂成分分析、制药过程控制以及实时在线分析等方面被广泛接受^[70-73]。Koller 等人^[74]利用近红外光谱监测固型药物的生产过程，对主要成分的含量进行控制。Moerman 等^[75]将近红外光谱技术应用于胸腺动脉瘤修复手术中，监测脊柱部位的血压变化。Bittner 等人^[76]使用近红外光谱鉴定药品阿莫西林别。Poellinger 等^[77]用近红外成像技术研究乳腺癌前期和晚期的切片样品，用于临床诊断乳腺癌患者，并区分病变不同时期。邵学广等^[78]将近红外漫反射光谱法结合化学计量学对不同种类药片中的有效成分含量进行了定量分析，取得了良好的预测效果。Leff 等人^[79]利用近红外光谱研究大脑活动时的血液动力学，为神经性疾病恢复提供依据。

在我国的现代近红外光谱技术的研究和应用虽然起步较晚，但发展迅猛，在农副产品及石化领域已经较为成熟。随着计算机和化学计量学方法的不断发展，以及仪器硬件的不断开发，近红外光谱技术越来越成熟、稳定、实用，还将会出现在更多的领域一显身手。

第三节 MATLAB 软件开发环境

现代近红外光谱技术是借助于硬件和软件两方面来实现从复杂、重叠、变动的光谱中提取有用信息的^[80-87]。国际国内有很多近红外公司或者研究机构已经开发出了很多种不同类型的近红外光谱信息处理软件，包括挪威 CAMO 公司的 The Unscrambler 软件、荷兰 Avantes 公司的 AvaSoft 软件、瑞士 BUCHI 实验室开发的 NIRCal 操作软件和中国石油化工科学研究院的 RIPP 化学计量学光谱分析软件等^[84]。MATLAB 科研软件软件作为一款功能强大的数据处理与分析工具，主要用于算法开发、数据可视化、数据分析以及数值计算。MATLAB 是 Matrix 和 Laboratory 两个词的组合，意为矩阵实验室，是由美国 Mathworks 公司发布的主要面对科学计算、可视化以及交互式程序设计的高技术计算环境。它将数值分析、矩阵计算、科学数据可视化^[88]以及非线性动态系统的建模和仿真等诸多强大功能集成在一个易于使用的视窗环境中，为科学研究、工程设计以及必须进行有效数值计算的众多科学领域提供了一种全面的解决方案，并在很大程度上摆脱了传统非交互式程序设计语言的编辑模式，代表了当今国际科学计算软件的先进水平，与 Mathematica、Maple 并称为三大数学软件，已经在自然科学、社会生产和科学研究等各个领域得到了广泛的应用^[89-95]。本论文的近红外光谱信息处理软件的开发就是基于 MATLAB 展开的。

1.4.1 MATLAB 语言及编程环境

- 1.桌面工具和开发环境。MATLAB 为用户提供了一个集成的开发环境，以方便利用 MATLAB 自带的函数和文件进行产品开发。“桌面工具和开发环境”是指 MATLAB 的一些工具集，大多具有图形用户界面，包括 MATLAB 桌面、命令窗口、工作空间、程序编辑器和调试器、代码分析器、帮助信息浏览器、文件以及其他工具等。
- 2.数学函数库。MATLAB 中有一个功能强大的数学函数库，涵盖大量的数值计算算法，其中包括化学计量学中经常用到的主成分分析 (PCA)，协方差 (COV) 以及小波变换等，使基于 MATLAB 的近红外光谱信息处理软件的开发过程变得简单而有效率。
- 3.编程语言。MATLAB 编程语言是以矩阵/数组为基本运算单元的交互式高级编程语言，又称为 M 语言。它包括程序流控制、函数、数据结构、输入/输出、面

向对象编程功能等。利用 MATLAB 编程语言，用户既可以快速编写简单的小程序，也可以进行大规模应用程序的开发。

4.缺点。所谓“尺有所短寸有所长”，MATLAB 在矩阵运算方面性能优越，但在循环以及字符串处理方面却表现一般，故而在进行软件开发过程中应当尽可能地避免循环，以矩阵的思维来解决问题。

1.4.2 图形用户界面设计

图形用户界面（Graphical User Interface，GUI）是一个包含设备或图形对象的图形化显示工具，允许用户通过界面与应用程序进行交互式操作。图形用户界面的设备包括鼠标、键盘、显示器等，图形对象可以是菜单、工具条、下压按钮、单选按钮、列表框和滑动条等^[96,97]。用户通过一定的方法（例如点击鼠标或敲击键盘）来选择、激活这些组件，使计算机产生某些动作，例如进行数据处理、绘制近红外光谱图等。MATLAB 软件早期虽然以矩阵运算为主，从第四版开始推出了句柄图形，之后所有的 Demos 都包含友好的用户界面，MATLAB 的图形界面设计功能也日益完善。在 MATLAB 中，GUI 是使用图形对象（Graphical Object）创建的用户界面，这些图形对象包括：按钮、文本框、菜单、滑动条、工具条、坐标轴以及面板等^[98,99]。用户通过与这些按钮和对话框交互实现对仪器的控制操作，软件可以将扫描的光谱数据以谱图形式显示在计算机屏幕上，并且可以把计算得到的样品的漫反射率和吸光度以谱图形式显示。

第四节 本论文选题

综上所述，近红外光谱技术高效、快速、无损和准确的优势使其成为最具前景的研究方向之一。目前国内外的近红外光谱信息软件虽然种类较多，但是并没有嵌入模型建立和预测功能，无法实现光谱在线检测功能。在农业、制药、质量控制等行业领域中的一些应用需要将仪器带到分散的分析现场使用，为满足这一要求，大量体积小、便于携带的近红外光谱仪器被研制出来，这些仪器种类繁多，采用了很多不同的原理，和不同的光路结构，与此同时，这些仪器的分辨率低，扫描结果的准确度有限。本论文的具体研究工作如下：

- 1、使用MATLAB软件开发近红外仪器的配套软件，并嵌入模型的建立和预

测等功能，使用户通过一些按钮和对话框就能操作近红外仪器，并在线建模和预测，将会简化工作流程。

2、使用便携式近红外光谱仪器结合富集技术定量检测稀溶液中的重金属离子，为实际应用中使用便携式近红外仪器进行定量分析提供依据。

3、使用便携式近红外光谱仪器对四种维库溴铵样品进行判别分析，通过不同的光谱预处理方法和距离判据组合，找出判别效果最好的模型，从而探索便携式近红外仪器用于准确的判别分析的可行性。

参考文献

- [1] 丁莹. 近红外光谱仪器的研究进展与发展趋向[J]. 红外, 2012, 33(7): 1-5.
- [2] 韩立苹, 须海荣, 倪君. 近红外光谱技术及其在茶叶上应用研究[J]. 茶叶, 2007, 33(1): 4-7.
- [3] 陆婉珍, 袁洪福. 近红外光谱仪器[M]. 北京: 化学工业出版社, 2016: 15-18.
- [4] 吉海彦. 近红外光谱仪器技术[J]. 现代科学仪器, 2001, 1(6): 25-28.
- [5] 张银桥, 张小超. 基于 MSP430 单片机控制的近红外光谱仪用光源[J]. 测控技术, 2009, 28(9): 30-33.
- [6] Fairuz Omar A, Zubir Matjafri M. Development of optical instrument as turbidimeter: a comparative study[J]. Sensor Review, 2012, 32(2): 134-141.
- [7] LU Wan-zhen. Modern Near Infrared Spectroscopy Analytical Technology 2nd ed[M]. Beijing: China Petrochemical Press, 2006:16-20.
- [8] Liu R, Deng B, Chen W, *et al.* Next step of non-invasive glucose monitor by NIR technique from the well controlled measuring condition and results[J]. optical and quantum electronics, 2005, 37(13-15): 1305-1317.
- [9] Deutsch E R, Reyes D, Schildkraut E R, *et al.* High-resolution miniature FTIR spectrometer enabled by a large linear travel MEMS pop-up mirror[J]. SPIE Defense, Security, and Sensing, 2009, 7319(12):1-8.
- [10] Manzardo O, Herzig H P, Guldemann B, *et al.* New design for an integrated fourier transform spectrometer[J]. Micromachining and Microfabrication, 2000, 5(12): 310-319.
- [11] Weeranantanaphan J, Downey G, Allen P, *et al.* A review of near infrared spectroscopy in muscle food analysis: 20052010[J]. J Near Infrared Spectrosc, 2011, 19(2): 61-64.
- [12] Ahlström A, Wikland M, Rogberg L, *et al.* Cross-validation and predictive value of near-infrared spectroscopy algorithms for day-5 blastocyst transfer[J]. Reproductive biomedicine online, 2011, 22(5): 477-484.
- [13] Armenta S, Garrigues S, De La Guardia M. Determination of edible oil parameters by near infrared spectrometry[J]. Anal Chim Acta, 2007, 596(2): 330-337.
- [14] Casselgren J, Rosendahl S, Sjö Dahl M, *et al.* Road condition analysis using NIR illumination and compensating for surrounding light[J]. Optics and Lasers in Engineering, 2016, 77: 175-182.
- [15] Egusquiguirre S P, Beziere N, Pedraz J L, *et al.* Optoacoustic imaging enabled biodistribution study of cationic polymeric biodegradable nanoparticles[J]. Contrast media & molecular imaging, 2015, 10(6): 421-427.
- [16] González-Martín I, Villaescusa-García V, López-González F, *et al.* Control of quality and silo storage of sunflower seeds using near infrared technology[J]. grasas y aceites, 2013, 64(1):

- 30-35.
- [17] Kurth C D, Levy W J, Mccann J. Near-infrared spectroscopy cerebral oxygen saturation thresholds for hypoxia–ischemia in piglets[J]. *Journal of Cerebral Blood Flow & Metabolism*, 2002, 22(3): 335-341.
- [18] Ma Y, Bai R, Du G, *et al.* Rapid determination of four tobacco specific nitrosamines in burley tobacco by near-infrared spectroscopy[J]. *Analytical Methods*, 2012, 4(5): 1371-1376.
- [19] Ouyang A-G. Partial least squares regression variable screening studies on apple soluble solids nir spectral detection[J]. *Spectroscopy and Spectral Analysis*, 2012, 32(10): 2680-2684.
- [20] Sinelli N, Cerretani L, Di Egidio V, *et al.* Application of near (NIR) infrared and mid (MIR) infrared spectroscopy as a rapid tool to classify extra virgin olive oil on the basis of fruity attribute intensity[J]. *Food research international*, 2010, 43(1): 369-375.
- [21] Altieri G, Genovese F, Admane N, *et al.* On-line measure of donkey's milk properties by near infrared spectrometry[J]. *LWT-Food Science and Technology*, 2016, 69: 348-357.
- [22] Arakawa M, Yamashita Y, Funatsu K. Genetic algorithm - based wavelength selection method for spectral calibration[J]. *J Chemom*, 2011, 25(1): 10-19.
- [23] Balabin R M, Smirnov S V. Variable selection in near-infrared spectroscopy: benchmarking of feature selection methods on biodiesel data[J]. *Anal Chim Acta*, 2011, 692(1): 63-72.
- [24] Cao H, Wang Y, Yang S, *et al.* A wavelength selection method based on random decision particle swarm optimization with attractor for near - infrared spectral quantitative analysis[J]. *J Chemom*, 2015, 29(5): 289-299.
- [25] Cernuda C, Lughofer E, Hintenaus P, *et al.* Enhanced genetic operators design for waveband selection in multivariate calibration based on NIR spectroscopy[J]. *J Chemom*, 2014, 28(2): 123-136.
- [26] Diniz P H G D, Gomes A A, Pistonesi M F, *et al.* Simultaneous classification of teas according to their varieties and geographical origins by using NIR spectroscopy and SPA-LDA[J]. *Food analytical methods*, 2014, 7(8): 1712-1718.
- [27] Goyens C, Jamet C, Schroeder T. Evaluation of four atmospheric correction algorithms for MODIS-Aqua images over contrasted coastal waters[J]. *Remote Sensing of Environment*, 2013, 131: 63-75.
- [28] Jiang H, Liu G, Mei C, *et al.* Measurement of process variables in solid-state fermentation of wheat straw using FT-NIR spectroscopy and synergy interval PLS algorithm[J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2012, 97: 277-283.
- [29] Li J, Huang W, Chen L, *et al.* Variable selection in visible and near-infrared spectral analysis for noninvasive determination of soluble solids content of ‘Ya’pear[J]. *Food Analytical Methods*, 2014, 7(9): 1891-1902.
- [30] Xing G, Cao J, Wang D, *et al.* Near Infrared Spectroscopic Combined with Partial Least Squares and Radial Basis Function Neural Network to Analyze Paclitaxel Concentration in Rat Plasma[J]. *Combinatorial chemistry & high throughput screening*, 2015, 18(8): 704-711.

- [31] Beghi R, Giovenzana V, Marai S, *et al.* Rapid monitoring of grape withering using visible near - infrared spectroscopy[J]. J Sci Food Agric, 2015, 95(15): 3144-3149.
- [32] Ignat T, Schmilovitch Z, Feföldi J, *et al.* Nonlinear methods for estimation of maturity stage, total chlorophyll, and carotenoid content in intact bell peppers[J]. Biosystems engineering, 2013, 114(4): 414-425.
- [33] Li L, Liu S, Peng Y, *et al.* Overview of principal component analysis algorithm[J]. Optik-International Journal for Light and Electron Optics, 2016, 127(9): 3935-3944.
- [34] Li Y, Guo M, Shi X, *et al.* Online near-infrared analysis coupled with MWPLS and SiPLS models for the multi-ingredient and multi-phase extraction of licorice (Gancao)[J]. Chinese medicine, 2015, 10(1): 1-10.
- [35] Lyapustin A I, Wang Y, Laszlo I, *et al.* Multi-angle implementation of atmospheric correction for MODIS (MAIAC): 3. Atmospheric correction[J]. Remote Sensing of Environment, 2012, 127(2): 385-393.
- [36] Minasny B, Mcbratney A B, Bellon-Maurel V, *et al.* Removing the effect of soil moisture from NIR diffuse reflectance spectra for the prediction of soil organic carbon[J]. Geoderma, 2011, 167(12): 118-124.
- [37] Miralbés C. Discrimination of European wheat varieties using near infrared reflectance spectroscopy[J]. Food Chem, 2008, 106(1): 386-389.
- [38] Pontes M J C, Rocha A M J, Pimentel M F, *et al.* Determining the quality of insulating oils using near infrared spectroscopy and wavelength selection[J]. Microchem J, 2011, 98(2): 254-259.
- [39] Rosa S S, Barata P A, Martins J M, *et al.* Development and validation of a method for active drug identification and content determination of ranitidine in pharmaceutical products using near-infrared reflectance spectroscopy: A parametric release approach[J]. Talanta, 2008, 75(3): 725-733.
- [40] Sánchez M-T, Pérez-Marín D, Flores-Rojas K, *et al.* Use of near-infrared reflectance spectroscopy for shelf-life discrimination of green asparagus stored in a cool room under controlled atmosphere[J]. Talanta, 2009, 78(2): 530-536.
- [41] Wang J, Liu J, Liu G, *et al.* Near-infrared spectroscopy using digital phase-sensitive detection[J]. Instrumentation Science & Technology, 2016, 44(2): 199-209..
- [42] 许禄, 邵学广. 化学计量学 (第二版) [M]. 北京: 科学出版社, 2004.
- [43] 梁逸曾, 俞汝勤. 化学计量学. 北京: 高等教育出版社[M], 2003.
- [44] Albuquerque J S, Pimentel M F, Silva V L, *et al.* Silicone sensing phase for detection of aromatic hydrocarbons in water employing near-infrared spectroscopy[J]. Anal Chem, 2005, 77(1): 72-77.
- [45] De Bei R, Cozzolino D, Sullivan W, *et al.* Non - destructive measurement of grapevine water potential using near infrared spectroscopy[J]. Australian Journal of Grape and Wine Research, 2011, 17(1): 62-71.

