



“十三五”科学技术专著丛书

光谱数据处理与 定量分析技术

李志刚 著

Spectral Data Processing and
Quantitative Analysis



北京邮电大学出版社
www.buptpress.com



“十三五”科学技术专著丛书

光谱数据处理与定量分析技术

李志刚 著



北京邮电大学出版社
www.buptpress.com

内 容 简 介

光谱数据处理与定量分析技术是光谱信号处理领域很重要的研究方向。本书分为8章:第1章是绪论,介绍近红外光谱分析技术的发展历程和展望;第2章介绍光谱分析技术的基础知识;第3章系统介绍傅里叶光谱分析仪器的的工作原理;第4章重点介绍拉曼光谱仪的基本原理;第5章介绍光谱数据处理与建模技术;第6章系统介绍基于奇异摄动技术的导数光谱获取技术,并利用其对光谱进行平滑与求导,同时与经典SG算法比较分析;第7章重点介绍数据集的划分技术以及对建模性能的影响,详细介绍了改进型SPXY数据集划分策略以及基于多空间的共识数据集划分策略,并进行了实验分析;第8章介绍集成建模策略,尤其以多空间集成建模作为核心内容进行系统的介绍和深入的分析,并通过实验与其他集成模型进行了详尽比较。

本书全面、系统地展示了近红外和拉曼光谱数据处理与定量分析的研究内容和最新成果,具有完整性、实用性和学术性。本书适合作为生物医学工程、电子信息工程、测试计量技术及仪器等专业在学习信号处理领域知识时的教学拓展,还可供光谱分析与建模领域从事科研工作和工程应用的人员参考。

图书在版编目(CIP)数据

光谱数据处理与定量分析技术 / 李志刚著. -- 北京: 北京邮电大学出版社, 2017.7

ISBN 978-7-5635-5074-6

I. ①光… II. ①李… III. ①光谱分析 IV. ①O433.4

中国版本图书馆CIP数据核字(2017)第085147号

书 名: 光谱数据处理与定量分析技术

著作责任者: 李志刚 著

责任编辑: 刘 颖

出版发行: 北京邮电大学出版社

社 址: 北京市海淀区西土城路10号(邮编:100876)

发行部: 电话: 010-62282185 传真: 010-62283578

E-mail: publish@bupt.edu.cn

经 销: 各地新华书店

印 刷: 北京九州迅驰传媒文化有限公司

开 本: 787 mm×1 092 mm 1/16

印 张: 11.75

字 数: 288千字

版 次: 2017年7月第1版 2017年7月第1次印刷

ISBN 978-7-5635-5074-6

定 价: 30.00元

• 如有印装质量问题,请与北京邮电大学出版社发行部联系 •

前 言

红外及拉曼光谱都是分子振动光谱。通过谱图解析可以获取分子结构的信息。任何气态、液态、固态样品均可进行红外光谱测定,这是其他仪器分析方法难以做到的。由于每种化合物均有红外吸收,尤其是有机化合物的红外光谱能提供丰富的结构信息,因此红外光谱法是有机化合物结构解析的重要手段之一,广泛应用于有机化学、高分子化学、无机化学、化工、催化、石油、材料、生物、医药、环境等领域。但是目前在低含量、复杂溶液的检测分析方面,在样本空间构建、数据集划分、数据信息挖掘、建模预测和稳定性方面,仍存在着很多尚未成熟解决的问题,这些问题制约了光谱分析技术进一步的应用。本书侧重于从算法层面提升光谱分析技术的预测性能,具有一定的参考价值。

本书主要对光谱数据处理及定量分析技术加以介绍,在结构上分为 8 章:

第 1 章是绪论,介绍近红外光谱分析技术的发展历程,对其国内外发展状态进行了综述,并进行了展望。

第 2 章是光谱分析技术的基础知识的介绍。

第 3 章重点介绍傅里叶变换红外光谱仪基本原理及主要特点。

第 4 章重点介绍拉曼光谱仪基本原理及主要特点。

第 5 章是常用经典光谱数据处理与建模技术的介绍。

第 6 章是基于奇异摄动技术的导数光谱获取技术的系统介绍,并利用其对光谱进行平滑与求导,同时与经典 SG 算法比较。

第 7 章是介绍数据集划分技术以及对建模性能的影响,详细介绍了改进型 SPXY 数据集划分策略,以及基于多空间的共识数据集划分策略,并进行了实验分析与探讨。

第 8 章是介绍集成建模策略,尤其是多空间集成建模作为核心内容进行了系统的介绍和深入的分析,并通过实验与其他集成模型进行了详尽的比较。

本书具有以下鲜明特色。

(1) 前沿性:内容丰富全面,结构合理,体系完整,将近红外光谱的数据处理和定量分析技术进行全面和系统的介绍。针对近红外光谱分析研究的挑战性难点问题进行分析,以目前和预期内的研究成果为支撑,给出严谨合理的论述。

