

小型西瓜品质近红外无损检测的光谱信息采集*

韩东海¹ 常冬² 宋曙辉³ 蒋圣楠⁴ 赵洪卫¹

(1. 中国农业大学食品科学与营养工程学院, 北京 100083; 2. 中粮营养健康研究院食品质量与安全中心, 北京 100020; 3. 北京市农林科学院蔬菜研究中心, 北京 100097; 4. 中国农业大学信息与电气工程学院, 北京 100083)

摘要: 为了实现小型西瓜生长过程中品质的快速无损检测, 对其果实测定时的基础信息采集方式进行了定量分析研究。由 PLS 建模结果可知, 无论是组织含水率还是可溶性固形物含量, 在瓜顶部位和赤道部位处都可获得较好的效果。针对不同深度的化学基础信息, 其建模效果存在一定差异。无论是组织含水率还是可溶性固形物含量, 距瓜皮 2 ~ 3 cm 的中层信息区域都能取得最佳的效果。其中瓜顶部位和赤道部位组织含水率的模型相关系数分别为 0.953 和 0.924, RMSEP 分别为 0.510 和 0.766; 瓜顶部位和赤道部位可溶性固形物含量的相关系数分别为 0.952 和 0.929, RMSEP 分别为 0.666 和 0.732。整体说明小型西瓜果实(品种京秀)近红外检测的最佳基础信息区域为中层信息区。

关键词: 小型西瓜 近红外 无损检测 信息采集

中图分类号: O657.33; S651 **文献标识码:** A **文章编号:** 1000-1298(2013)07-0174-05

Information Collection of Mini Watermelon Quality Using
Near-infrared Non-destructive Detection

Han Donghai¹ Chang Dong² Song Shuhui³ Jiang Shengnan⁴ Zhao Hongwei¹

(1. College of Food Science and Nutritional Engineering, China Agricultural University, Beijing 100083, China
2. Food Quality and Safety Centre, COFCO Nutrition and Health Research Institute, Beijing 100020, China
3. Vegetable Research Center, Beijing Academy of Agriculture and Forestry Science, Beijing 100097, China
4. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China)

Abstract: In order to achieve the fast and non-destructive detection of mini watermelon quality during growth process, the basic information collection methods for fruit measuring were analyzed quantitatively. The PLS modeling results showed that both the tissue water content and soluble solids content could obtain good results at the top position and the equator. The modeling performance had some differences with basic chemical information at different depths, but both the tissue water content and soluble solids content could got the best results in the middle information area which was under 2 ~ 3 cm from the rind. The correlation coefficients of tissue water content at the top location and the equator were 0.953 and 0.924, and the RMSEP were 0.510 and 0.766, respectively. The correlation coefficients of soluble solid content at the top location and the equator were 0.952 and 0.929, and the RMSEP were 0.666 and 0.732, respectively. For the mini Jingxiu watermelon, the best foundation information area for NIR detection was the middle area.

Key words: Mini watermelon Near-infrared Non-destructive detection Information collection

引言

我国是西瓜生产大国,其栽培面积及产量均居

世界之首。西瓜果实多汁,除不含脂肪和胆固醇外,几乎含有人体所需的各种营养成分,具有良好的保健功效及很高的经济价值,西瓜的内在品质也备受

收稿日期: 2012-07-03 修回日期: 2012-11-02

* 国家自然科学基金资助项目(31071555)和现代农业产业技术体系建设专项资金资助项目(CARS-26-22)

作者简介: 韩东海,教授,博士生导师,主要从事食品质量无损检测研究, E-mail: handh@cau.edu.cn

消费者关注。在大规模生产中,西瓜采收主要依靠经验判断,缺乏多元稳定的现代无损分选技术,特别是早期上市的小型西瓜,因其品质好,附加值高,在市场流通过程中经常出现生熟瓜混销现象,这些问题极大地影响了西瓜的消费。因此,在西瓜生产过程和流通领域中,都急需现代无损检测技术对西瓜成熟度和内部品质进行判别。近红外光谱分析技术作为一种应用非常广泛的无损检测技术^[1],它的作用已经得到了多方面的证实,特别是在农产品领域应用非常广泛^[2~3]。本文针对小型西瓜果实品质及成熟度表征因子进行可见/近红外定量建模分析,主要探讨光谱信息采集位置和化学基础信息采集深度因素对建模效果的影响。光谱采集位置按瓜顶部位和赤道部位分别建立可溶性固形物含量、组织含水率以及果实授粉后生长期的相关性模型,分析光谱采集的最佳位置。信息深度主要分为浅层信息、中层信息和深层信息,针对瓜顶部位采集的光谱,建立每个深度层次对应的的可溶性固形物含量、组织含水率的定量模型,分析最佳的信息采集深度,为进一步的成熟度、品质分析研究奠定基础。

1 试验材料、方法及仪器

选择品种为京秀的小型西瓜果实 102 个,在室温条件下,测定其不同位置和深度的可溶性固形物含量及组织含水率等基础信息指标^[4]。其中,可溶性固形物含量测定采用日本爱宕公司生产的 ATAGO PAL-1 型数显折光仪;组织含水率的测定采用国标 GB/T 8858—1988《水果、蔬菜产品中干物质和水分含量的测定方法》^[5]。根据采集深度的不同,取样位置依次为:浅层,瓜皮白色不可食部位;中层,2~3 cm 深处瓜瓤;深层,瓜心处瓜瓤。为了保证数据的准确性,可溶性固形物含量取样为边长 1 cm 的立方体,而组织含水率采样为其周围位置取 5 g 左右。为了保证数据的准确性,瓜顶部位和赤道部位采用相同的信息采集方式。