- [46] Freschet G T, Barthès B G, Brunet D, *et al.* Use of near infrared reflectance spectroscopy (NIRS) for predicting soil fertility and historical management[J]. *Communications in soil science and plant analysis*, 2011, 42(14): 1692-1705.
- [47] Garde-Cerdán T, Lorenzo C, Alonso G L, *et al.* Employment of near infrared spectroscopy to determine oak volatile compounds and ethylphenols in aged red wines[J]. *Food Chem*, 2010, 119(2): 823-828.
- [48] Guo W-L, Du Y-P, Zhou Y-C, *et al.* At-line monitoring of key parameters of nisin fermentation by near infrared spectroscopy, chemometric modeling and model improvement[J]. *World J Microbiol Biotechnol*, 2012, 28(3): 993-1002.
- [49] Lomborg C J, Thomsen M H, Jensen E S, *et al.* Power plant intake quantification of wheat straw composition for 2nd generation bioethanol optimization—a near infrared spectroscopy (NIRS) feasibility study[J]. *Bioresour Technol*, 2010, 101(4): 1199-1205.
- [50] Maltesen M J, Bjerregaard S, Hovgaard L, *et al.* Multivariate analysis of phenol in freeze-dried and spray-dried insulin formulations by NIR and FTIR[J]. *AAPS PharmSciTech*, 2011, 12(2): 627-636.
- [51] Martins A, Daffner L A, Fenech A, *et al.* Non-destructive dating of fiber-based gelatin silver prints using near-infrared spectroscopy and multivariate analysis[J]. *Analytical and bioanalytical chemistry*, 2012, 402(4): 1459-1469.
- [52] Peng X, Chen H. Rapid estimation of single cell oil content of solid-state fermented mass using near-infrared spectroscopy[J]. *Bioresour Technol*, 2008, 99(18): 8869-8872.
- [53] Pfautsch S, Macfarlane C, Ebdon N, *et al.* Assessing sapwood depth and wood properties in *Eucalyptus* and *Corymbia* spp. using visual methods and near infrared spectroscopy (NIR)[J]. *Trees*, 2012, 26(3): 963-974.
- [54] Sarraguça M C, Cruz A V, Soares S O, *et al.* Determination of flow properties of pharmaceutical powders by near infrared spectroscopy[J]. *J Pharm Biomed Anal*, 2010, 52(4): 484-492.
- [55] Smyth H, Cozzolino D, Cynkar W, *et al.* Near infrared spectroscopy as a rapid tool to measure volatile aroma compounds in Riesling wine: possibilities and limits[J]. *Analytical and bioanalytical chemistry*, 2008, 390(7): 1911-1916.
- [56] Sulub Y, Small G W. Spectral simulation protocol for extending the lifetime of near-infrared multivariate calibrations[J]. *Anal Chem*, 2009, 81(3): 1208-1216.
- [57] Wang K, Chi G, Lau R, *et al.* Multivariate calibration of near infrared spectroscopy in the presence of light scattering effect: a comparative study[J]. *Anal Lett*, 2011, 44(5): 824-836.
- [58] 孙君明, 韩粉霞, 闰淑荣, *et al.* 傅里叶近红外反射光谱法快速测定大豆脂肪酸含量[J]. *光谱学与光谱分析*, 2008, 28(6): 1290-1295.
- [59] Phil Williams, Pierre Dardenne, Vincent Baeten, *et al.* Applications of Near Infrared Spectroscopy in Agriculture, NIR Conference in Brazil, 2015, 101-102.
- [60] Ahmed M K, Levenson J. Application of near-infrared spectroscopy to the quality assurance of ethanol and butanol blended gasoline[J]. *Pet Sci Technol*, 2012, 30(2): 115-121.

- [61] Alves J C L, Poppi R J. Diesel oil quality parameter determinations using support vector regression and near infrared spectroscopy for hydrotreated feedstock monitoring[J]. J Near Infrared Spectrosc, 2012, 20(4): 419-425.
- [62] Mignani A G, Ciaccheri L, Ottevaere H, *et al.* Visible and near-infrared absorption spectroscopy by an integrating sphere and optical fibers for quantifying and discriminating the adulteration of extra virgin olive oil from Tuscany[J]. Analytical and bioanalytical chemistry, 2011, 399(3): 1315-1324.
- [63] Summers D, Lewis M, Ostendorf B, *et al.* Visible near-infrared reflectance spectroscopy as a predictive indicator of soil properties[J]. Ecological Indicators, 2011, 11(1): 123-131..
- [64] 许禄, 邵学广. 化学计量学 (第二版) [M]. 北京: 科学出版社, 2004.
- [65] Blanco M, Coello J, Iturriaga H, *et al.* Determination of water in lubricating oils by mid-and near-infrared spectroscopy[J]. Microchimica Acta, 1998, 128(3-4): 235-239.
- [66] Araujo A M, Santos L M, Fortuny M, *et al.* Evaluation of Water Content and Average Droplet Size in Water-in-Crude Oil Emulsions by Means of Near-Infrared Spectroscopy†[J]. Energy & Fuels, 2008, 22(5): 3450-3458.
- [67] Kelly J J, Barlow C H, Jingui T M, *et al.* Prediction of gasoline octane numbers from near-infrared spectral features in the range 660-1215 nm[J]. Anal Chem, 1989, 61(4): 313-320.
- [68] Balabin R M, Syunyaev R Z, Schmid T, *et al.* Asphaltene adsorption onto an iron surface: combined near-infrared (NIR), Raman, and AFM study of the kinetics, thermodynamics, and layer structure[J]. Energy & Fuels, 2010, 25(1): 189-196.
- [69] Kong S-H, Noh Y-W, Suh Y-S, *et al.* Evaluation of the novel near-infrared fluorescence tracers pullulan polymer nanogel and indocyanine green/ γ -glutamic acid complex for sentinel lymph node navigation surgery in large animal models[J]. Gastric Cancer, 2015, 18(1): 55-64.
- [70] Yip W L, Gausemel I, Sande S A, *et al.* Strategies for multivariate modeling of moisture content in freeze-dried mannitol-containing products by near-infrared spectroscopy[J]. J Pharm Biomed Anal, 2012, 70: 202-211.
- [71] Warnecke S, Rinnan , Allesø M, *et al.* Measurement of active content in escitalopram tablets by a near - infrared transmission spectroscopy model that encompasses batch variability[J]. J Pharm Sci, 2013, 102(4): 1268-1280.
- [72] Burggraef A, Monteyne T, Vervaeke C, *et al.* Process analytical tools for monitoring, understanding, and control of pharmaceutical fluidized bed granulation: A review[J]. European Journal of Pharmaceutics and Biopharmaceutics, 2013, 83(1): 2-15.
- [73] Märk J, Andre M, Karner M, *et al.* Prospects for multivariate classification of a pharmaceutical intermediate with near-infrared spectroscopy as a process analytical technology (PAT) production control supplement[J]. European Journal of Pharmaceutics and Biopharmaceutics, 2010, 76(2): 320-327.
- [74] [1] Koller D M, Posch A, Hörl G, *et al.* Continuous quantitative monitoring of powder

- mixing dynamics by nearinfrared spectroscopy[J]. Powder Technol, 2011, 205(1): 87-96.
- [75] Moerman A, Van Herzele I, Vanpeteghem C, *et al.* Near-infrared spectroscopy for monitoring spinal cord ischemia during hybrid thoracoabdominal aortic aneurysm repair[J]. Journal of Endovascular Therapy, 2011, 18(1): 91-95.
- [76] Bittner L, Heigl N, Petter C, *et al.* Near-infrared reflection spectroscopy (NIRS) as a successful tool for simultaneous identification and particle size determination of amoxicillin trihydrate[J]. J Pharm Biomed Anal, 2011, 54(5): 1059-1064.
- [77] Poellinger A, Burock S, Grosenick D, *et al.* Breast cancer: early-and late-fluorescence near-infrared imaging with indocyanine green—a preliminary study[J]. Radiology, 2011, 258(2): 409-416.
- [78] Li P, Du G, Cai W, *et al.* Rapid and nondestructive analysis of pharmaceutical products using near-infrared diffuse reflectance spectroscopy[J]. J Pharm Biomed Anal, 2012, 70(5): 288-294.
- [79] Leff D R, Orihuela-Espina F, Elwell C E, *et al.* Assessment of the cerebral cortex during motor task behaviours in adults: a systematic review of functional near infrared spectroscopy (fNIRS) studies[J]. NeuroImage, 2011, 54(4): 2922-2936.
- [80] 吉海彦. 近红外光谱仪器技术 [J]. 现代科学仪器, 2001, 20.
- [81] 刘青格, 李署坚, 李雪山. 多尺度边缘检测方法提取近红外光谱信息特征的研究[J]. 光谱学与光谱分析, 2005, 25(6): 894-897.
- [82] Candolfi A, Wu W, Massart D L, *et al.* Comparison of classification approaches applied to NIR-spectra of clinical study lots. J Pharm Biomed Anal, 1998, 16(5): 1329-1347.
- [83] 褚小立, 袁洪福, 陆婉珍. 近年来我国近红外光谱分析技术的研究与应用进展[J]. 分析仪器, 2006, (2): 1-10.
- [84] 丁莹. 近红外光谱仪器的研究进展与发展趋向[J]. 红外, 2012, 33(7): 1-5.
- [85] 陆婉珍, 褚小立. 近红外光谱(NIR)和过程分析技术(PAT)[J]. 现代科学仪器, 2007, 2(4): 13-17.
- [86] Candolfi A, Wu W, Massart D L, *et al.* Comparison of classification approaches applied to NIR-spectra of clinical study lots[J]. Journal of Pharmaceutical & Biomedical Analysis, 1998, 16(8): 1329-1347.
- [87] 齐晓, 韩建国, 李曼莉. 近红外光谱分析仪器的的发展概况[J]. 光谱学与光谱分析, 2007, 27(10): 2022-2026.
- [88] 罗华飞. MATLAB GUI 学习手记(第2版)[M]. 北京:北京航空航天大学出版社, 2011.
- [89] Wang Y, Gao J, Yang H, *et al.* Compensating the Degradation of Near-Infrared Absorption of Black Silicon Caused by Thermal Annealing[J]. Nanoscale research letters, 2016, 11(1): 1-6.
- [90] Goudarzi M A, Cocard M, Santerre R. EPC: Matlab software to estimate Euler pole parameters[J]. Gps Solutions, 2014, 18(1): 153-162.
- [91] Joshi S, Schultz T P. Family planning and women and children health: long-term consequences of an outreach program in Matlab, Bangladesh[J]. Demography, 2013, 7(1): 52-55.

- [92] Michalak P. The simple hourly method of EN ISO 13790 standard in Matlab/Simulink: A comparative study for the climatic conditions of Poland[J]. Lancet, 2014, 375(9731): 2073-2081.
- [93] Rama S, Dilip C S, Perumal R N. A software program to investigate the nucleation kinetics of solution grown crystals using MATLAB platform [J]. Comput Phys Commun, 2014, 185(2): 661-669.
- [94] Tsarapkina D, Jeffrey D J. Exploring Rounding Errors in Matlab Using Extended Precision [J]. Procedia Computer Science, 2014, 29(2): 1423-1432.
- [95] 郭媛君. 基于DMD的微型近红外光谱仪光谱信息处理及其应用软件[D]. 重庆大学, 2011.
- [96] 李俊. 基于 Matlab 的语音信号时域分析系统[J]. Computer Science\&\sapplication, 2013, 03(8): 369-373.
- [97] 刘海洋. 近红外光谱仪基础软件开发[D]. 吉林大学, 2010.
- [98] 裴卫平. 尿液成分的近红外光谱分析原理及实现研究[D]. 重庆大学, 2012.
- [99] 谢彩侠, 胡亚楠, 左春芳等. 近红外漫反射光谱(NIRS)结合 TQ 软件建立盾叶薯蓣中薯蓣皂苷元定量模型[J]. 复旦学报:医学版, 2013, 40(4): 431-435.

第二章 基于 MATLAB 的近红外光谱软件开发

第一节 引言

近红外光谱指波长范围在 780~2526nm 的电磁波,利用分子的倍频和合频吸收来表达物质的结构和组成信息^[1]。随着光学技术、电子技术与计算机软硬件技术、化学计量学的发展,使现代近红外光谱分析成为一种成熟的分析技术,广泛应用于农业^[2-5]、药物产品分析^[6-13]、烟草^[14-18]、石油化工^[19-23]、食品工业^[24-27]、生命科学^[28-31]及环境保护^[32-35]等领域,尤其在测定农产品中的水分、淀粉、脂类、蛋白质与纤维素等成分方面得到了广泛应用。与此同时,近红外光谱仪器也得到了长足发展,市场上各种类型的仪器层出不穷,主要有滤光片型近红外光谱仪器、发光二极管型近红外仪器、光栅单色器型近红外仪器、傅里叶变换型近红外光谱仪器以及声光可调滤光片型近红外光谱仪器以及多通道检测型近红外光谱仪器。这些仪器以各自所具有的不同特点和优势在近红外领域发挥着重要作用。

然而,目前我们使用近红外仪器采集到样品光谱后,通常需要先将光谱数据导入第三方数据分析软件中然后才能进行数据处理和分析,过程比较繁琐,在一定程度上降低了工作效率,对于在一些需要现场检测分析的在线使用情形就更为不便。因而有必要开发一个集光谱采集、数据处理分析于一体的综合型近红外软件,操作者只需要通过简单的操作就可以完成光谱的采集、定性定量模型的建立和校正、预测以及生成报告等一系列工作,简化了工作流程。

MATLAB 是矩阵实验室的简称,和 Mathematica、Maple 并成为三大数学软件,它在数学类科技应用软件中的数值计算方面首屈一指。MATLAB 可以进行矩阵运算、实现算法、创建用户界面等,主要应用于工程计算、控制设计、信号处理与通讯、图像处理、信号检测、金融建模设计与分析等领域^[36-40]。借助于可视化 GUIDE 工具箱, MATLAB 可以便捷地设计出直观、方便的菜单化和控件式的人机交互界面。

本章主要介绍借助于 MATLAB 平台设计近红外软件的过程、软件的框架、设计思路、设计原则以及各主要功能模块的作用。

第二节 软件设计

2.2.1 设计原则和模式

软件设计是依赖编程或者代码语言优雅地表达并且解决现实需求的一门科学和艺术。软件应当很容易被人理解和接受。每一款软件的功能定位和目标群体不同，所以其操作界面也是有差别的，而且各具特色。然而，软件设计的出发点和所要达到的最终目的是一致的，那就是所设计的界面应遵循以下的特性：

(1) 直观性

整个软件界面简单明了，力争使用户看一次界面就能理解如何操作。软件以提高工作效率，减少操作者的键盘输入频次为主旨。

(2) 可靠性

可靠性是指在常规的软件设计中应用各种方法和技术使软件能够在规定的条件下和规定的时间区间完成规定功能的能力。

(3) 交互性

交互性是一种人与计算机之间传递、交换信息的媒介和对话接口，是评价软件性能的重要指标。

设计模式是一系列被反复使用的、大多数人认可的设计经验的总结，使用设计模式的根本目的是为了代码的重复使用、更易被他人理解以及更加可靠。软件模式等于一定条件下所出现的问题以及解法。一般情况下，在软件结构设计中应遵循以下原则：

(1) 单一职责原则

单一职责原则是实现低耦合、高内敛的指导方针，很多代码重用的脚本中都能看到它的影子，需要发现待解决问题之间不同职责并将其分离。

(2) 开闭原则

一个软件实体应当对扩展开放，对修改封闭。在设计一个模块时，应当尽可能使这模块可以在不被改动的情况下实现功能性扩展，即实现在不修改源代码的前提下重新定义这模块的行为。

(3) 依赖倒转原则

依赖倒转原则是指在设计高层模块时，不要过分依赖于底层模块，而应该依赖于抽象，而抽象不应该依赖于细节。

2.2.2 设计思路

(1) 首先确定软件所要实现的功能, 以及为了实现这些功能需要哪些具体模块。本软件的目的是为了与近红外仪器的通讯, 通过软件操作仪器来扫描光谱, 以及实现判别和定量模型的在线建立、样品的在线预测、生成报告等功能。因此, 本软件的模块主要包括光谱扫描模块、模型建立模块、预测模块以及生成建模结果报告模块。其中光谱扫描模块用以实现与仪器建立连接, 仪器相关参数设置, 仪器自检, 光谱文件存放位置和存放格式设置, 自动扫描功能等与光谱的取得和存放有关的功能; 模型建立模块主要是对光谱数据进行化学计量学处理, 包括各种光谱预处理方法, 判别模型所用到的常用距离判别、定量模型的主因子数选择、建模效果的图形表达以及模型的保存; 预测模块主要用以使用已有模型来处理未知样品的光谱数据并将预测结果以图形的形式展示出来; 报告生成模块是将所建立模型的名称、时间、性质等各种参数信息、建模效果信息以及预测信息(如果进行了预测)等以文档的形式输出到计算机硬盘当中, 以便之后查看。

(2) 绘制用户界面草图, 并根据直观性、交互性原则对其修改, 最终确定一个可行性最强的方案。本软件所使用的界面可分为菜单部分、按钮部分、扫描光谱名称显示、光谱名称编辑、谱图展示以及历史记录部分。其中菜单部分包含了软件中所有的功能, 而工具栏和按钮部分则是将最常用的功能单独布局, 以方便软件操作者使用, 光谱名称显示部分用以标记扫描过的光谱名称并将名称与右侧的光谱图对应起来, 光谱名称编辑是供操作者编辑光谱名称的接口, 谱图展示区域不仅可以展示扫描到的光谱, 而且可以通过切换下方的单选按钮来切换不同的显示区域, 包括与模型建立有关的图形, 预测的结果等; 历史记录区域用以显示最近进行的 20 项操作动作, 用户可通过这个窗口来查看最近进行的操作。

(3) 使用 MATLAB GUI 创建交互界面, 对每个模块进行代码实现并逐项进行检测。在编写各个模块的代码时, 首先梳理各模块之间的逻辑关系, 规划好实现的先后顺序, 同时由于 MATLAB 的瓶颈之一在于循环效率过低, 在设计代码时, 尽可能以矩阵的思想来完成任务从而避免循环的出现。