(2) 应用性:本书提供的分析方法及研究对象都以解决实际应用中的问题和难点为出发点,给出具体的应用实例,具有很强的实用性。

(3) 学术性:本书具有一定的理论高度和学术价值,书中绝大部分内容取材于作者近期在国际、国内高端学术期刊和重要国际会议上发表的论文,具有很高的学术参考价值。

在本书的编写工作中,东北大学秦皇岛分校的王巧云副教授参加了第4~8章部分内容的编写工作,作者在此深表谢意。

由于作者水平有限,书中难免有疏漏与错误之处,诚请读者批评指正。

李志刚

2016年10月于东北大学秦皇岛分校

目 录

第 1 章 绪论	1
1.1 光谱分析技术的发展历程	1
1.2 光谱分析化学计量学概述	2
1.3 光谱分析技术应用领域	3
1.4 光谱分析技术展望	5
本章参考文献	6
第 2 章 光谱分析技术的基础知识	10
2.1 分子振动的吸收频率	10
2.2 分子振动能级(红外吸收产生的条件)	12
2.3 朗伯-比尔定律	13
2.4 漫反射理论	14
2.5 ATR 理论(衰减全反射)	16
本章参考文献	19
第 3 章 傅里叶光谱分析仪器的基本原理	21
3.1 傅里叶变换红外光谱仪	21
3.1.1 仪器组成	21
3.1.2 光学系统及工作原理	22
3.2 仪器的类型及性能指标	26
3.2.1 仪器的类型	26
3.2.2 仪器的性能指标	28
本章参考文献	30
第 4 章 拉曼光谱仪器的基本原理	32
4.1 拉曼散射基本原理概述	33
4.2 拉曼散射的经典及量子解释	34
4.3 常用拉曼光谱技术简介	37
4.3.1 共振拉曼光谱	37
4.3.2 表面增强拉曼光谱	38
4.3.3 SERS 联用技术	40

4.4	拉曼光谱的应用	42
4.5	拉曼光谱定量分析方法概述	45
4.5.1	拉曼光谱定量分析的理论基础	45
4.5.2	拉曼光谱定量分析流程及应用案例	45
	本章参考文献	49
第5章	光谱数据处理与建模技术	52
5.1	光谱常用的预处理方法	52
5.1.1	基线校正方法	53
5.1.2	数据平滑	53
5.1.3	数据求导	57
5.1.4	标准正态变量校正法	57
5.1.5	多元散射校正法	58
5.1.6	正交信号校正	58
5.1.7	快速傅里叶变换	59
5.1.8	基于小波变换的滤波方法	59
5.2	光谱定量分析建模方法	61
5.2.1	多元线性回归分析法	61
5.2.2	主成分回归分析法	62
5.2.3	偏最小二乘回归分析法	63
5.2.4	间隔偏最小二乘法	64
5.2.5	非线性建模方法	66
5.3	模型评价指标	68
	本章参考文献	69
第6章	基于奇异摄动技术的导数光谱获取技术	71
6.1	跟踪-微分器介绍	71
6.1.1	渐近稳定系统的概念	71
6.1.2	跟踪-微分器的一般形式	72
6.2	奇异摄动高增益积分链式微分器	74
6.3	逆泰勒级数导数光谱估计器的设计	76
6.4	导数光谱定量分析应用对比	78
	本章参考文献	87
第7章	数据集的划分	90
7.1	随机划分	90
7.2	KS划分算法	90
7.3	SPXY划分算法	91
7.4	Y空间划分算法	91

7.5	λ -SPXY 可调参数算法	92
7.6	共识划分策略	96
7.6.1	啤酒样本验证结果	99
7.6.2	葡萄酒样本验证结果	100
	本章参考文献	102
第 8 章	集成建模技术	104
8.1	叠加偏最小二乘建模	104
8.2	双叠加偏最小二乘	105
8.3	PLS 全谱多导数谱空间集成建模	105
8.3.1	全谱集成建模方法	106
8.3.2	多空间样本划分策略	107
8.3.3	全血葡萄糖和胆固醇测定应用	108
8.4	iPLS 多导数空间集成建模	112
8.4.1	iPLS 全谱范围多导数空间集成建模	112
8.4.2	iPLS 选择区间多空间集成建模	119
8.5	基于杂化光谱空间 iPLS 集成建模	129
8.5.1	基于多空间的样本集综合划分策略	129
8.5.2	杂化谱空间集成区间 PLS 建模	129
8.5.3	实验验证	130
	本章参考文献	135
附录	程序	138