2.2.3 窗口界面的设计

窗口界面总体布局按照功能分为操作区和图形显示区两大部分，界面左方区域主要为用户的操作提供各种接口，右方大部分区域用以显示相应操作输出的图形。界面整体设计布局如图 2.1 所示：

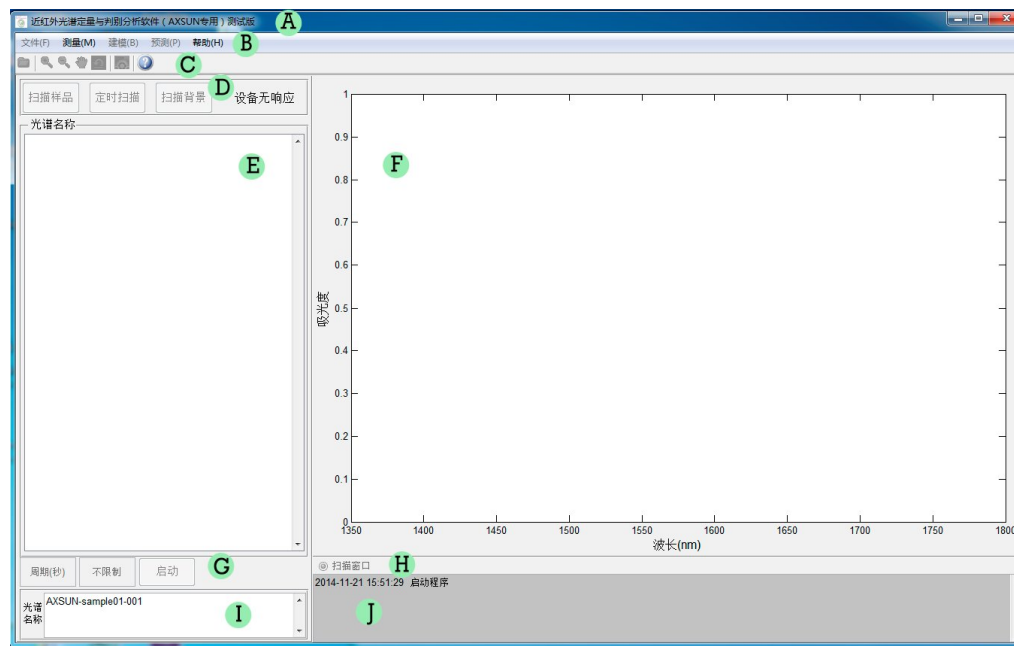


图 2.1 软件界面

- A) 软件用户界面的标题栏，显示本软件的名称。
- B) 集成本软件所有功能的下拉菜单，本软件所支持的所有操作都可以在这个栏目的子菜单下找到，包括文件、测量、建模、预测以及帮助五个大类。
- C) 快速调用常用工具，从左往右依次为打开文件工具，放大、缩小、移动以及复位图像工具，仪器参数设置工具以及帮助工具。
- D) 当前仪器状态指示，根据当前设备状态显示以下不同的信息：
- 未安装驱动（指仪器的硬件驱动）
 - 设备未响应
 - 设备预热中
 - 设备已就绪
- 扫描样品，定时扫描以及扫描背景这三个相对其他功能更为常用，将其单独

拿出，置于界面上，方便用户操作。

E) 用以显示当前已扫描样品的名称，在该区域右击鼠标右键还可以实现删除光谱以及打开存放光谱文件的目标文件夹等功能。

F) 用以显示当前已扫描样品的光谱图，在该位置还有若干个显示图形的坐标轴。

G) 定时扫描功能。考虑到有些情况下需要对样品重复多次扫描，此时用户只需要设置将要进行的扫描次数和两次扫描之间的时间间隔就可以使软件自动工作。

H) 单选按钮区域，用以切换 F 区域的图像显示。

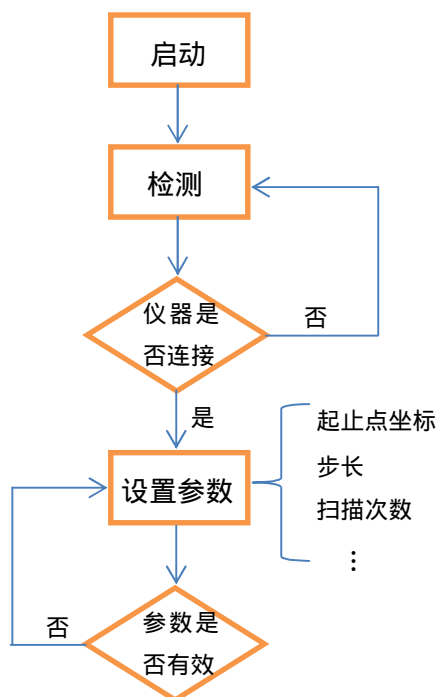
I) 该区域用以显示待扫描样品的名称。

J) 报告窗口，记录并显示用户所执行过的最近的 20 条操作，保存在本软件所在目录下的“CommandHistory.txt”。

2.2.4 主要模块的设计

(1) 测量模块

测量模块主要实现与仪器间通讯以及采集光谱等功能，是本软件的第一个模块，也是最主要的模块，其下设四个子模块：与仪器建立连接模块，仪器参数设置模块、存放光谱文件设置模块以及扫描模块。本模块的流程图如下，



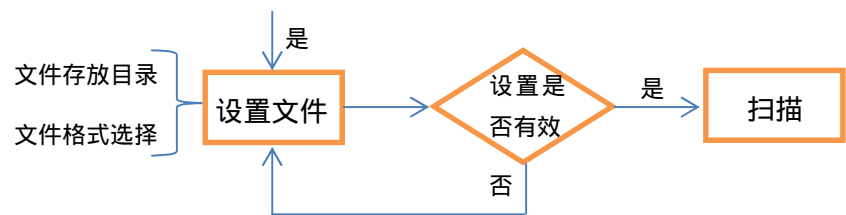


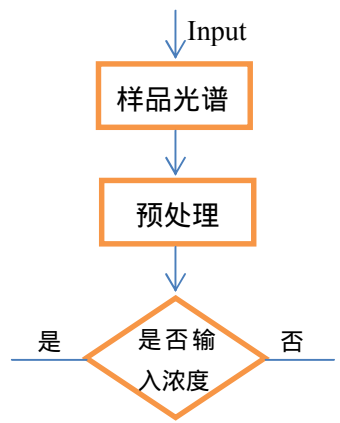
图 2.2 测量模块流程图

本模块以仪器连接与否为标识建立了“熔断”机制。只有当软件与仪器成功建立连接之后，该模块的其他子模块才被激活，而一旦断开连接，则本模块所有功能即刻停止，进入待激活状态。这样的设计在保证扫描结果准确性的同时可以尽量避免未知错误的发生。在连接上仪器之后，设置参数模块被激活，用户此时可以对仪器的各项参数进行设置，包括选择波长范围（注意，不能超过仪器固有的上下限范围），设置两个波长点之间的步长，扫描次数等参数，完成之后，用户需要设置存放光谱文件的路径以及选择光谱文件以何种格式保存（诸如 SPC，SPA，CSV 等）。这些设置操作如果用户未做出选择，系统将按照默认设置运行，光谱文件将被保存至程序所在目录。最后，扫描按钮将被激活，用户就可以进行扫描操作了。如果软件非初次运行，程序会自动使用最近一次的设置参数。

准确并且及时地判断仪器是否连接是最基础的工作，同时也是本次软件开发过程中难度最大的部分，它直接关系到软件系统的稳定性以及其他子模块的可用性。在对软件调试的过程中，相比于其他模块，这部分所需要考虑的内容更多，这是最后使用“熔断”机制的直接原因。

（2）数据处理模块

本模块用来建立定性和定量模型，包括样品光谱的选择，光谱预处理，建立相应模型以及保存模型。本模块的流程图如下，



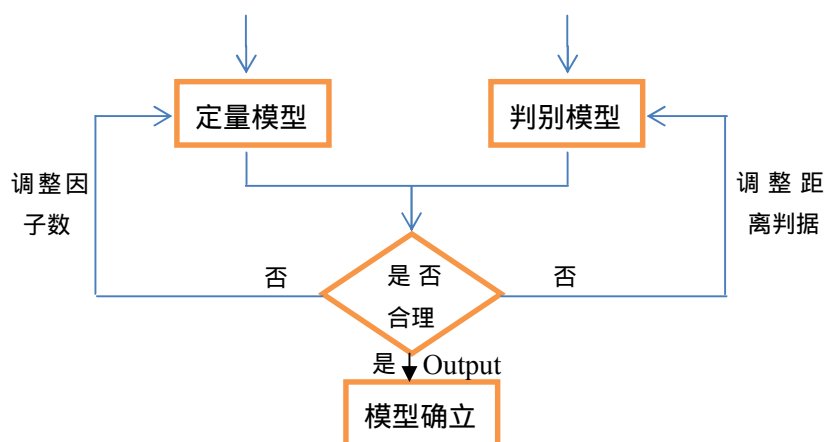


图 2.3 建模模块流程图

本模块首先要求用户给出建立模型所需要的光谱文件，可通过“建模”→“确定校正集”组合菜单来指定目标光谱，然后光谱预处理菜单被激活，该菜单下集合了若干种化学计量学经常用到的光谱预处理方法，包括连续小波变换（CWT）、标准正态变量（SNV）等以及它们的组合，用户可以根据自己的需要调整输入参数以达到最佳的预处理效果。预处理完成之后，程序会根据之前用户是否输入浓度来判断要建立定量模型还是判别模型（如果用户输入了浓度则建立定量模型，否则建立定性模型），进而激活相应的菜单项。建立定量模型时，执行“PLS 定量模型”菜单，程序会将模型通过留一交叉验证得到的预测值与实际值之间的关系图画出，并将模型的相关系数与交叉验证均方根误差以图例形式标识出来，如果模型不合理（例如相关系数远低于 1），可以通过改变主因子数等方式来重新构建定量模型；建立判别模型时，执行“判别分析”菜单下的若干子菜单，这些子菜单包含了常用到的距离判据，包括欧氏距离、马氏距离、街区距离以及对应的主成分空间距离等，可通过比较不同距离判据下的模型效果来选择最优的判据。模型建立完成之后，可以以独立的文件形式保存至软件之外，以便下次直接调用而不需要重新建立模型。

本模块通过将建立模型中所需要的输入参数设计为菜单形式，简化了用户建立模型的难度和步骤，只需要执行相应的菜单组合，便可以迅速建立相应模型并对模型进行优化。

（3）预测模块

模型建立完成之后，预测模块被激活，通过输入待预测样品光谱，程序可以依据模型给出对应的预测结果，预测结果以图形的形式在程序窗口中展示，相关系数，预测均方根误差等关键参数以图例形式标识，并且将预测结果以文本方式在程序之外生成。本模块的流程图如下所示：

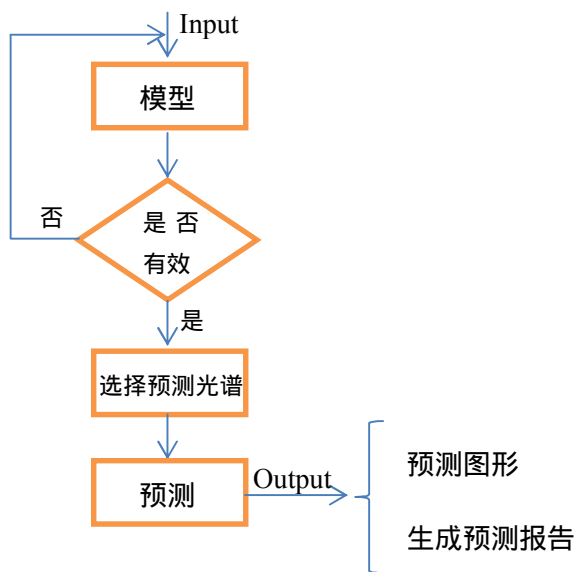


图 2.4 预测模块流程图

有两种方式为程序输入模型：1.直接读取一个保存在程序之外的模型文件，这个文件必须由本程序按照特定的数据结构生成（在本程序中是一个 MAT 格式的文件）；2.通过数据处理模块，现建一个模型，然后用以预测未知样品。预测结果的具体数值（定量模型是预测的浓度值，定性模型是预测的距离值）以表格的形式在图形的左上角显示。预测完成之后“生成报告”菜单将被激活，该菜单将启动 word 形式的性能报告模板，保存数据和图像等运行结果至所指定的硬盘位置。

本模块用以使用模型预测未知样品，并将预测结果输出。在所生成的预测报告中包含了有关模型预测的全部信息，通过浏览报告文件，可以直观地了解模型的性质、质量以及预测结果等信息。

第三节 近红外光谱软件测试

2.3.1 光谱测量

启动软件，在连接仪器之前，工具栏中部分工具以及“扫描按钮”、“定时扫描”、“扫描背景”三个按钮处于灰色非激活状态，右侧信息栏中显示“设备未连接”字样，如图 2.5 所示；将仪器与程序连接，这些按钮被激活，执行“测量”菜单下的仪器参数设置选项，弹出仪器参数设置对话框如图 2.6，填写合理的参数，点击“确定”保存设置，然后就可以通过程序控制仪器进行扫描操作了，图 2.7 为扫描一个样品得到的谱图。

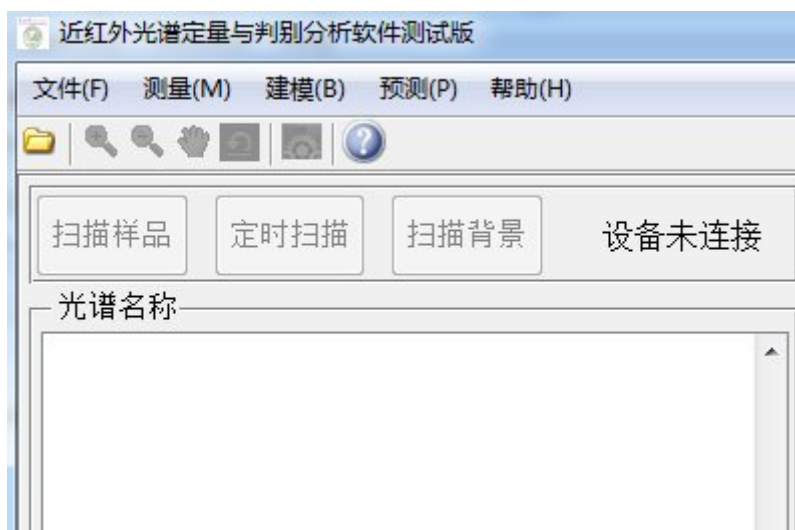


图 2.5 连接仪器之前



图 2.6 参数设置

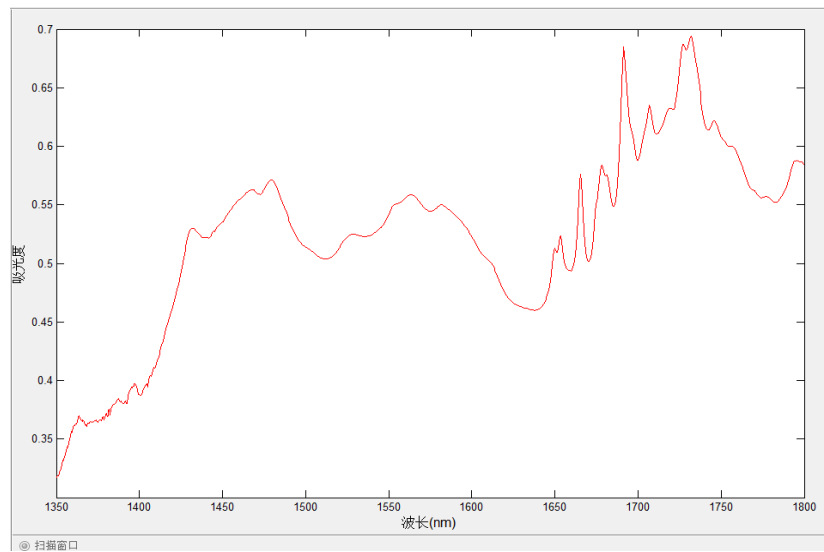


图 2.7 样品光谱

2.3.2 建立定量模型

执行“建模”-“确定校正集”菜单，弹出如图 2.8 所示确定校正集窗口，用户可通过按钮选择要用于建立模型的样品光谱，然后程序会根据是否输入了浓度值来激活不同的菜单项（定性定量模型的菜单选项是不同的）。以建立定量模型为例，接下来选择光谱预处理方法，执行如图 2.9 所示的预处理菜单，默认情况下是 None，即不做任何预处理，选择第二种 SNV 方法（标准正态变换）处理光谱，结果如图 2.10 所示，然后执行“PLS 定量模型”菜单建立定量模型，建模结果如图 2.11，此时如果对建模结果不满意，可以通过调整因子数或者选择不同的预处理方法来重新建立模型直到模型符合要求为止。