第1章 绪论

1.1 光谱分析技术的发展历程

红外光是指波长介于可见光与微波之间的一种电磁波。1672年牛顿把通过玻璃棱镜的太阳光分解成了从红光到紫光的各种颜色的光谱,发现太阳光(白光)是由各种颜色的光复合而成的,同时做出了单色光在性质上比白光更简单的著名结论。1800年英国物理学家赫谢尔(Herschel)用棱镜使太阳光色散,研究各部分光的热效应,意外发现在红光的外侧具有最大的热效应,说明红色光的外侧还有辐射存在,当时把它称为“红外线”或“热线”^[1]。1802年渥拉斯顿、夫琅和费通过圆孔衍射观察到了光谱线,之后公布了太阳光谱中的许多条暗线,并以字母命名。19世纪中期,瑞典物理学家埃斯特朗探测出氢原子光谱中最强的一条谱线,此后氢原子光谱一直是光谱学研究的重要课题之一。同时期,基尔霍夫与本生发展了实用光谱学,证明了光谱学可以用作定性化学分析的新方法,并利用这种方法发现了几种当时还未知的元素,还证明了太阳里也存在着多种已知的元素。19世纪末,Abney和Festing在近红外光谱的短波区域第一次记录到了有机化合物的近红外区的谱图。由于当时没有精密仪器可以检测,仅限于几个实验室的研究,所以一直未能得到实际应用^[2]。这是红外光谱的萌芽阶段。

20世纪50年代后,随着简易型近红外光谱仪器的陆续出现以及仪器研发水平的不断提高,美国的Karl Norris等人做了大量的工作,采用可见光透射与反射技术对鸡蛋、蔬菜和水果进行测定^[3-7],由于可见区信息量小,研究工作受到限制。但他们注意到水分对近红外光的强烈吸收,于是开始了运用近红外光谱测定农产品的水分含量的工作。他们采用多组分测定技术来排除样品中其他成分所带来的干扰,并同时测定几种成分的透射率,用统计学的方法分别计算出各自的含量;采用漫反射技术来克服固体样品中低透射率的难题;发展了一套检验方法,用来解决某种计算方法的取舍问题,即如果计算测定值与样品的真实值相符就采用该方法,反之舍去^[8]。Norris等人在近红外光谱漫反射技术上所做的大量研究,使近红外光谱技术在农副产品分析上得到了广泛的应用^[9-14]。这是红外光谱技术的发展期。

20世纪60年代中期,红外光谱分析技术出现灵敏度低、抗干扰能力差、分辨率低、单次测量噪声大、谱带重叠及尖峰测量困难等问题,再加上其他新兴分析技术的冲击,红外光谱分析技术的研究工作陷入低谷,除最初开展的农副产品的研究工作外,在其他新的领域基本上没有发展。

20世纪80年代,仪器分析技术与化学计量学随着计算机技术和数学应用的发展而迅速发展^[15]。由于通过化学计量学的方法来提取光谱信息及背景干扰的实现,加上红外光谱

分析技术本身的优点,红外光谱技术得到了人们重新认识,进而被快速地推进到各个领域,近红外光谱分析技术开始有了起色^[16]。

20世纪90年代后,红外光谱技术逐渐引起了国际分析界的关注,近年来近红外光谱技术一直是匹兹堡(PITTCO)会议的热点^[17]。在2000年的匹兹堡会议上,该技术被认定为此次会议上最受重视的光谱分析方法^[18]。在每届关于近红外光谱研究的国际会议上,都会刊登大量涉及近红外光谱仪、化学计量学、光谱预处理和各种新应用的文章。随着光纤技术在近红外光谱仪上的应用,出现了现场光谱技术,从而实现在线分析。同时,化学计量学的不断发展也使分析技术在生物化学、医药、农业、食品工业等领域得到了广泛的应用,从而红外光谱技术进入一个蓬勃发展的时期。

1.2 光谱分析化学计量学概述

红外光谱分析技术的发展与化学计量学(Chemometrics)的发展息息相关。化学计量学是光谱分析中的重要技术,为光谱分析提供光谱预处理、数据集划分、定性定量分析的理论基础。化学计量学是一门集数学、统计学、计算机科学与化学于一体的交叉学科,应用数学、统计学方法研究化学测量中的信号分辨与解析、数据处理、化学分类与预测等问题,最大限度地获取有关物质的化学信息并构建更稳健的预测模型^[19]。

化学计量学最初是由瑞典化学家沃尔德(S. Wold)在1971年为基金定名时首次提出的,并在1972年瑞典的科学杂志上发表了化学计量学的相关论文。这一命名得到科瓦斯基(B. R. Kowalski)的支持,并将他的研究小组称为化学计量学实验室^[20]。1974年B. R. Kowalski和S. Wold成立国际化学计量学学会,并在美国华盛顿大学召开成立大会,同时正式创办学会的机关刊物《化学情报与计算机科学杂志》(*Journal of Chemical Information and Computer Science*),标志着这一学科的诞生^[21]。1975年国际化学计量学学会给化学计量学下了一个明确的定义:化学计量学是化学的一个分支学科,应用数学和统计学方法设计和选择测量过程和实验,并通过解析化学数据最大限度地获取化学信息。随后,有多种有关化学计量学的专著及期刊出版,有多个大型国际学术会议召开,如美国化学会计算机分会的*Chemometrics: Theory and Applications*、1978年Massart等人的*Evaluation and Optimization of Laboratory Methods and Analytical Procedures*、1981年Kateman和Pijpers的*Quality Control in Analytical Chemistry*、1986年Sharaf等人的*Chemometrics*等。随着分析测试技术、测量仪器以及计算机科学的迅速发展,化学计量学在统计学、数学与化学以及计算机科学技术的交叉点处也得以迅速发展。目前,化学计量学已作为化学量测的基础理论与方法,成为化学发展的前沿领域之一^[22]。在红外光谱分析中,化学计量学主要包括以下三个方面应用。

(1) 光谱数据预处理

光谱数据的预处理是红外光谱分析的关键部分,是建立稳健模型的重要基础。红外光谱测量中会出现由于光谱不同程度吸收而导致的光谱图重叠,会包含与待测样品性质无关的干扰信息,会出现光的散射、杂散光和光谱仪器响应造成的光谱重现性差和基线漂移等现象,致使光谱数据中包含了噪声。因此需要对光谱数据进行预处理,以优化测量精度,获得更为稳健、可靠的校正模型。

光谱数据预处理包括光谱数据均值中心、标准化、归一化、Savitzky-Golay(SG)平滑与导数法、基于奇摄动技术的光谱导数估计器(Singular Perturbations Spectrum Estimator, SPSE)法等多种方法。详细内容将在第5章进行具体介绍。

(2) 定量分析

定量分析,即多元校正,是在物质浓度(或其他物化浓度)与分析仪器响应值之间建立的定量关系^[23],可通过校正模型直观反映出来。光谱分析中常用的多元校正的方法有:多元线性回归法(MLR)、主成分回归法(PCR)和偏最小二乘(PLS)等线性校正方法,以及人工神经网络(ANN)和支持向量机(SVM)等非线性校正方法。定量分析广泛应用于制药、化工、食品检测等诸多领域。

(3) 定性分析

在红外光谱技术实际应用中,经常遇到只需知道物质的类别或质量等级,无须知道样品的组分数及含量,也就是化学计量学中的定性分析问题。化学计量学中定性分析类似于计算机技术中的模式识别,是将红外光谱数据根据某一特征或某一原则进行分类识别,把具有共同特征的样本归为一类,获得决策性信息。定性分析方法一般包括两大范畴:有监督式和无监督式。有监督式是用一组已知类别的样本作为训练集,利用计算机对其进行学习,然后用未知的样本进行预测,得到未知样本的类别信息。而无监督式是利用某一固定规则直接对未知样本进行分类识别。常见的有监督的识别方法有K-最近邻法、BP神经网络、贝叶斯(Bayes)判别法;无监督的识别方法有聚类分析法、无监督神经网络法。

1.3 光谱分析技术应用领域

光谱分析技术作为一种速度快而且效率高的分析方法,在农业、石油化工业、食品业和制药工业等方面都得到了广泛的应用,也可以对气体、液体、粉末、固体以及生物的组织等各种样品进行快速准确的定性及定量分析。由于光纤技术的引入,光谱分析技术可以实现恶劣或危险环境下的在线分析。

(1) 农业领域方面的应用

农业产品分析是近红外光谱的传统应用领域,大多采用的是漫反射原理,因此在农业和畜牧业方面,它的应用和发展非常快。近红外光谱分析方法最早用于测量小麦中所含的水分的量^[24],我国曾经将此项技术应用在小麦优良品种的选取上,这样在很大程度上提高了工作效率。加拿大的谷作物研究室使用此项技术快速测定硬质小麦中黄色颜料的含量,分析得到的结果与用标准方式测定的结果十分相近^[24]。近红外光谱技术应用于油料作物方面,主要是在育种的过程中快速提取某一品种;应用于果蔬方面,主要是测量水果和蔬菜内部维生素或蛋白质等的含量,这种检测已经逐步发展为在线分析技术;应用于饲料方面,可分析饲料中所含有的粗蛋白、纤维素、水分以及脂肪等组分的含量;应用于土壤方面,可分析土壤中含有的水分和有机质等一些指标;在检测茶叶成分方面,可用于对茶叶制品中不同组分的快速分析测定。Barton等人根据不同的波长区域建立的化学计量学模型对稻米的化学组成(淀粉、蛋白质、水分)和物理性质(透明度、研磨程度)进行质量评估^[25]。卢利军等人利用红外光谱分析技术快速测定了黄豆粕中的蛋白质、水分、粗脂肪含量,检测结果令人满意^[26]。Mouazen等人设计了一种土壤在线分析系统,使用红外漫反射光谱分析土壤pH