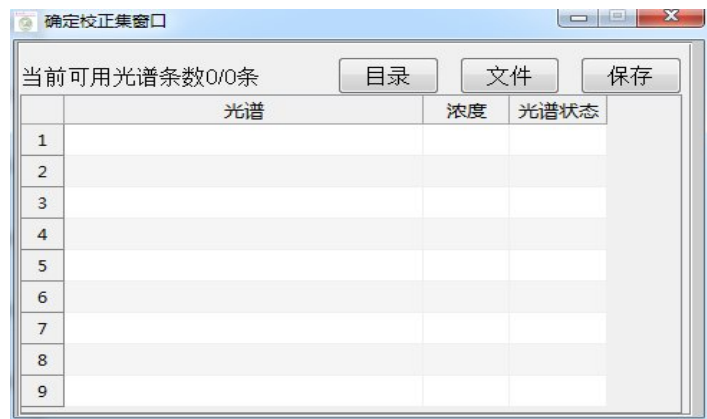


图 2.8 确定校正集窗口

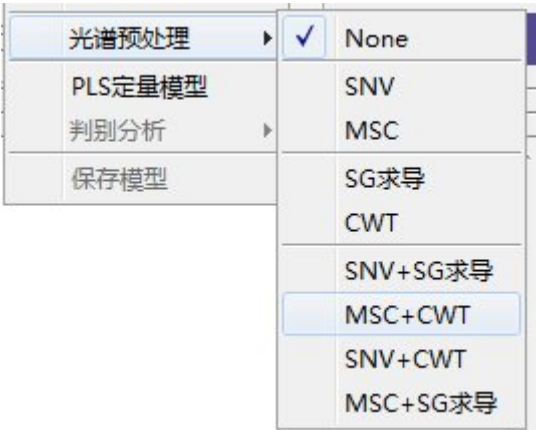


图 2.9 预处理菜单

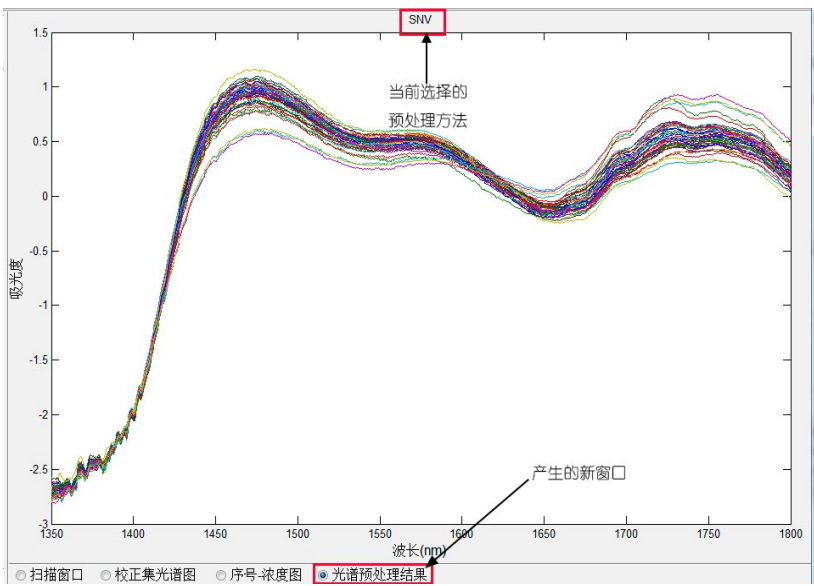


图 2.10 SNV 预处理结果

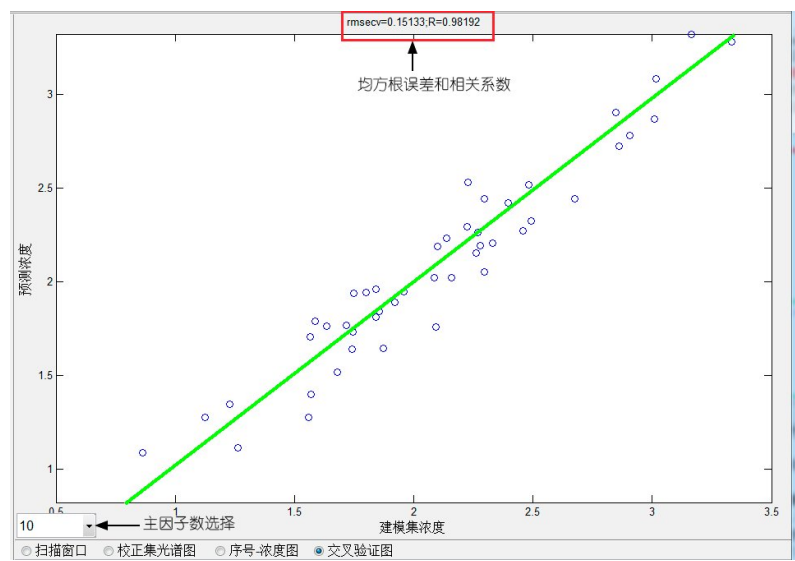


图 2.11 模型建立（定量）

执行预测菜单中的“选择光谱”菜单，确定用以预测的样品光谱，程序将生成如图 2.14 所示的结果图。主坐标轴中显示被预测样品光谱的谱图，左上角以表格形式将被预测光谱的名称与预测角度具体数值列示出来。

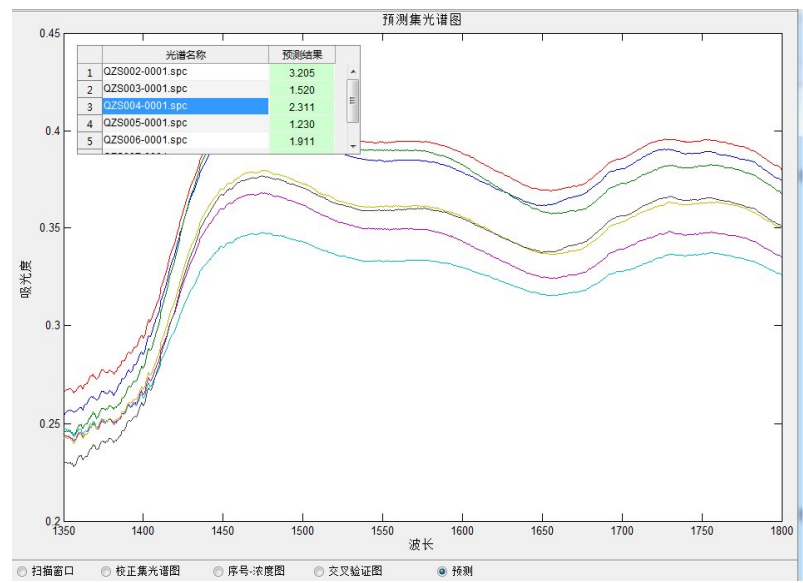


图 2.12 预测结果

2.3.3 建立判别模型

判别分析与定量分析的步骤基本相同，在选择校正集和光谱预处理时是相同的，在预处理完成之后，在“判别分析”菜单下选择不同的距离判据建立判别模型，如图 2.13 所示，如果某个距离判据下的模型效果不好，可通过更换距离判据来重新建立模型直到符合要求为止。建模结果如图 2.14 所示。

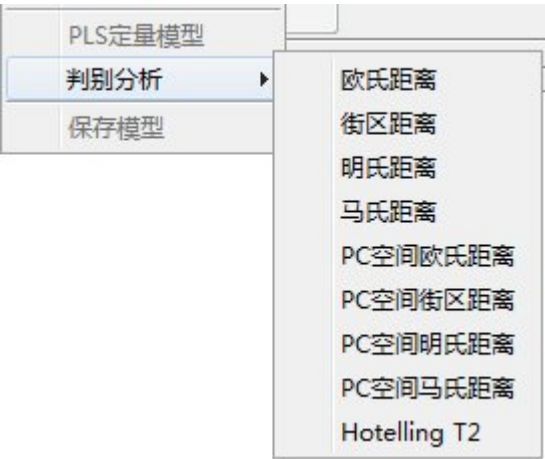


图 2.13 距离判据菜单

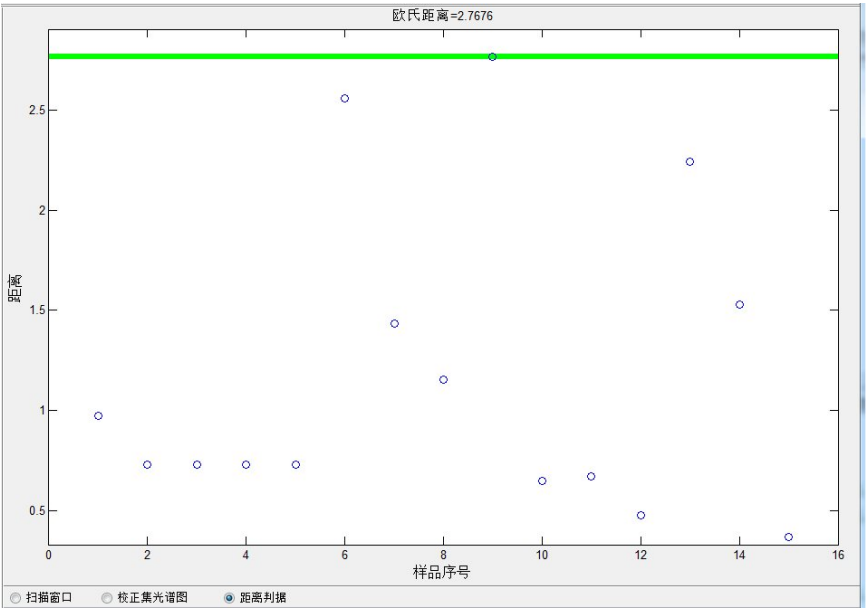


图 2.14 模型建立（判别）

通过执行“选择光谱”菜单，程序会使用建立好的判别模型对所选择的光谱进行处理，结果如图 2.15 所示，与定量模型直接显示预测浓度值不同的是，

判别模型显示的预测结果为“是”或者“否”，以表示对应样品与建模集样品是否属于同一类。

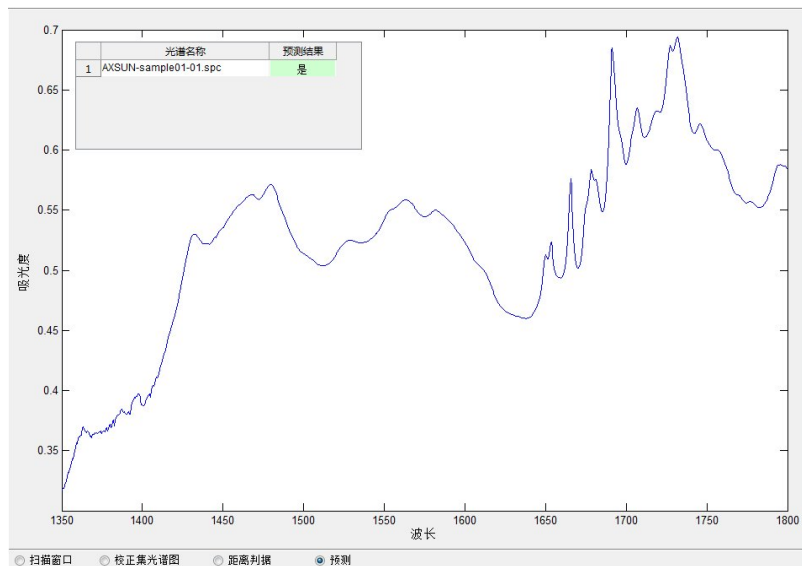


图 2.15 预测结果

2.3.4 生成建模报告

当经过筛选模型参数，确认模型满足要求之后，用户通常需要获取有关该模型的分析报告。为此，本软件利用 MATLAB 数据接口，将数据和图像链接到对应的 word 模板文档，当点击“生成报告”菜单时，系统将有关报告的时间、建立模型的光谱预处理方法、模型的性质（定量/定性）、主因子数等信息写入该 word 文档，图 2.16 所示为建立定量模型时所生成的一个报告截图。

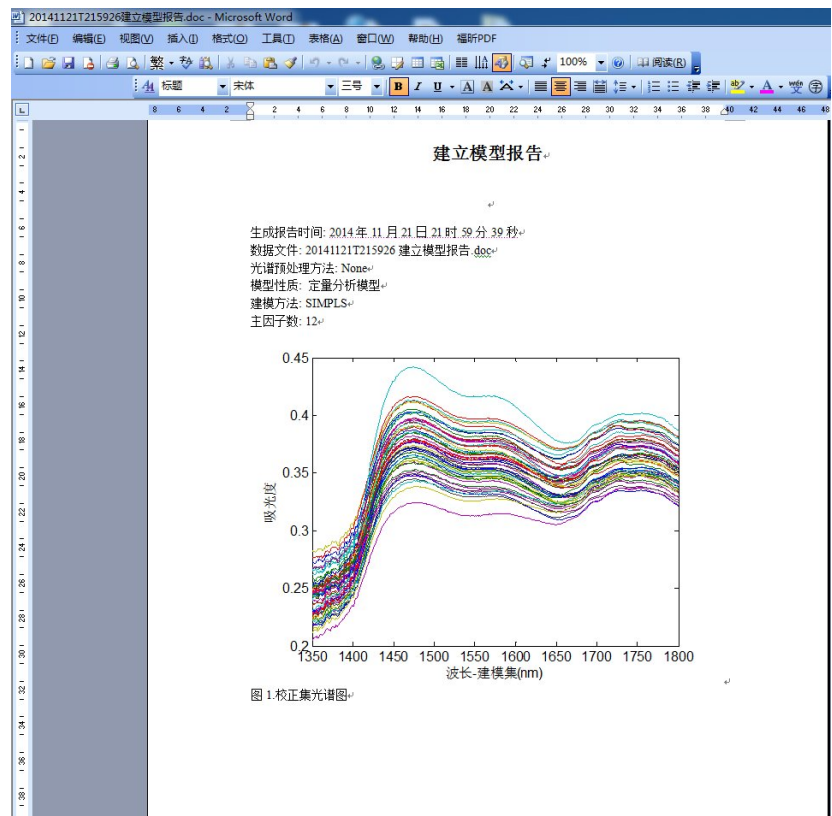


图 2.16 生成报告

第四节 结论

本章首先介绍了近红外仪器配套软件开发的意义。其次，详细介绍了基于 MATLAB 环境进行软件开发的思路、原则、软件的框架以及应用。通过开发近红外仪器配套软件，可以方便仪器的用户按照自己的要求来操作仪器，简化了操作的流程；通过将模型建立的功能集成到本程序中，使操作者只需要点击若干个按钮就可以建立起各种定性定量模型，并预测未知样品，将预测结果以文档报告的形式保存下来，便于查看。

参考文献

- [1] 陆婉珍, 袁洪福, 徐广通等. 现代近红外光谱分析技术[M]. 北京: 中国石化出版社, 2000.
- [2] 梁高峰, 贾宏汝, 谷运红, 秦广雍, 王卫东. 近红外光谱分析技术及其在农业研究中的应用. 安徽农业科学 2007, (29), 9113-9115.
- [3] Bo S, Jonsson A, Börjesson T. Use of near infrared reflectance spectroscopy to predict nitrogen uptake by winter wheat within fields with high variability in organic matter[J]. Plant & Soil, 2005, 269(1): 251-258.
- [4] Chen H, Tu Z, Qing Z, *et al.* Feasibility Study of Veterinary Drug Residues in Honey by NIR Detection[J]. Ifip Advances in Information & Communication Technology, 2013, 392: 150-156.
- [5] Chen X J, Di W U, Yong H E, *et al.* Study on Discrimination of Tea Based on Color of Multispectral Image[J]. Guang pu xue yu guang pu fen xi = Guang pu, 2008, 28(11): 2527-30.
- [6] Chong A K, Mathieu R. Near infrared photography for craniofacial anthropometric landmark measurement[J]. Photogrammetric Record, 2006, 21(113): 16-28.
- [7] Godinho M S, Blanco M R, Neto F F G, *et al.* Evaluation of transformer insulating oil quality using NIR, fluorescence, and NMR spectroscopic data fusion[J]. Talanta, 2014, 129(1): 143-149.
- [8] Huang Y, Min S G, Cao J L, *et al.* Near-infrared imaging for quantitative analysis of active component in counterfeit dimethomorph using partial least squares regression[J]. Chemical Papers, 2012, 66(11): 1065-1072.
- [9] Inarejos-García A M, Gómez-Alonso S, Fregapane G, *et al.* Evaluation of minor components, sensory characteristics and quality of virgin olive oil by near infrared (NIR) spectroscopy[J]. Food Research International, 2013, 50(1): 250-258.
- [10] Jia D, Yue H, Li Z, *et al.* Determination of 27 chemical constituents in Chinese southwest tobacco by FT-NIR spectroscopy[J]. Industrial Crops & Products, 2012, 40(1): 21-26.
- [11] Lankmayr E, Mocak J, Serdt K, *et al.* Chemometrical classification of pumpkin seed oils using UV-Vis, NIR and FTIR spectra[J]. Journal of Biochemical & Biophysical Methods, 2004, 61(1-2): 95-106.
- [12] Li J. Chiroptical Activity of Plasmonic Silver Nanocaps and Gas Sensing[J]. Faraday Discuss, 2015, 20(3): 201-206.
- [13] Liu T. Determination of Quality Properties of Soy Sauce by Support Vector Regression Coupled with SW-NIR Spectroscopy[J]. Chemical Research in Chinese Universities, 2011, 27(3): 385-391.
- [14] Liu X, Chen H C, Liu T A, *et al.* Application of PCA-SVR to NIR Prediction Model for

- Tobacco Chemical Composition[J]. Spectroscopy and Spectral Analysis, 2007, 27(12): 2460-2463.
- [15] Liu Y, Luo W, Wang W, *et al.* Quantitative Analysis of Uncoated Eszopiclone Tablets by Near-Infrared Spectroscopy[J]. Anal Lett, 2014, 47(11): 1938-1951.
- [16] Magwaza L S, Opara U L, Terry L A, *et al.* Evaluation of Fourier transform-NIR spectroscopy for integrated external and internal quality assessment of Valencia oranges[J]. Sage Open, 2013, 31(1): 144-154.
- [17] Mantanus J, Ziémons E, Rozet E, *et al.* Building the quality into pellet manufacturing environment – Feasibility study and validation of an in-line quantitative near infrared (NIR) method[J]. Talanta, 2010, 83(2): 305-311.
- [18] Mazur L, Peralta-Zamora P G, Demczuk B, *et al.* Application of multivariate calibration and NIR spectroscopy for the quantification of methylxanthines in yerba mate (*Ilex paraguariensis*) [J]. Journal of Food Composition & Analysis, 2014, 35(2): 21-23.
- [19] Merluzzi P, Mercurio A, Haines C P, *et al.* ACCESS: NIR Luminosity Function and Stellar Mass Function of Galaxies in the Shapley Supercluster Environment[J]. Monthly Notices of the Royal Astronomical Society, 2009, 402(2): 753-766.
- [20] Palou A, Cruz J, Blanco M, *et al.* Determination of drug, excipients and coating distribution in pharmaceutical tablets using NIR-CI[J]. Journal of Pharmaceutical Analysis, 2012, 2(2): 90-97.
- [21] Ren Y, Wang R, Liu Y, *et al.* A hematoporphyrin-based delivery system for drug resistance reversal and tumor ablation[J]. Biomaterials, 2013, 35(8): 2462–2470.
- [22] Sandak A, Sandak J, Babiński L, *et al.* Spectral analysis of changes to pine and oak wood natural polymers after short-term waterlogging[J]. Polymer Degradation & Stability, 2013, 99(13): 68-79.
- [23] Schrader B, Dippel B, Erb I, *et al.* NIR Raman spectroscopy in medicine and biology: results and aspects[J]. J Mol Struct, 1999, s 480–481: 21-32.
- [24] Udelhoven T, Emmerling C, Jarmer T. Quantitative analysis of soil chemical properties with diffuse reflectance spectrometry and partial least-square regression: A feasibility study[J]. BMC Evolutionary Biology, 2013, 13(3): 867-907.
- [25] Yi W, Xiang M A, Dong W Y, *et al.* Tobacco Quality Analysis of Producing Areas of Yunnan Tobacco Using Near-Infrared(NIR) Spectrum[J]. Spectroscopy & Spectral Analysis, 2013, 33(1): 78-80.
- [26] You J, Zhang P, Hu F, *et al.* Near-infrared light-sensitive liposomes for the enhanced photothermal tumor treatment by the combination with chemotherapy[J]. Pharm Res, 2014, 31(3): 554-65.
- [27] Yuan L, Lin W, Chen H. Analogs of Changsha near-infrared dyes with large Stokes Shifts for bioimaging[J]. Biomaterials, 2013, 34(37): 9566-9571.
- [28] Zhang L D, Zhao L L, Zhao L L, *et al.* Study on the Application for Near-Infrared Spectroscopy Quantitative Analysis and Selecting Optimum Wavelength by the MAXR

- Regression Procedure[J]. Spectroscopy & Spectral Analysis, 2005, 25(8): 1227-1229.
- [29] 梁高峰, 贾宏汝, 谷运红等. 近红外光谱分析技术及其在农业研究中的应用[J]. 安徽农业科学, 2007, 35(29): 9113-9115.
- [30] 梁亮, 刘志霄, 杨敏华等. 近红外光谱技术在生物学研究中的应用[J]. 生命科学研究, 2007, 12(9): 11-13.
- [31] 刘浩, 相秉仁, 屈凌波等. 芳酸类药物的近红外及二维相关(2DCorr)光谱分析[J]. 光谱学与光谱分析, 2007, 27(2): 265-269.
- [32] 施丰成, 邓发达, 朱立军等. NIR 光谱法测定烟草中棕榈酸、亚麻酸及亚油酸[J]. 烟草科技, 2010, (1): 32-36.
- [33] 孙露萍, 王举涛. 近红外在食品及药品农残检测中的应用研究进展[J]. 广州化工, 2013, 41(15): 12-13.
- [34] 王艳斌, 郭庆洲, 陆婉珍等. 近红外分析方法测定润滑油基础油的化学族组成[J]. 石油化工, 2001, 30(3): 224-227.
- [35] 闻明, 吉海彦. 小波分析消噪及其在 Matlab 中的实现[J]. 现代电子技术, 2003, 26(24): 47-49.
- [36] 吴静珠, 石瑞杰, 陈岩, *et al.* 食用油油酸的近红外特征谱区优选[J]. 中国粮油学报, 2015, 30(2): 118-121.
- [37] 殷飞, 李珂. 近红外光谱药品快速鉴别技术在基层药品检验中的应用[J]. 药物分析杂志, 2010, (10): 1996-2000.
- [38] 袁天军, 王家俊, 者为等. 近红外光谱法的应用及相关标准综述[J]. 中国农学通报, 2013, 中国农学通报, 2013, 29(7): 190-196.
- [39] 赵妮娜. 基于 NIR/FTIR 的烟草淀粉和总氮测定方法的研究[J]. 西北农林科技大学, 2013, (101): 99-101.
- [40] 罗华飞. MATLAB GUI 学习手记(第2版)[M]. 北京:北京航空航天大学出版社, 2011.

第三章 基于便携式近红外光谱仪的重金属离子定量分析研究

第一节 引言

近年来，重金属污染成为一个受到广泛关注的环境问题。砷、镍、铜、汞、铅、锌、铬、镉等重金属在工业和农业上被普遍使用^[1-2]，其不仅不能生物降解，反而由于在水中具有很高溶解度，很容易被生物体吸收并富集，当这些重金属离子在人体中沉积过量时便会导致疾病的发生。因此，对水环境等复杂体系中重金属离子含量的监测具有重要意义。利用原子光谱、质谱、电分析化学等分析技术可以实现对重金属离子的高灵敏检测^[3-6]。但是，这些技术往往需要繁杂的样品处理步骤，且得到组成相对简单的样品后才能进行检测^[7]。因此，发展样品前处理简单且快速的分析方法是十分必要的。

近红外光谱（NIRS）具有透射、反射、透反射等多种灵活的测量模式，几乎不需要样品处理，可直接对实际样品进行快速、无损和在线分析^[8]。近年来，近红外光谱分析技术已被广泛应用于食品^[9]、烟草^[10-11]、石化^[12-13]和医药^[14-15]等领域。近红外光谱来源于化合物中含氢基团基频振动的倍频和合频吸收，而重金属离子在近红外光谱区并没有吸收^[16-17]，因此很难利用近红外光谱响应直接对其进行定量分析。但有些重金属可以与背景基质中有机官能团络合实现富集，从而产生近红外光谱响应，使得近红外光谱分析技术能够间接对微量的重金属离子进行定量分析^[18]。Du 等结合近红外光谱技术在线检测水溶液中微量的铜离子^[19]和铅离子^[20]，Shao 等使用分散吸附结合近红外光谱技术检测了水溶液中微量的汞、铅、镉离子^[21]以及铜、钴、镍离子^[22]。采用富集技术不仅降低了近红外光谱的检测限，还实现了不具有近红外光谱活性的重金属离子的间接分析。

便携式近红外仪器结构紧凑、携带方便，适合于现场检测，然而由于其分辨率较低等因素，使得它在对目标物的定量分析方面受到一定的限制。为了探索其定量分析的可行性，本研究使用便携式近红外光谱仪结合富集技术定量检测水中的重金属离子。实验结果表明，结合化学计量学方法和富集技术，便携

式近红外光谱仪器也能够快速定量分析水中的重金属离子，这在很大程度上提高了检测效率。

本章旨在探讨便携式近红外仪器结合富集技术定量检测水中重金属离子的可行性。在研究中，首先采用胺基改性聚合淀粉对水中的重金属离子进行吸附，然后使用便携式近红外仪器直接采集吸附目标物的 AMS 的近红外漫反射光谱，结合偏最小二乘回归多方法对其进行建模分析。建模结果表明便携式近红外光谱仪器用于水中重金属离子的定量检测是可行的。

第二节 样品准备与光谱采集

3.2.1 样品准备

本研究采用制备好的现有样品，该批样品由本课题组张洋^[23]制备，采用粒度一致的胺基改性聚合淀粉（由广州大学化学化工学院提供）吸附 50 个含有不同浓度镍离子、铜离子、锌离子、镉离子、铅离子、铬离子的混合溶液（该混合溶液由硝酸镍（ $\text{Ni}(\text{NO}_3)_2 \cdot 6\text{H}_2\text{O}$ ，Purity=98.0%），硝酸铜（ $\text{Cu}(\text{NO}_3)_2 \cdot 3\text{H}_2\text{O}$ ，Purity $\geq 99.5\%$ ），硝酸锌（ $\text{Zn}(\text{NO}_3)_2 \cdot 6\text{H}_2\text{O}$ ，Purity=98.0%），硝酸镉（ $\text{Cd}(\text{NO}_3)_2 \cdot 4\text{H}_2\text{O}$ ，Purity $\geq 98.5\%$ ），硝酸铅（ $\text{Pb}(\text{NO}_3)_2$ ，Purity=99.0%），硝酸铬（ $\text{Cr}(\text{NO}_3)_3 \cdot 9\text{H}_2\text{O}$ ，Purity=98.5%）提供重金属离子，硝酸铜从天津市光复科技发展有限公司处购买，其余来自阿法埃沙天津化学有限公司）然后将吸附剂与残余液体分离，使胺基改性聚合淀粉自然风干，去除残余水分，得到 50 个样品。将样品编号，依次装入对应编号的透明无色平底干燥的玻璃瓶中，装入的样品量能够完全覆盖玻璃瓶底部。同时，将各样品标号与其所含有的金属离子浓度的对应信息录入到电子表格中。

3.2.2 近红外光谱的采集

室温下，使用 MicroNIRTM 1700 光谱仪采集所有样品的近红外漫反射光谱。首先将光谱仪固定在支架上，光源方向朝上，扫描样品时，将装有样品的无色透明玻璃置于光源中心，样品完全覆盖玻璃瓶底部以确保光源与样品充分接触。

每条光谱由 125 个波长点组成，波长范围在 900~1700nm ($11111\sim5882\text{cm}^{-1}$)，由于波长点数非常少，因此本文选择全波段进行分析。与此同时，为了降低噪声，提高光谱的信噪比，每条光谱均扫描 50 次，为保证准确性每个样品重新装样扫描三次并取其平均光谱。

第三节 结果与讨论

3.3.1 光谱预处理

图 3.1 为使用 MicroNIRTM 1700 光谱仪扫描样品得到的近红外漫反射光谱图。

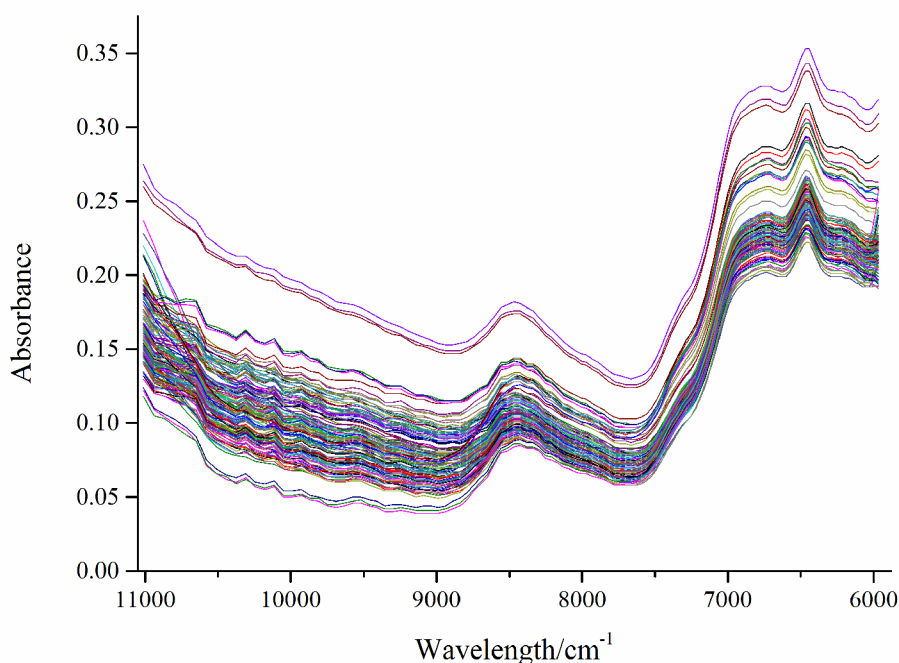


图 3.1 样品的近红外光谱

从图 3.1 可以看出原始光谱中存在明显的基线漂移和噪声，并且在 $11111\sim9000\text{cm}^{-1}$ 之间存在干涉现象，因此在建立模型之前需要对原始光谱进行预处理。常用的光谱预处理方法有连续小波变换 (Continuous Wavelet Transform, CWT)。

Savitzky-Golay 求导法 (SG 求导)、标准正态变量变换 (Standard Normal Variate, SNV) 和多元散射校正 (Multiplicative Scatter Correction, MSC) 等。连续小波变换利用小波系数进行回归分析, 通过合适的尺度参数和小波基有效地去除信号中的无效信息, 并保留有效信息, 常用于平滑、滤噪、背景扣除和基线校正; SG 求导用于求取导数光谱, 可有效消除基线和其他背景的干扰, 分辨重叠峰, 提高分辨率和灵敏度; SNV 和 MSC 主要用来消除样品颗粒大小、表面散射以及光程变化对近红外漫反射光谱的影响(本研究中所使用到的 CWT 和 SG 求导均为二阶导数)。以 SG 求导为例说明光谱预处理的作用。图 3.2 为经过 Savitzky-Golay 求导之后的光谱, 可以看出原始光谱中存在的基线漂移被消除, 并且光谱具有相对较高的分辨率。

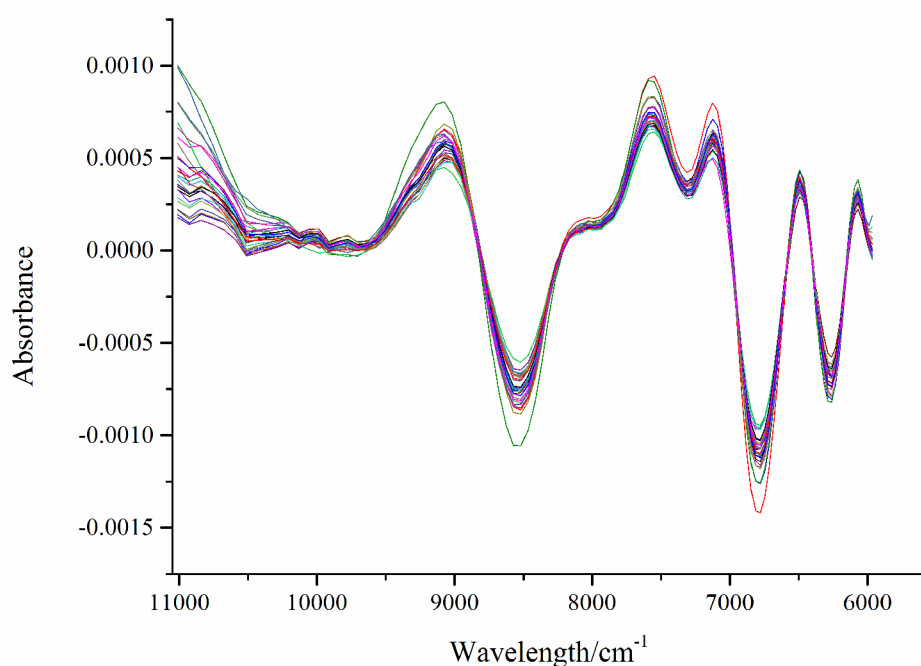


图 3.2 卷积求导之后的光谱

3.3.2 校正模型的建立

对样品光谱进行主成分分析, 选用主成分得分为特征变量进行 Kennard-Stone 分组, 得到由 30 个样品组成的校正集和 20 个样品组成的验证集,

然后在 $11111 \sim 9000 \text{ cm}^{-1}$ 的光谱范围内,采用偏最小二乘方法建立定量模型。在建立模型之前,先进行偏最小二乘模型因子数的选择。合理确定参加建立模型的主因子数是充分利用光谱信息和滤除噪声的有效方法,如果主因子数太小,将会丢失原始光谱较多的有用信息,导致拟合不充分,称为欠拟合;太大则会将测量噪声过多地包括进来,所建模型的预测误差会显著增大,称为过拟合。以图 3.1 中经过 Savitzky-Golay 求导之后的光谱为例,对该光谱进行留一交叉验证,得到该光谱的交叉验证均方根误差随主因子数变化的趋势图(图 3.3)。从图中可以看出,交叉验证均方根误差随着主因子数的增加,呈递减趋势,在主因子数为 6 之后趋于稳定,并有小幅上升,说明在这点以后,增加的主成分是与被测组分无关的噪声成分。因此使用 Savitzky-Golay 求导处理样品光谱时,建立定量模型所用的最佳主因子数可以取为 6、7、8。