值、磷、碳及水分含量以评价土壤性能^[27]。张艳哲、于洁等人利用近红外光谱技术结合 MLR 法对活性米品质进行检测研究,其所建模型可以有效预测活性米的品质,具有较高的预测能力和精度^[28]。

(2) 石油化工领域的应用

近红外光谱技术在国内外石油和化工方面有着广泛的应用,并且已经给我们带来了极大的经济效益。20 世纪 80 年代后期,美国华盛顿的过程分析中心对石油品质的分析进行了大量的研究工作^[29]。研究表明,近红外光谱分析辛烷值在时间和成本上的效益都极其显著,并且可用于在线控制分析,这一突出研究成果得到了各国的广泛应用^[30]。邵波等人利用红外光谱法,快速检测了成品罐汽油的有关性质,可在 5 min 内快速、准确地一次测定其辛烷值、芳烃含量^[31]。史月华等人利用红外光谱法,采用主成分回归算法快速测量了汽油中辛烷的含量,使模型的预测能力有所改善^[32]。随着分析技术的快速发展,近红外光谱技术的应用方面也得到扩展,比如对润滑油和燃油的性质和组成进行分析,汽油炼制中各种化学物质含量的测定,石油加工过程的参数检测。Choquette 研究了用红外光谱测定汽油中含氧化物的含量^[33]。在发达国家中,几乎全部的炼油制造厂在各种化工和炼油工艺中都采用近红外光谱的在线分析技术^[34],使提炼油的过程变得低成本、高经济效益以及高质量,与此同时减少了测量时间、提高了效率。

(3) 食品工业的应用

应用近红外光谱技术对肉制品中水分、淀粉、脂肪以及蛋白质的含量进行在线检测,方便快捷,这种检测方法已经广泛应用于对肉制品的检测^[35]。此项技术还可用于检测火腿的 pH 值及其中的氯化钠、生肉的卡路里以及肉块的一些物理化学的特性。Savenije B. 等人利用可见/红外反射光谱测定猪肉的色度、pH 值等质量参数。实验证明,此法可用于猪肉质量的快速鉴定^[36],同时也可以检测奶制品中脂肪、蛋白质和乳糖等的含量,而且精度高、分析速度快。另外,应用近红外光谱技术测定各种液体饮料及食用油样品的营养成分含量,如水分、蛋白质、纤维、维生素等。原娇娇以茶油为研究对象,研究茶油的脂肪酸含量,形成其特征图谱,并运用红外光谱法,提出了茶油品质的快速鉴别方法^[37]。不仅如此,近红外光谱技术在分析啤酒、葡萄酒和黄酒样品的酒精含量等方面也起到了很重要的作用,不但可以鉴定酒的品种,还可以定量分析其中的酒精度、总糖和总酸。樊双喜等人通过研究黄酒的近红外光谱和利用化学计量学的技术,采用偏最小二乘法建立快速检测黄酒的酒精度、总糖和总酸的近红外模型,研究结果表明,近红外光谱法可用于快速检测黄酒的酒精度、总糖和总酸^[38]。

(4) 制药工业方面的应用

在制药业中,近红外光谱技术不仅可以用于对药物的活性组分以及其他组分的含量进行在线控制和鉴定,还可以用于鉴别中药材。在对药物进行定性分析方面,已经采取了多种分析方法,如化学计量学方法和多类支持向量机等分析方法^[8],对药物的真品和赝品以及类别进行分析研究。而在对药物进行定量分析方面,可用于测得不同种类中药的掺入量以及药物中含有的有效成分含量,并且已经开始应用于药物生产过程中的实时监测。瞿海斌等人利用红外漫反射光谱法,对黄连浸膏粉中的生物碱含量进行了快速、无损检测^[39],结果表明,此法可以用于工业现场的在线检测。Merckle P 等人利用红外光谱法,测定了 3 种不同泡腾片中阿司匹林的含量,结果表明 NIR 法可快速、有效地测定阿司匹林的含量^[40]。Ulm-

schneider M. 等人用红外光谱结合相应的多元校正方法快速地鉴别了 11 种固体药用辅料,该方法可在生产线和实验室使用^[41]。

除了在农业、石油化工业、食品业和制药工业方面得到了广泛应用以外,近红外光谱技术还可以应用于对血糖和血氧量的测量以及对细胞病理的鉴别等方面,因此近红外光谱技术在临床医学方面已经表现出极大的发展潜力。Yoan-Joo、Kim 等人利用中红外光谱研究样本校正集与验证集划分方法对全血中的葡萄糖浓度变化的影响^[42]。暨南大学的黄富荣、罗云瀚等人采用近红外透射光谱结合化学计量学方法建立了全血中胆固醇和甘油三酯的定量分析模型,验证了近红外光谱测量全血中胆固醇和甘油三酯的可行性^[43]。在环境保护方面,此项技术还可以用于对海平面上溢出浮油的种类和水里的有机污染物来进行快速鉴别^[44]。