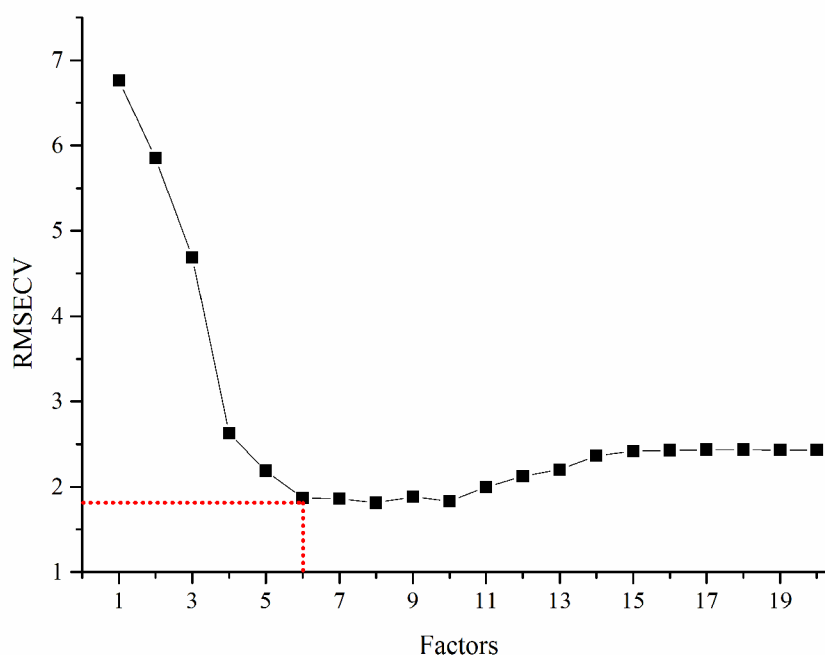


图 3.3 交叉验证均方根误差随主因子数变化的趋势图

分别对镍离子和铜离子建立定量模型,为筛选出最佳模型,将经过各种预处理方法后所建立的定量模型的相关系数(R)、交叉验证均方根误差(RMSECV)以及 RPD 值列于表 3.1 中进行比较。样品参考值与预测浓度值的相关系数以及

交叉验证均方根误差用于评价校正模型，通常模型的 RPD 值大于 3.0 说明模型可以用于筛查和过程控制，而大于 5.0 则说明模型可以用于准确的定量分析^[24]。由表可知，经过预处理之后所建立的定量模型均得到不同程度的优化。对于镍离子，使用连续小波变换时，模型的相关系数明显优于其他预处理方法，定量模型的交叉验证均方根误差值相对较低，RPD 值为 5.25，这说明该定量模型适用于对镍离子的准确定量分析；对于铜离子，各种预处理方法下的定量模型的相关系数均在 0.96 左右，但多元散射校正结合 Savitzky-Golay 求导的预处理方法对应一个较低的均方根误差值和较高的 RPD 值，所以采用此方法对铜离子进行定量建模。

表 3.1 镍离子和铜离子建立 SIMPLS 模型的交叉验证结果

	Preprocessing	Factor	R	RMSECV	RPD		Preprocessing	Factor	R	RMSECV	RPD
Ni ²⁺	None	10	0.9391	2.90	2.86		None	9	0.9473	2.67	3.07
	SNV	3	0.9546	2.50	3.40		SNV	8	0.9523	2.37	3.33
	MSC	3	0.9545	2.50	3.40		MSC	7	0.9620	2.12	3.72
	CWT	9	0.9819	1.62	5.25		CWT	9	0.9611	2.20	3.60
	SG	8	0.9766	1.81	4.69	Cu ²⁺	SG	7	0.9564	2.27	3.48
	MSC-CWT	3	0.9559	2.45	3.46		MSC-CWT	9	0.9641	2.08	3.80
	MSC-SG	5	0.9529	2.55	3.34		MSC-SG	6	0.9654	2.03	3.89
	SNV-CWT	3	0.9569	2.43	3.51		SNV-CWT	9	0.9616	2.14	3.70
	SNV-SG	5	0.9529	2.55	3.34		SNV-SG	6	0.9636	2.09	3.79

3.3.3 模型的预测效果

为了验证模型的预测效果，使用 20 个外部验证集样品对所建立的定量模型进行检验。在验证实验中，吸附操作、光谱测量条件、建模方法与校正集样品相同。20 个样品中镍离子和铜离子的浓度范围分别为 3.30 ~ 29.13 mg · L⁻¹ 和 3.29 ~ 29.60 mg · L⁻¹。图 3.4 为验证集样品中镍离子和铜离子的参考浓度和预测浓度之间的关系图，分别线性回归，得到镍离子的拟合相关系数为 0.9665，而

铜离子的拟合相关系数为 0.9544。并且通过计算得到镍离子的预测均方根误差为 1.8528，铜离子的预测均方根误差为 2.2316。以上结果表明，建立的镍离子和铜离子的校正集模型对验证集样品的预测效果较好，模型可以用于定量分析。

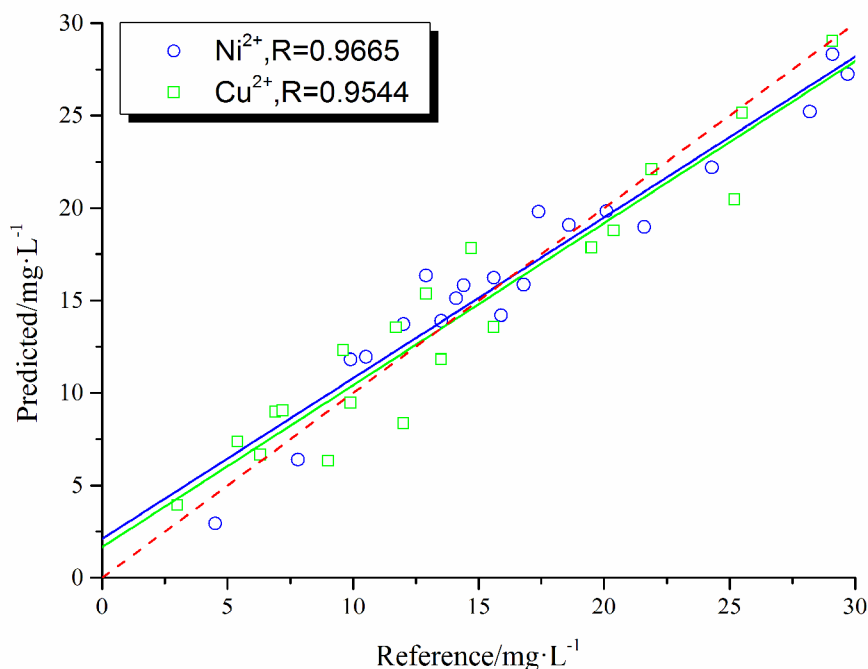


图 3.4 镍离子和铜离子验证集样品组分浓度参考值和近红外光谱预测值的关系图,直线通过线性拟合得到,虚线表示对角线

第四节 结论

为研究便携式近红外光谱仪对水中微量重金属离子定量分析的可行性，本文首先通过胺基改性聚合淀粉富集稀溶液中的镍离子和铜离子，然后采用样品的近红外漫反射光谱结合偏最小二乘回归法建立定量模型，之后使用外部验证的方法对回归模型进行准确地验证。镍离子和铜离子的定量检测范围分别为 $3.30 \sim 29.13 \text{ mg} \cdot \text{L}^{-1}$ 和 $3.29 \sim 29.60 \text{ mg} \cdot \text{L}^{-1}$ ，样品的预测值与参考值相关系数分别为 0.9894 和 0.9524。结合化学计量学方法和富集技术，便携式近红外光谱仪同样能够对重金属离子定量分析。同时本方法不但提高了近红外光谱技术的检

测灵敏度，而且是一种使用便携式近红外光谱仪对微量金属离子定量分析的有益探。

参考文献

- [1] Chen G, Yuan M, Wei T, *et al.* Erratum to “Micro near infrared spectroscopy (MicroNIRS) based on on-line enrichment: Determination of trace copper in water using glycidyl methacrylate-based monolithic material” [Anal. Chim. Acta 670 (2010) 39–43][J]. Anal Chim Acta, 2010, 670(1-2): 380-387.
- [2] Chen Y P, Liu Q, Yue X Z, *et al.* Ultrasonic vibration seeds showed improved resistance to cadmium and lead in wheat seedling[J]. Environmental Science & Pollution Research International, 2013, 20(7): 4807-16.
- [3] Ding X, Guo Y, Ni Y, *et al.* A novel NIR spectroscopic method for rapid analyses of lycopene, total acid, sugar, phenols and antioxidant activity in dehydrated tomato samples[J]. Vib Spectrosc, 2015, 82: 1-9.
- [4] Giner - Casares J J, Henriksen - Lacey M, Garc, *et al.* Plasmonic Surfaces for Cell Growth and Retrieval Triggered by Near-Infrared Light[J]. Angew Chem, 2015, 9(3): 12-15.
- [5] González-Martín I, lvarez-García N, González-Pérez C, *et al.* Determination of inorganic elements in animal feeds by NIRS technology and a fibre-optic probe[J]. Talanta, 2006, 69(3): 711-715.
- [6] Guo M, Law W C, Liu X, *et al.* Plasmonic Semiconductor Nanocrystals as Chemical Sensors: Pb²⁺ Quantitation via Aggregation-Induced Plasmon Resonance Shift[J]. Plasmonics, 2014, 9(4): 893-898.
- [7] Huang Z, Tao W, Fang J, *et al.* Multivariate calibration of on-line enrichment near-infrared (NIR) spectra and determination of trace lead in water[J]. Chemometrics & Intelligent Laboratory Systems, 2009, 98(2): 195-200.
- [8] Kol, Aacute. Quick method (FT&2010;NIR) for the determination of oil and major fatty acids content in whole achenes of milk thistle (Silybum marianum (L.) Gaertn.)[J]. J Sci Food Agric, 2014,6(2): 7-9.
- [9] Lau W K, Yeung C W, Lui P W, *et al.* Different trends in modulation of NMDAR1 and NMDAR2B gene expression in cultured cortical and hippocampal neurons after lead exposure[J]. Brain Research, 2002, 932(1-2): 10-24.
- [10] Li J, Zhang Y, Cai W, *et al.* Simultaneous determination of mercury, lead and cadmium ions in water using near-infrared spectroscopy with preconcentration by thiol-functionalized magnesium phyllosilicate clay[J]. Talanta, 2011, 84(3): 679-83.
- [11] Liu X, Chen H C, Liu T A, *et al.* Application of PCA-SVR to NIR Prediction Model for Tobacco Chemical Composition[J]. Spectroscopy and Spectral Analysis, 2007, 27(12): 2-3.
- [12] Liu Y, Ning Y, Cai W, *et al.* Micro-analysis by near-infrared diffuse reflectance spectroscopy

- with chemometric methods[J]. *Analyst*, 2013, 138(21): 6617-6622.
- [13] Pó A, #X, Torak, *et al.* Correlation between instrumental texture and colour quality attributes with sensory analysis of selected cheeses as affected by fat contents[J]. *International Journal of Food Science & Technology*, 2015, 50(4): 999–1008.
- [14] Sheng N, Cai W, Shao X. An approach by using near-infrared diffuse reflectance spectroscopy and resin adsorption for the determination of copper, cobalt and nickel ions in dilute solution[J]. *Talanta*, 2009, 79(2): 339-343.
- [15] Shenk J S, Westerhaus M O, Hoover M R, *et al.* Analysis of Forages by Infrared Reflectance1[J]. *J Dairy Sci*, 1979, 62(5): 807-812.
- [16] Tseng J Y, Chang C Y, Chang C F, *et al.* Kinetics and equilibrium of desorption removal of copper from magnetic polymer adsorbent[J]. *J Hazard Mater*, 2009, 171(1-3): 370-373.
- [17] Xu Z, Yan L, Li X, *et al.* Discriminant analysis of Chinese patent medicines based on near-infrared spectroscopy and principal component discriminant transformation[J]. *Spectrochimica Acta Part A Molecular & Biomolecular Spectroscopy*, 2015, 149: 985-990.
- [18] Yi W, Xiang M A, Dong W Y, *et al.* Tobacco Quality Analysis of Producing Areas of Yunnan Tobacco Using Near-Infrared(NIR) Spectrum[J]. *Spectroscopy & Spectral Analysis*, 2013, 33(1): 78-80.
- [19] Zhang C, Zhang M, Chang Q. Preparation of mercaptoacetyl chitosan and its removal performance of copper ion and turbidity[J]. *Desalination & Water Treatment*, 2013, 53(7): 1-8.
- [20] Zou D, Feng W, Shi G, *et al.* Hetero-binuclear near-infrared (NIR) luminescent Zn–Nd complexes self-assembled from the benzimidazole-based ligands[J]. *Spectrochimica Acta Part A-Molecular and Biomolecular Spectroscopy*, 2012, 22(98): 126-130.
- [21] 林梅莹, 尚小琴, 李淑妍等. 氨基改性淀粉重金属废水处理剂的制备及应用[J]. *化工进展*, 2011, 30(4): 854-856.
- [22] 邵学广, 宁宇, 刘凤霞等. 近红外光谱在无机微量成分分析中的应用[J]. *化学学报*, 2012, 70(20): 2109-2114.
- [23] 于新洋, 卢启鹏, 高洪智等. 便携式近红外光谱仪器现状及展望[J]. *光谱学与光谱分析*, 2013, (11): 2983-2988.
- [24] 张洋. 吸附材料的改性及其在微量成分近红外光谱分析中的应用研究[D]. 南开大学, 2013.

第四章 基于便携式近红外光谱仪的判别分析研究

第一节 引言

近年来,便携式近红外光谱仪的研制与应用成为热点^[1]。美国 Axsun 科技公司研制的 Axsun-XL410 型便携式近红外光谱仪采用法布里-珀罗干涉仪作为单色器^[2-6],光学系统采用微机电系统(Micro Electro-Mechanical Systems, MEMS)制造,光源、单色器和检测器等核心部件的整体长度仅 14mm,并且包括波长和光强(Wavelength and Amplitude)两个校正模块,整体尺寸与字典相当^[7-10],重量约 3 公斤,具有很好的便携型、良好的抗震性、稳定性以及环境适应性^[11-14];美国 JDSU 公司成功研发出体型微小的近红外光谱仪——MicroNIRTM 1700 光谱仪^[15,16]。JDSU MicroNIRTM 是一款超紧凑、超轻便的近红外光谱仪,其核心分光元件是 JDSU 设计和制造的线性渐变滤光片(LVF, Linear Variable Filter)。LVF 是一种特殊的带通滤光片,使用了 JDSU 先进的光学镀膜和制造技术——制作时特意向特定方向形成楔形镀层。由于通带的中心波长与膜层厚度相关,因此滤光片的穿透波长在楔形方向发生了线性变化,从而起到分光作用。将 LVF 与线阵 InGaAs 探测器(128 线元非制冷铟镓砷二极管阵列)耦合,组成了可以提供光谱信息的近红外传感器,可实时评估,适合现场检测,易于使用,可用于无损检测、快速测试、在线过程监控^[17,18]。

便携型近红外光谱仪不仅要求便携性,而且要求仪器具有实用性。因此,对便携型近红外光谱仪所采集的光谱建立的校正模型的性能进行评价变得非常重要。在用化学计量学方法建立模型时,各种各样的预处理方法经常被用到,常用的预处理方法有连续小波变换、Savitzky-Golay 卷积求导法、多元散射校正、标准正态变量等。同时,建立判别模型所用到的距离判据也有多种(如欧氏距离,马氏距离以及 Hotelling T² 等)。这些不同的预处理方法以及不同的距离判据直接影响判别模型的判别效果。在不同的预处理方法和距离判据中取得一个最佳的组合搭配将是本章重点解决的问题。

本章将首先使用 JDSU MicroNIRTM 仪器扫描出实验样品的谱图,然后根据测试结果,结合化学计量学方法比较不同的光谱预处理方法和距离判据条件下

建立判别模型的预测效果，根据预测效果提出每一种样品下最佳的预处理方法和距离判据。

第二节 样品准备与光谱采集

4.2.1 实验样品及测试条件

由 A、B、C、D、E 公司生产的注射用维库溴铵样品（该批样品由中国食品药品检定研究院提供），物理状态为白色粉末或块状物，规格为 4 mg/瓶。

室温条件下测量，将样品玻璃瓶垂直向上放在 JDSU 检测器上，隔玻璃瓶底测试，扫描次数设定为 50 次，每个样品采集三条光谱（降低噪声，提高信噪比）。镀金的漫反射体作为参考暗电流和背景。仪器在 908.1nm ~ 1676.2nm 范围内取样，波长点间隔为 6.2nm，每条光谱由 125 个点构成。

4.2.2 近红外光谱的采集

实验中采用 JDSU MicroNIR™ 1700 对样品进行扫描。金镜（镀金的漫反射体）作为光谱扫描的暗电流（Dark Current）和参考背景。为了增加光谱信号的信噪比（Signal / Noise），设置样品扫描次数为 50 次，每个样品扫描三条光谱，取这三条光谱的平均光谱用于数据处理。扫描时，先用高级擦镜纸将盛放样品的无色玻璃瓶底擦干净，然后将玻璃瓶口朝上放置于仪器检测器上，每个样品三次扫描中，每次完成后都重新放置玻璃瓶，以确保所得到的光谱能够尽可能全面反映样品的信息。常用的光谱预处理方法有连续小波变换（CWT）、Savitzky-Golay 卷积求导法、标准正态变量变换（SNV）和多元散射校正（MSC）。连续小波变换是本课题组经常使用的近红外光谱预处理方法。小波变换将光谱信号分解成一系列小波函数的叠加，小波分析在时域和频域同时具有良好的局部化性质，它可以对高频成分采用逐渐精细的时域或空间域取代步长，从而可以聚焦到对象的任意细节，故而小波变化被誉为分析信号的“数学显微镜”，在分析化学的信号处理中有着较为广泛的应用；对光谱求取导数是光谱分析中常用的基线校正和光谱分辨预处理方法。对光谱求导一般有两种方法：直接差分法和 Savitzky-Golay 卷积求导法。对于分辨率高、波长采样点多的光谱，直接差

分法求取的导数光谱与实际相差不大，但对于稀疏波长采样点的光谱，该方法所求的导数则存在较大误差。考虑到本实验中所用到的 JDSU MicroNIRTM 1700 扫描出来的光谱只有 125 个波长点，属于标准的“稀疏波长”，故而这里采用 Savitzky-Golay 卷积求导法计算。标准正态变量变换主要用来消除样品颗粒大小、表面散射以及光程变化对近红外漫反射光谱的影响。多元散射校正的目的与标准正态变量变换相同，主要消除颗粒分布不均匀及颗粒大小产生的散射影响，故而将这两种预处理方法放在一起进行讨论。

第三节 结果与讨论

4.3.1 光谱预处理

图 4.1 为 A 公司生产的维库溴铵样品的原始谱图，从图中可以明显看出光谱中存在基线背景漂移以及大量的重叠峰。为使所建立的判别模型更加稳健、可靠，需要在建立模型之前首先运用光谱预处理方法对光谱进行处理。而不同的预处理方法也对最终模型产生影响，因此，有必要比较多种预处理方法下所建模型的预测效果。

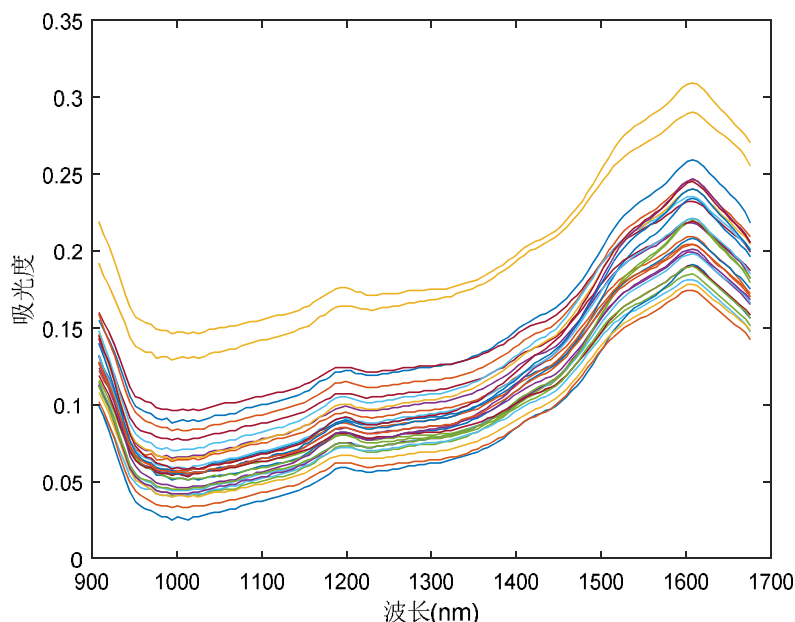


图 4.1 A 公司生产的维库溴铵样品的原始光谱

图 4.2 为经过连续小波变换预处理后样品的光谱,可以看出基线以及重叠峰得到了明显校正,处理后的光谱具有较高的分辨率。

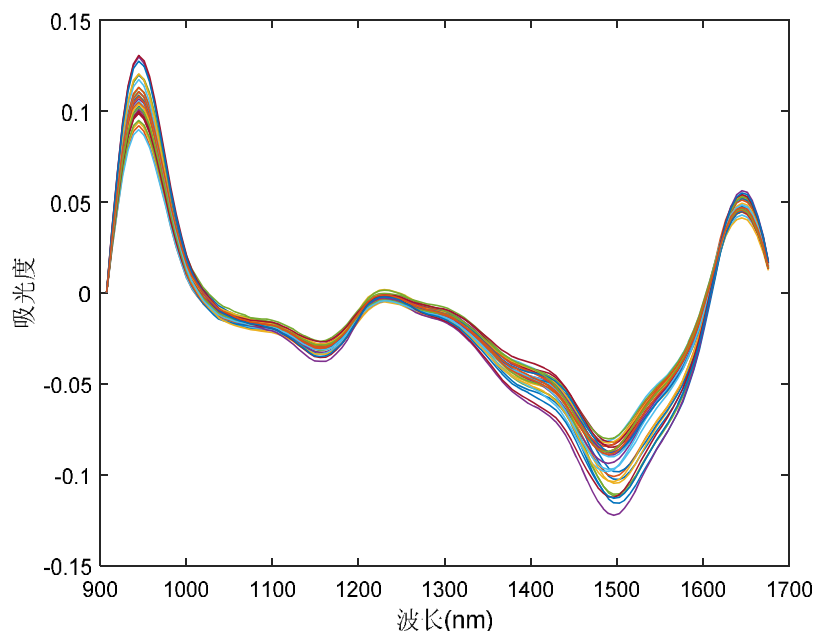


图 4.2 CWT 处理后的光谱

4.3.2 判别模型建立及验证

待分析样品共有五组,分别以这五种样品为校正集建立五个判别模型。具体到每一组样品(例如由 A 公司所生产),从该组中随机选取一半样品作为校正集(model),另外一半样品自动划分为预测集(test)。以校正集 model 对应的光谱为校正集,分别以欧氏距离、明氏距离、街区距离、PC 空间欧氏距离、PC 空间明氏距离、PC 空间街区距离、PC 空间马氏距离以及 Hotelling T^2 为距离判据建立判别模型。预测集 test 对应的光谱作为检验集,用于模型的内部验证;其余四组样品作为检验集,用于模型的外部验证。根据模型的内部和外部验证综合结果来评价便携式仪器 JDSU MicroNIRTM 1700 对于本组实验样品的判别能力。

光谱预处理完成之后,分别以欧氏距离、明氏距离、街区距离、PC 空间欧氏距离、PC 空间明氏距离、PC 空间街区距离、PC 空间马氏距离以及 Hotelling T^2 为距离判据建立判别模型(置信区间取常用的 0.95),并进行内部和外部验证。以明氏距离为距离判据建立模型,结果如图 4.3(a)、(b)所示。

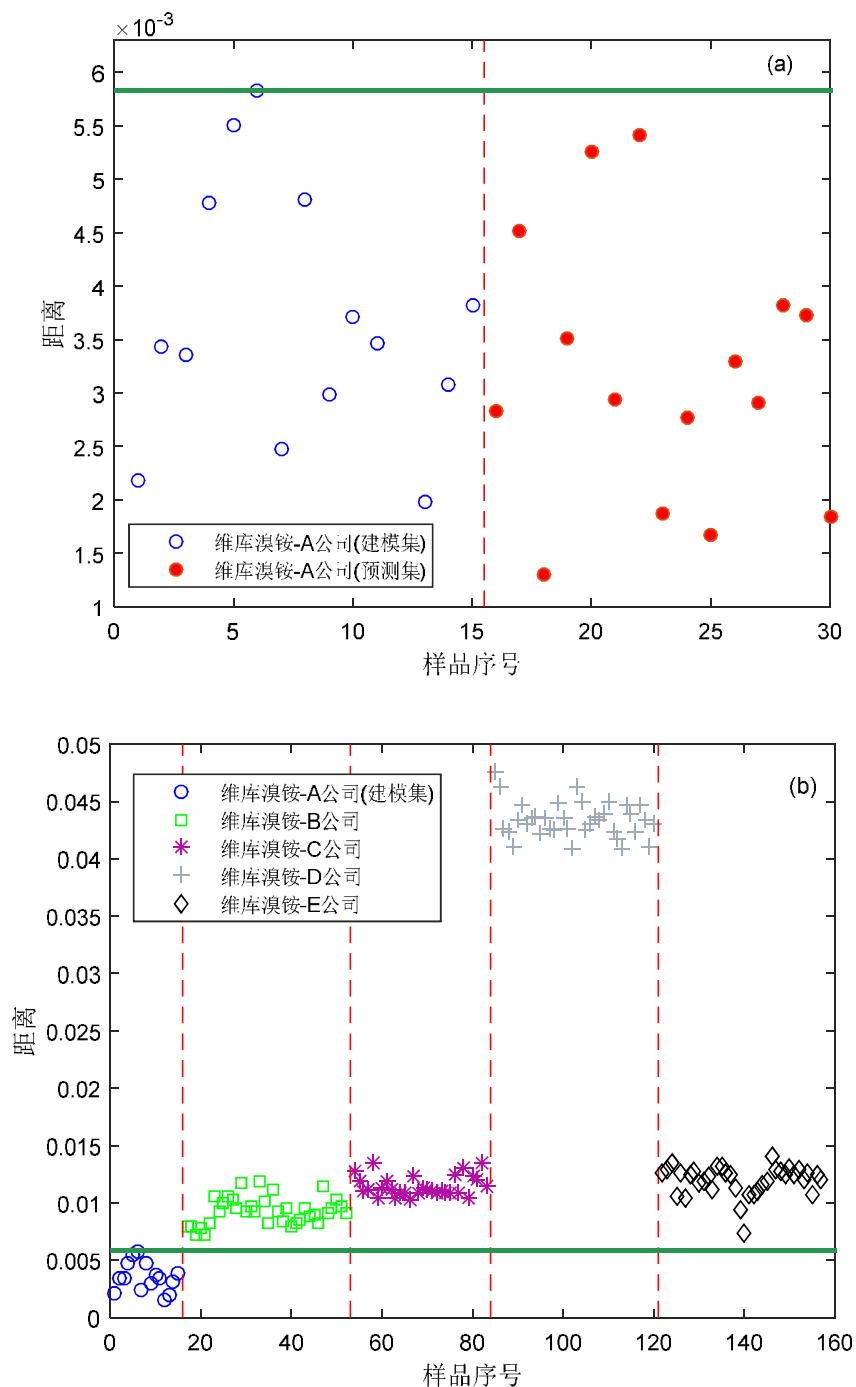


图 4.3 距离判据为明氏距离下的预测结果

(a) 为类内验证结果 (b) 为类外验证结果

图 (a) 和 (b) 中横坐标均表示样品的序号, 纵坐标为每个样品对应的距离数值, 蓝色空心圆点为用来建立模型的样品在对应距离下的位置, 绿色加粗横向直线表示相应距离下的阈值, 它是进行判别分析的依据, 所有在该直线以下

的样品点都被判别为属于建模集样品类(这里为 A 公司生产的维库溴铵),在该直线以上则不属于该类;(a)中实心圆点表示用于类内验证的样品;(b)中绿色空心方块代表 B 公司生产的维库溴铵样品,紫色星号代表 C 公司,浅灰色十字表示 D 公司,黑色空心菱形为 E 公司生产维库溴铵样品对应的距离值。

从图(a)中可以看出,所有的类内验证集样品均落在了阈值线以下,说明模型全部成功识别出了同建模集样品属于一类的类内验证集样品,图(b)中所有非 A 公司生产的维库溴铵都出现在阈值线之上,说明模型将这些样品划分到建模集之外,与事实相符,说明结合相应的光谱预处理方法,明氏距离模型在类内和类外预测均得到了正确的结果,没有出现误判现象,效果比较好。通过表 4-1 可以看到在相同的预处理方法下其他距离判据模型的判别效果。

表 4-1 八种距离下的预测结果统计

距离类型	类内验证(样品数:15)		类外验证(样品数:138)	
	准确数	误判数	准确数	误判数
欧氏距离	15	0	138	0
街区距离	15	0	138	0
明氏距离	15	0	138	0
PC 空间欧氏距离	15	0	138	0
PC 空间街区距离	15	0	138	0
PC 空间明氏距离	15	0	138	0
PC 空间马氏距离	10	5	138	0
Hotelling T^2	13	2	138	0

4.3.3 模型的预测效果

五种样品在 CWT、SG、SNV 等不同的预处理方法以及不同的距离判据下所建模型的预测结果统计见表 4-2。

表 4-2 五种样品的预测结果统计

样品(维库 溴铵)	预处理方法	距离类型	类内验证		类外验证	
			准确数	误判数	准确数	误判数
A 公司(共 15 个)	不进行预处理	欧氏距离	15	0	47	91
		明氏距离	15	0	43	95
		街区距离	15	0	47	91
		PC 空间欧氏距离	15	0	47	91

第四章 基于便携式近红外光谱仪的判别分析研究

		PC 空间明氏距离	15	0	56	82
		PC 空间街区距离	15	0	47	91
		PC 空间马氏距离	13	2	138	0
		Hotelling T2	15	0	138	0
		欧氏距离	15	0	98	40
		明氏距离	15	0	83	55
		街区距离	15	0	98	40
	连续小波变换 (CWT)	PC 空间欧氏距离	15	0	97	41
		PC 空间明氏距离	15	0	136	2
		PC 空间街区距离	15	0	97	41
		PC 空间马氏距离	8	7	138	0
		Hotelling T2	12	3	138	0
		欧氏距离	15	0	137	1
		明氏距离	15	0	131	7
		街区距离	15	0	137	1
		PC 空间欧氏距离	15	0	130	8
		PC 空间明氏距离	15	0	138	0
		PC 空间街区距离	15	0	130	8
		PC 空间马氏距离	11	4	138	0
	Savitzky-Golay 卷积求导	Hotelling T2	14	1	138	0
		欧氏距离	15	0	78	60
		明氏距离	15	0	44	94
		街区距离	15	0	78	60
		PC 空间欧氏距离	15	0	66	72
		PC 空间明氏距离	15	0	121	17
		PC 空间街区距离	15	0	66	72
		PC 空间马氏距离	15	0	138	0
		Hotelling T2	15	0	138	0
		欧氏距离	15	0	136	2
		明氏距离	15	0	130	8
		街区距离	15	0	136	2
	标准正态变量变 换 (SNV) 结合连 续小波变换 (CWT)	PC 空间欧氏距离	15	0	135	3
		PC 空间明氏距离	15	0	138	0
		PC 空间街区距离	15	0	135	3
		PC 空间马氏距离	8	7	138	0
		Hotelling T2	11	4	138	0
		欧氏距离	15	0	138	0
		明氏距离	15	0	138	0
		街区距离	15	0	138	0
		PC 空间欧氏距离	15	0	138	0
	标准正态变量变 换 (SNV) 结合 Savitzky-Golay 卷积求导	PC 空间明氏距离	15	0	138	0
		PC 空间街区距离	15	0	138	0
		PC 空间马氏距离	15	0	138	0
		PC 空间欧氏距离	15	0	138	0

第四章 基于便携式近红外光谱仪的判别分析研究

B 公司(共 18 个)		PC 空间明氏距离	15	0	138	0
		PC 空间街区距离	15	0	138	0
		PC 空间马氏距离	10	5	138	0
		Hotelling T2	13	2	138	0
		欧氏距离	18	0	0	132
		明氏距离	18	0	0	132
		街区距离	18	0	0	132
	不进行预处理	PC 空间欧氏距离	18	0	0	132
		PC 空间明氏距离	18	0	12	120
		PC 空间街区距离	18	0	0	132
		PC 空间马氏距离	11	7	132	0
		Hotelling T2	13	5	132	0
		欧氏距离	16	2	52	80
		明氏距离	16	2	47	85
		街区距离	16	2	52	80
	连续小波变换 (CWT)	PC 空间欧氏距离	16	2	52	80
		PC 空间明氏距离	17	1	83	49
		PC 空间街区距离	16	2	52	80
		PC 空间马氏距离	10	8	132	0
		Hotelling T2	14	4	132	0
		欧氏距离	17	1	70	62
		明氏距离	16	2	65	67
		街区距离	17	1	70	62
	Savitzky-Golay 卷积求导	PC 空间欧氏距离	17	1	68	64
		PC 空间明氏距离	17	1	86	46
		PC 空间街区距离	17	1	68	64
		PC 空间马氏距离	10	8	132	0
		Hotelling T2	11	7	132	0
		欧氏距离	18	0	63	69
		明氏距离	18	0	46	86
		街区距离	18	0	63	69
	标准正态变量变 换 (SNV) / 多元 散射校正 (MSC)	PC 空间欧氏距离	18	0	57	75
		PC 空间明氏距离	18	0	95	37
		PC 空间街区距离	18	0	57	75
		PC 空间马氏距离	15	3	132	0
		Hotelling T2	16	2	132	0
		欧氏距离	18	0	104	28
		明氏距离	18	0	108	24
		街区距离	18	0	104	28
	标准正态变量变 换 (SNV) 结合连 续小波变换 (CWT)	PC 空间欧氏距离	18	0	104	28