1.4 光谱分析技术展望

随着生活水平的日益提高,人们对健康、安全和环保等方面的需求越来越高,食品安全、药物安全等诸多关系民生的领域已经成为大家关注的焦点。另外,安全无损的医疗诊断技术也是人们的迫切需求。但近年来,在经济利益的驱使下,劣质油品、假酒、假药等伪劣食品、药品事件频发。因此对食品、药品进行快速、精确地定性、定量检测,做到及早发现问题,把安全风险降到最低显得尤为重要。而在这些领域中,复杂混合溶液是其最常见的物质形态之一,如牛奶等液态食品、中药注射液等液态药物。另外,在医疗诊断方面,人们对环保、快速、无害的诊疗与分析技术的需求日益高涨,传统的有损、复杂、高成本、高污染的检测手段与分析技术已经不能满足医疗诊断领域的需求,而基于光学的检测分析技术对人体没有任何伤害,也不需要额外的化学试剂,其正在逐渐显示在医学诊断方面的巨大优势,尤其光谱分析方法在血液检测、肿瘤早期诊断等方面,已经成为全世界重点研究的热点问题。

但是,至今为止,无论是拉曼光谱还是红外光谱,在实际应用过程中,尤其是针对复杂混合溶液的分析应用过程中很容易受到各种干扰信息的影响,给光谱定量分析带来很大的困难,严重制约了光谱技术的应用。目前已知的干扰信息主要有检测器噪声、激光器功率波动,外界环境温度波动,更主要的是由被检测对象(复杂混合液体)成分复杂性引起的噪声、扰动以及非线性变化。这些干扰因素将引起波数、强度的漂移以及光谱信号的混叠,给有效的检测与分析带来极大的困难,因此仍然存在很多亟待解决的、具有挑战性的难题。事实上,光谱信息的质量在很大程度上依赖于后端算法对谱图的处理过程,如果后端信号处理不充分,即便仪器的硬件非常精密,光谱的质量也不能得到保证,没有高质量的谱图数据就无法构建高精度的模型,无法进行定量分析。针对光谱分析实际应用中面临的这些困难和瓶颈,所涉及的问题可以归结为一个学术问题和一个应用问题。(1)学术问题如下:光谱信号解析与融合回归建模方法。光谱解析与融合回归建模方法主要涉及对未知物的解析,研究如何从复杂混合溶液光谱中分析出有哪些物质及各主要成分的含量。已经有学者对光谱分析及建模进行过广泛的应用研究,但多年来,很多共性的光谱处理方法仍采用 20 世纪 70 年代出现的信号处理技术,没有根本性的改进。从当前技术手段来看,处理性能仍有较大提升空间;还有一些环节研究工作才刚刚起步,仍有大量问题有待解决。而建模方面全谱的 PLS 建模方法仍然是主流。近年来,在此基础上发展出间隔偏最小二乘(Interval PLS)、叠

加偏最小二乘(Stacked PLS)以及基于它们的多模型共识集成建模方法。这些方法的核心思想是分割全谱区域为多个系列的子区间,形成子模型,通过不同的策略优选出光谱的特征波段区间(子模型),然后进行加权融合建模,以达到获得更高预测精度的目的。但是,目前所有建模方法都是基于单一光谱形式的融合,其融合策略方面有较大的局限性,信息挖掘的深度也远远不够。本项目从新的角度、基于新的条件和新的应用来研究和解决噪声、扰动、基线漂移以及谱峰混叠等基础问题,深度挖掘光谱信息,建立合理、准确的定量分析模型是光谱技术应用的关键所在。(2)应用问题如下:以复杂混合溶液光谱分析为应用背景(①血液检测分析,研究全血光谱分析方案;②啤酒、红酒等酒品的光谱分析),解决实际应用中的瓶颈和难题。目前基于血液的诊断指标检测与分析主要基于生化分析,从全血到血清、血浆需要进行离心和分离等多个步骤的预处理且需要大量化学试剂,其检测周期长,检测成本高,污染环境。另外,需要专业人员操作,不利于需要血液检测分析的流行病筛查,更主要的是不利于在基层及社区医院的推广。而光谱分析具有无须试剂、分析快捷、环保、低成本的优势,对其进行深入的理论与方法研究,能够促进新型便携血液检测分析设备的研发,促进其在基层和社区医院推广,完善我国基层和社区医院医疗服务水平;光谱分析方法具有进行多指标同时测定(例如,同时定量分析血红蛋白、血糖、胆固醇、甘油三酯等)、非破坏性测定和连续测定等特点,非常适合在线实时定量分析,另外,还可以推广到重症监护中的连续血糖监测应用中。目前基于全血无须分离预处理的血液光谱分析技术的研究工作刚刚开展,具有广阔的研究空间和发展前景。

光谱分析技术应用于在线检测分析、应用于复杂物质成分检测分析已成为研究工作的热点。在这个领域,除了仪器结构上的改进外,核心问题与需要突破的瓶颈是光谱仪对实际应用场合中噪声和大量干扰因素的鲁棒性,也就是光谱分析与建模过程中要发展具有高抗扰能力的光谱数据处理方法,获取高质量的光谱数据,进而完成在线高质量建模与分析的目的。而复杂混合溶液是药品、食品领域的常见物质形态,其光谱具有复杂的干扰因素,对其进行在线、离线的检测均具有一定的挑战性。本书以红酒、啤酒为研究对象,重点研究复杂混合溶液光谱信号的特征以及从具体应用中归纳提炼出的基础问题。这些问题中有些是与其他分析仪器、其他应用领域所共有的问题,如温度校正、波数校正、基线校正、噪声去除、定量分析等。本书从新的角度、基于新的条件和新的应用来研究这些基础问题。

目前,光谱分析技术作为无损、环保、安全的检测技术已经在许多领域进行了成功的应用,尤其是相对于经典物理、化学方法,核磁共振法,色谱法等。光谱分析技术可脱离实验室环境,可降低检测成本,可避免环境污染,可实现快速、精确的在线检测。因此,推进光谱技术在食品、药品、医疗诊断方面的基础应用研究具有现实意义。

本章参考文献

- [1] Herschel W. Experiments on the refrangibility of the invisible rays of the sun. by william herschel, II. dfrs[J]. Philosophical Transactions of the Royal Society of London, 1800, 284-292.
- [2] Abney W, Festing E R. Near-Infrared Spectral of Organic Liquids[J]. Philos Trans R Soc, 1881, 172: 887.

- [3] Birth G. S., Norris K H. An instrument using light transmittance for nondestructive measurement of fruit maturity[J]. *Food Technology*, 1958, 12(11): 592-595.
- [4] Norris K H, Rowan J D. Automatic detection of blood in eggs[J]. *Agricultural Engineering*, 1962, 43(3): 154-159.
- [5] Hart J R, Norris K H, Golumbic C. Determination of the moisture content of seeds by near-infrared spectrophotometry of their methanol extracts[J]. *Cereal Chem*, 1962, 39(2): 94-99.
- [6] Massie D R, Norris K H. Spectral reflectance and transmittance properties of grain in the visible and near infrared[J]. *Trans. ASAE*, 1965, 8(4): 598-600.
- [7] Norris K H. Design and development of a new moisture meter[J]. *Agric. Eng*, 1964, 45(7): 370-372.
- [8] 张小超, 吴静珠, 徐云. 近红外光谱分析技术及其在现代农业中的应用[M]. 北京: 电子工业出版社, 2012.
- [9] Birth G S, Zachariah G L. Spectrophotometry of agricultural products[J]. *Trans. ASAE*, 1973, 16(3): 548-552.
- [10] Watson C A. Near infrared reflectance spectrophotometric analysis of agricultural products[J]. *Analytical Chemistry*, 1979, 49(9).
- [11] Williams P C, Stevenson S G, Irvine G N. Testing wheat for protein and moisture with the automated digital analyzer[J]. *Cereal Chemistry*, 1978.
- [12] Birth G S. RADIOMETRIC MEASUREMENT OF FOOD QUALITY - A REVIEW[J]. *Journal of Food Science*, 1979, 44(4): 949-957.
- [13] Barton F E, Windhatn E T, Champagne E T, et al. Optimal geometrics for the development of rice quality spectroscopy chemometric models [J]. *Cereal Chem*, 1988(3): 315-319.
- [14] Albanell E, Plaixats J, Ferret A. Evaluation of near-infrared reflectance spectroscopy for predicting stover quality trait in semi-exotic populations of maize [J]. *Sci Food Agric*, 1995, (69): 269-273.
- [15] 徐广通, 袁洪福. 现代近红外光谱技术及应用进展[J]. *光谱学与光谱分析*, 2000, 20(2): 134-142.
- [16] 严衍禄, 赵龙莲, 韩东海, 等. 近红外光谱分析基础与应用[M]. 北京: 中国轻工业出版社, 2005.
- [17] 金钦汉. 从 '99 匹兹堡会议看分析仪器的跨世纪发展动向[J]. *分析仪器*, 1999(2): 1-5.
- [18] 金钦汉. 从 2001 年匹兹堡会议看分析化学和分析仪器发展的几个重要动向[J]. *现代科学仪器*, 2001(3): 21-22.
- [19] 梁逸曾, 吴海龙, 俞汝勤. 化学计量学[J]. *现代科学仪器*. 1988(5): 3-6.
- [20] Currie L A, Filliben J J, Devoe J R. Statistical and mathematical methods in analytical chemistry[J]. *Analytical Chemistry*, 1972, 44(5): 497.
- [21] 史永刚, 冯新泸, 李子存. 化学计量学[M]. 北京: 中国石化出版社, 2002.