第四章 基于便携式近红外光谱仪的判别分析研究

C 公司(共 15 个)	标准正态变量变 换 (SNV) 结合 Savitzky-Golay 卷积求导	PC 空间明氏距离	16	2	126	6
		PC 空间街区距离	18	0	104	28
		PC 空间马氏距离	8	10	132	0
		Hotelling T2	14	4	132	0
		欧氏距离	18	0	116	16
		明氏距离	17	1	126	6
		街区距离	18	0	116	16
		PC 空间欧氏距离	18	0	107	25
		PC 空间明氏距离	18	0	129	3
		PC 空间街区距离	18	0	107	25
		PC 空间马氏距离	10	8	132	0
		Hotelling T2	15	3	132	0
		欧氏距离	14	1	11	127
		明氏距离	14	1	8	130
		街区距离	14	1	11	127
	不进行预处理	PC 空间欧氏距离	14	1	11	127
		PC 空间明氏距离	14	1	42	96
		PC 空间街区距离	14	1	11	127
		PC 空间马氏距离	10	5	138	0
		Hotelling T2	14	1	138	0
		欧氏距离	15	0	65	73
		明氏距离	15	0	72	66
		街区距离	15	0	65	73
		PC 空间欧氏距离	15	0	65	73
		PC 空间明氏距离	15	0	101	37
	连续小波变换 (CWT)	PC 空间街区距离	15	0	65	73
		PC 空间马氏距离	6	9	138	0
		Hotelling T2	7	8	138	0
		欧氏距离	15	0	64	74
		明氏距离	15	0	64	74
		街区距离	15	0	64	74
		PC 空间欧氏距离	15	0	62	76
		PC 空间明氏距离	15	0	98	40
		PC 空间街区距离	15	0	62	76
		PC 空间马氏距离	8	7	138	0
	Savitzky-Golay 卷积求导	Hotelling T2	11	4	138	0
		欧氏距离	15	0	98	40
		明氏距离	15	0	100	38
		街区距离	15	0	98	40
		PC 空间欧氏距离	15	0	89	49
	标准正态变量变 换 (SNV) / 多元 散射校正 (MSC)	欧氏距离	15	0	98	40
		明氏距离	15	0	100	38
		街区距离	15	0	98	40
		PC 空间欧氏距离	15	0	89	49
		欧氏距离	15	0	98	40

第四章 基于便携式近红外光谱仪的判别分析研究

D 公司(共 18 个)	标准正态变量变 换 (SNV) 结合连 续小波变换 (CWT)	PC 空间明氏距离	15	0	123	15
		PC 空间街区距离	15	0	89	49
		PC 空间马氏距离	15	0	138	0
		Hotelling T2	15	0	135	3
		欧氏距离	15	0	100	38
		明氏距离	15	0	111	27
		街区距离	15	0	100	38
		PC 空间欧氏距离	15	0	88	50
		PC 空间明氏距离	15	0	136	2
		PC 空间街区距离	15	0	88	50
		PC 空间马氏距离	5	10	138	0
		Hotelling T2	6	9	138	0
	标准正态变量变 换 (SNV) 结合 Savitzky-Golay 卷积求导	欧氏距离	15	0	77	61
		明氏距离	15	0	75	63
		街区距离	15	0	77	61
		PC 空间欧氏距离	15	0	65	73
		PC 空间明氏距离	15	0	134	4
		PC 空间街区距离	15	0	65	73
		PC 空间马氏距离	8	7	138	0
		Hotelling T2	13	2	138	0
		欧氏距离	18	0	12	120
		明氏距离	18	0	11	121
		街区距离	18	0	12	120
	不进行预处理	PC 空间欧氏距离	18	0	12	120
		PC 空间明氏距离	18	0	41	91
		PC 空间街区距离	18	0	12	120
		PC 空间马氏距离	14	4	132	0
		Hotelling T2	14	4	132	0
		欧氏距离	18	0	132	0
		明氏距离	18	0	132	0
		街区距离	18	0	132	0
	连续小波变换 (CWT)	PC 空间欧氏距离	18	0	132	0
		PC 空间明氏距离	18	0	132	0
		PC 空间街区距离	18	0	132	0
		PC 空间马氏距离	17	1	132	0
		Hotelling T2	18	0	132	0
		欧氏距离	18	0	132	0
	Savitzky-Golay 卷积求导	明氏距离	18	0	132	0
		街区距离	18	0	132	0
		PC 空间欧氏距离	18	0	132	0

第四章 基于便携式近红外光谱仪的判别分析研究

E 公司(共 18 个)	标准正态变量变 换 (SNV) /多元 散射校正 (MSC)	PC 空间明氏距离	18	0	132	0
		PC 空间街区距离	18	0	132	0
		PC 空间马氏距离	18	0	132	0
		Hotelling T2	18	0	132	0
		欧氏距离	18	0	132	0
		明氏距离	18	0	132	0
		街区距离	18	0	132	0
		PC 空间欧氏距离	18	0	132	0
		PC 空间明氏距离	18	0	132	0
		PC 空间街区距离	18	0	132	0
		PC 空间马氏距离	18	0	132	0
		Hotelling T2	18	0	132	0
	标准正态变量变 换 (SNV) 结合连 续小波变换 (CWT)	欧氏距离	18	0	132	0
		明氏距离	18	0	132	0
		街区距离	18	0	132	0
		PC 空间欧氏距离	18	0	132	0
		PC 空间明氏距离	18	0	132	0
		PC 空间街区距离	18	0	132	0
		PC 空间马氏距离	17	1	132	0
		Hotelling T2	18	0	132	0
		欧氏距离	18	0	132	0
		明氏距离	18	0	132	0
		街区距离	18	0	132	0
		PC 空间欧氏距离	18	0	132	0
	标准正态变量变 换 (SNV) 结合 Savitzky-Golay 卷积求导	PC 空间明氏距离	18	0	132	0
		PC 空间街区距离	18	0	132	0
		PC 空间马氏距离	18	0	132	0
		Hotelling T2	18	0	132	0
		欧氏距离	17	1	17	115
		明氏距离	17	1	16	116
		街区距离	17	1	17	115
		PC 空间欧氏距离	17	1	16	116
		PC 空间明氏距离	17	1	39	93
		PC 空间街区距离	17	1	16	116
		PC 空间马氏距离	11	7	132	0
		Hotelling T2	12	6	132	0
	连续小波变换 (CWT)	欧氏距离	18	0	39	93
		明氏距离	18	0	43	89
		街区距离	18	0	39	93
		PC 空间欧氏距离	18	0	39	93

第四章 基于便携式近红外光谱仪的判别分析研究

		PC 空间明氏距离	18	0	84	48
		PC 空间街区距离	18	0	39	93
		PC 空间马氏距离	7	11	132	0
		Hotelling T2	9	9	132	0
		欧氏距离	18	0	36	96
		明氏距离	18	0	39	93
		街区距离	18	0	36	96
	Savitzky-Golay 卷积求导	PC 空间欧氏距离	18	0	36	96
		PC 空间明氏距离	18	0	85	47
		PC 空间街区距离	18	0	36	96
		PC 空间马氏距离	17	1	132	0
	标准正态变量变 换 (SNV) / 多元 散射校正 (MSC)	Hotelling T2	17	1	132	0
		欧氏距离	18	0	36	96
		明氏距离	18	0	36	96
		街区距离	18	0	36	96
		PC 空间欧氏距离	18	0	36	96
		PC 空间明氏距离	18	0	61	71
		PC 空间街区距离	18	0	36	96
		PC 空间马氏距离	17	1	131	1
		Hotelling T2	17	1	130	2
		欧氏距离	18	0	59	73
		明氏距离	18	0	46	86
		街区距离	18	0	59	73
	标准正态变量变 换 (SNV) 结合连 续小波变换 (CWT)	PC 空间欧氏距离	18	0	57	75
		PC 空间明氏距离	17	1	105	27
		PC 空间街区距离	18	0	57	75
		PC 空间马氏距离	11	7	132	0
		Hotelling T2	15	3	132	0
		欧氏距离	18	0	81	51
		明氏距离	18	0	82	50
		街区距离	18	0	81	51
	标准正态变量变 换 (SNV) 结合 Savitzky-Golay 卷积求导	PC 空间欧氏距离	18	0	78	54
		PC 空间明氏距离	17	1	115	17
		PC 空间街区距离	18	0	78	54
		PC 空间马氏距离	15	3	132	0
		Hotelling T2	17	1	132	0

从表 4-2 中可以看出, 对于 A 公司生产的维库溴铵样品建立近红外判别模型时, 预处理方法如果选用连续小波变换, 那么在八种不同距离判据下均出现

了一定数量的类内误判或类外误判，相比类外样品，模型判别类内样品效果更好一点；预处理方法选用 Savitzky-Golay 卷积求导，那么在 PC 空间明氏距离下可以在类内样品和类外样品中同时达到零误判，其它距离下尽管类内样品没有或很少出现误判，但是对于类外样品的判别效果并不理想，比较之下，Savitzky-Golay 卷积求导结合 PC 空间明氏距离建立的判别模型判别本类样品具有较好的效果；标准正态变量变换处理后，所有距离下的判别模型在类内样品判别均达到了零误判，然而在类外样品判别时大部分出现了误判情况，只有 PC 空间马氏距离和 Hotelling T2 判据没有出现误判；将标准正态变量变换结合连续小波变换方法处理样品光谱，在八种距离下均出现了一定数量的类内或者类外样品误判，相比之下标准正态变量结合 Savitzky-Golay 卷积求导方法预处理样品光谱时，八种距离下对应的八种判别模型有六种可以达到零误判，并且这种预处理方法组合下的判别模型在判别类外样品时具有较高的准确率。综合比较这五种预处理方法及其对应各种距离下的模型的判别效果，可以看出，对于 A 公司生产的维库溴铵样品建立判别模型时，采用标准正态变量变换结合 Savitzky-Golay 卷积求导的预处理方法可以是所建立的模型更为稳健有效，距离判据可以在欧氏距离、明氏距离、街区距离、PC 空间欧氏距离、PC 空间明氏距离以及 PC 空间街区距离中选择。五种样品的最优预处理方法以及距离判据的统计结果见表 4-3。

表 4-3 五种样品的最优预处理方法及距离判据组合

样品（维库溴铵）	预处理方法	距离类型	类内验证		类外验证	
			准确数	误判数	准确数	误判数
A 公司（共 15 个）	SG	PC 空间明氏距离	15	0	138	0
	SNV	PC 空间马氏距离	15	0	138	0
		Hotelling T2				
	SNV+CWT	PC 空间明氏距离	15	0	138	0
	SNV+SG	欧氏距离	15	0	138	0
		明氏距离				
		街区距离				
		PC 空间欧氏距离				
		PC 空间明氏距离				
	B 公司（共 18 个）	SNV	PC 空间街区距离	16	2	132
Hotelling T2						

第四章 基于便携式近红外光谱仪的判别分析研究

C 公司 (共 15 个)	SNV	PC 空间马氏距离	15	0	138	0
D 公司 (共 18 个)	CWT	欧氏距离等	18	0	132	0
E 公司 (共 18 个)	SNV	PC 空间马氏距离	17	1	131	1

从上表可以看出, A、B、C、D、E 对应的最佳预处理方法中都含有标准正态变量变换 (SNV), 但是相应的距离判据却不尽相同, 例如 A 公司样品中 SNV 对应的距离判据有 PC 空间马氏距离和 Hotelling T2 两种, 而 C 公司对应的距离判据中 Hotelling T2 就不是其最佳的距离判据。

建立判别模型时, 不同的样品有适合各自不同的预处理方法, 建立模型时, 需要根据实际情况灵活地选用最佳的预处理方法和距离判据组合。

第四节 结论

本章针对使用 JDSU MicroNIR™ 1700 仪器所扫描得到的五种样品光谱, 分别比较了不同预处理方式下所建立判别模型的预测效果, 并根据结果对每一种样品提出了建立判别模型使用的最佳光谱预处理方法和距离判据组合。实验结果也同时表明, JDSU MicroNIR™ 1700 仪器对实验中的五种样品具有比较准确的测量结果, 具有实际应用价值。

参考文献

- [1] 褚小立, 陆婉珍. 近红外光谱仪新进展[J]. 分析仪器, 2008, (1): 3-7.
- [2] Sfontouris I A, Lainas G T, Sakkas D, *et al.* Non-invasive metabolomic analysis using a commercial NIR instrument for embryo selection[J]. Journal of Human Reproductive Sciences, 2013, 6(2): 133-139.
- [3] Liu Y D, Gao R J, University E C J, *et al.* Review of Portable NIR Instruments for Detecting Fruit Interior Quality[J]. Guang pu xue yu guang pu fen xi, 2010, 30(10): 2874-2878.
- [4] Smith M R, Jee R D, Moffat A C, *et al.* A procedure for calibration transfer between near-infrared instruments--a worked example using a transmittance single tablet assay for piroxicam in intact tablets[J]. Analyst, 2004, 129(9): 105003-105007.
- [5] Cozzolino D, Cynkar W U, Damberg R G, *et al.* In Situ Measurement of Soil Chemical Composition by Near-Infrared Spectroscopy: A Tool Toward Sustainable Vineyard Management[J]. Communications in Soil Science & Plant Analysis, 2013, 44(44): 1610-1619.
- [6] Ir MPU, Instrumentation N S. Market Profile: Used IR and NIR Spectroscopy Instrumentation[J]. Spectroscopy, 2013, 72(3): 91-93.
- [7] Järvinen K, Hoehe W, Järvinen M, *et al.* In-line monitoring of the drug content of powder mixtures and tablets by near-infrared spectroscopy during the continuous direct compression tableting process[J]. European Journal of Pharmaceutical Sciences Official Journal of the European Federation for Pharmaceutical Sciences, 2013, 48(4-5): 680-688.
- [8] Wu Z, Bing X, Min D, *et al.* Validation of a NIR quantification method for the determination of chlorogenic acid in *Lonicera japonica* solution in ethanol precipitation process[J]. Journal of Pharmaceutical & Biomedical Analysis, 2012, 62(2): 1-6.
- [9] Krapf L C, Nast D, Gronauer A, *et al.* Transfer of a near infrared spectroscopy laboratory application to an online process analyser for in situ monitoring of anaerobic digestion[J]. Bioresour Technol, 2012, 129C(2): 39-50.
- [10] Snively J B, Taylor M J, Jenniskens P. Airborne imaging and NIR spectroscopy of the ESA ATV spacecraft re-entry: instrument design and preliminary data description[J]. International Journal of Remote Sensing, 2011, 32(11): 3019-3027.
- [11] Rocha J D, Bachilo S M, Ghosh S, *et al.* Efficient spectrofluorimetric analysis of single-walled carbon nanotube samples[J]. Anal Chem, 2011, 83(19): 7431-7433.
- [12] [1] Kapitán J, Barron L D, Hecht L. A novel Raman optical activity instrument operating in the deep ultraviolet spectral region[J]. Journal of Raman Spectroscopy, 2015, 46(4): 94-96.
- [13] Benedetti C, Abatzoglou N, Simard J S, *et al.* Cohesive, multicomponent, dense powder flow characterization by NIR[J]. Int J Pharm, 2007, 336(2): 292-301.

- [14] Poddar R, Cortés D E, Werner J S, *et al.* Publisher's Note: Three-dimensional anterior segment imaging in patients with type 1 Boston Keratoprosthesis with switchable full depth range swept source optical coherence tomography[J]. *Journal of Biomedical Optics*, 2013, 18(8): 86002-86002.
- [15] JDSU's Flex Product Enable DuPont Performance Coating to Create Breakthrough Custom Finishes for the Automotive Industry[J]. *Optics Mechanics & Electronics Information*, 2005, 2(12): 73-75.
- [16] M F, Löfstedt T, *et al.* Near infrared and skin impedance spectroscopy – a possible support in the diagnostic process of skin tumours in primary health care[J]. *Skin Research & Technology*, 2015, 21(4): 493-499.
- [17] Liu Y, Campbell B T, Delhom C, *et al.* Variation and relationship of quality and near infrared spectral characteristics of cotton fibers collected from multi-location field performance trials[J]. *Textile Research Journal*, 2015, 85(14): 53-59.
- [18] 孙景群. 光器件的回损测量[J]. *通讯世界*, 2012, (7): 73-74.