- [22] 陆晓华,施文赵. 化学计量学[J]. 自然杂志. 1992,15(4):286-289.
- [23] 陆婉珍. 现代近红外光谱分析技术[M]. 北京:中国石化出版社,2007.
- [24] 张小超,吴静珠,徐云. 近红外光谱分析技术及其在现代农业中的应用[M]. 北京:电子工业出版社,2012.
- [25] Barton F E, Windham W R, Chanpagne E T, et al. Optimal geometries for the development of rice quality spectroscopic chemometric models[J]. Cereal Chem, 1998, 75(3): 315-319.
- [26] 卢利军,庄树华,李爱军,等. 应用近红外技术测定黄豆粕中水分、蛋白质和粗脂肪[J]. 分子科学学报(中、英文),2001,17(2):115-120.
- [27] Mouazen A M, Maleki M R. On-line measurement of some selected soil properties using a Vis-NIR sensor[J]. Soil & Tillage Research, 2007, 93: 13-27.
- [28] 张艳哲,于洁,王磊,等. 基于近红外光谱的活性米品质检测研究[J]. 农机化研究,2016,38(8):141-145.
- [29] Kelly J J, Barlow C H, Jinguji T M, et al. Prediction of gasoline octane numbers from near-infrared spectral features in the range 660-1215 nm[J]. Anal. Chem.; (United States), 1989, 61:4(4):313-320.
- [30] Andrej Popkowski, Ara Barsamian. New gasoline blending unit started at Poland refinery. oil&gas, 1999,97(27):64-70.
- [31] 邵波,黄小英,王京华. 近红外光谱快速测定成品罐汽油的有关性质[J]. 石油化工,2002,31(10):848-851.
- [32] 史月华,陆勇,徐光明,等. 主成分回归残差神经网络校正算法用于近红外光谱快速测定汽油辛烷值[J]. 分析化学,2001,29(1):87-91.
- [33] Choquette S J, Chesler S N, Duewer D. Identification and Quantitation of Oxygenates in Gasoline Ampules Using Fourier Transform Near-Infrared and Fourier Transform Raman Raman Spectroscopy [J]. Anal. Chem, 1996, 68 (20): 3525-3533.
- [34] 陆婉珍,袁洪福,徐广通,等. 现代近红外光谱技术在石油产品分析中的应用[C]. 中国分析测试协会科学技术奖发展回顾. 2015.
- [35] 黄伟,杨秀娟,张燕鸣,等. 近红外光谱技术在肉类定性鉴别中的研究进展[J]. 肉类研究,2014,28(1):31-34.
- [36] Savenije B. Prediction of pork quality using visible/ near-infrared reflectance spectroscopy [J]. Meat Science, 2006, 73: 181-184.
- [37] 原姣姣. 近红外光谱法对茶油化学成分及快速鉴伪的研究[J]. 生物质化学工程,2013,47(2):64-64.
- [38] 樊双喜,钟其顶,李国辉,等. 近红外光谱法快速检测黄酒中氨基酸态氮的模型优化研究[J]. 酿酒科技,2015(5):11-14.
- [39] 瞿海斌,刘全,程翼宇. 近红外漫反射光谱法测定黄连浸膏粉中生物碱含量[J]. 分析化学,2004,32(4):477-480.
- [40] Merckle P, Kovar K A. Assay of effervescent tablets by Near-Infrared Spectroscopy in transmittance and reflectance mode: acetylsalicylic acid in mono and combination

- formulations[J]. *J Pharm Biomed Anal*, 1998, 17: 365-374.
- [41] Ulmschneider M, Penigault E. Direct identification of key intermediates in containers using Fourier-transform near-infrared spectroscopy through the protective polyethylene primary packaging [J]. *Analisis*, 2000, 28(2): 136-140.
- [42] Yoen-Joo Kim, Sangjoon Hahn, Gilwon Yoon, Determination of glucose in whole blood samples by mid-infrared spectroscopy [J]. *Applied Optics*, 2003, 42(4): 745-749.
- [43] 黄富荣, 罗云瀚, 郑仕富, 等. 全血胆固醇、甘油三酯近红外光谱分析与模型优化 [J]. *光学学报*, 2011, 31(10):267-272.
- [44] Li Z, Lv H, Li T, et al. Reagent-free simultaneous determination of glucose and cholesterol in whole blood by FTIR-ATR [J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2017, 178: 192-197.
- [45] 谈爱玲, 毕卫红, 赵勇. 基于稀疏非负矩阵分解和支持向量机的海洋溢油近红外光谱鉴别分析[J]. *光谱学与光谱分析*, 2011, 31(5):1250-1253.