光谱采集仪器为日本久保田株式会社生产的 Kubota K-BA100R 型便携式水果品质可见/近红外漫透射光谱分析仪^[6~7]。内置光源功率 50 W。

2 可见/近红外光谱的采集及预处理

2.1 光谱的采集

可见/近红外光谱采集时,将漫透射光纤头手动放置于西瓜样品待测部位,对准并且贴紧标定的采集区域进行采集,以防止漏光而造成不必要的干扰。仪器以内置分光作为背景,每 5 min 进行一次光谱自动修正,以消除环境温度变化等外部因素对采集

光谱的影响。样品积分时间与背景积分时间分别为 100 ms 和 50 ms,瓜顶部位采集 1 个点,共获得 102 条光谱,赤道部位均匀采集 4 个点,共获得 408 条光谱,分析时将 4 点进行平均处理。由于瓜顶和赤道位置光谱外观差异不大,故只列出瓜顶位置的原始光谱,如图 1 所示。

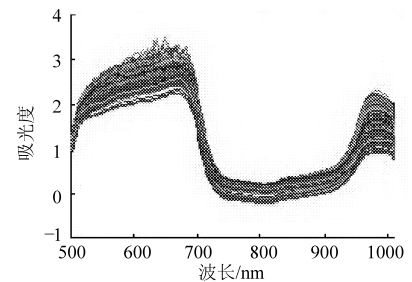


图 1 西瓜样品的漫透射原始光谱

Fig. 1 Original diffuse transmittance spectra of watermelon

2.2 可见/近红外光谱预处理

如图 1 所示,西瓜原始光谱的噪声较大,信噪比差,特别是光谱两端尤为明显。为了消除高频随机噪声,强化谱线特征,提高有用信息强度,需对光谱数据进行适当的预处理。因此,在剔除异常样品之后,采用 Matlab, Omnic 等数据分析软件对有效光谱进行了 SG (Savitzky-Golay) 卷积平滑及二阶导数处理^[8~9]。

光谱及化学异常值的剔除采用马氏距离、杠杆值与学生残差 t 检验准则相结合的方法。SG 平滑的窗口宽度为 7 点,计算次数为 3 次,结果如图 2 所示,经过平滑后光谱的噪声有了明显减弱,谱线得到改善。经二阶导数处理,可以消除光谱基线漂移,强化谱带特征,使信息差异得到一定的放大。

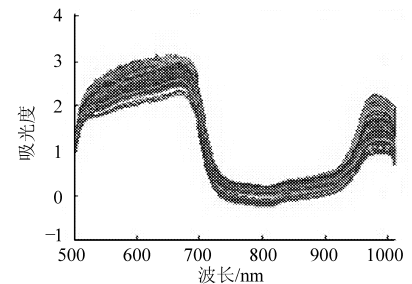


图 2 经 SG 平滑后西瓜样品的漫透射光谱

Fig. 2 Diffuse transmittance spectra of watermelon after SG smooth method

3 试验结果和讨论

3.1 可见/近红外光透过深度及中层信息位置确定

与苹果等成分分布相对均匀的样品种类不同,西瓜的内部结构更加复杂,为环形层级化的结构模式,而且随着生长期和成熟阶段的不同,成分分布还会发生很大的变化。

整体而言,西瓜果实的内部信息根据深度可分为浅层信息区、中层信息区和深层信息区。浅层区域为绿色硬质外皮以及白色不可食部位,即普通意义上的瓜皮;中层区域为白色不可食部位以内的红色可食部位直至瓜子带内侧;深层区域为西瓜中心半径约1~2 cm 的无子瓜心部位。如图3所示。

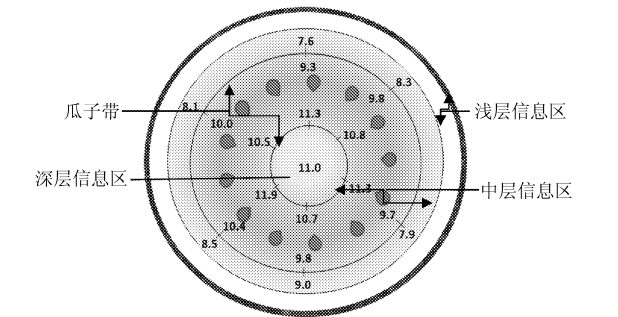


图3 西瓜切面糖度分布图
Fig.3 SSC section distribution of watermelon

由于浅层信息和深层信息分别为瓜皮和瓜心的信息,位置相对固定,易作区分,而中层信息包含范围较宽,且囊括了结构复杂的瓜子带(图3)。因此,为了确定在近红外分析检测时所定义的中层信息的化学基础数据采集位置,以及近红外光谱能够直接反映的西瓜内部品质的深度,对可见/近红外光能够透过西瓜果肉的深度进行了试验。依据光谱仪的光纤头大小在纵向长径为15cm 的目标样品瓜顶部位做圆形标记,固定光谱采集位置。将光纤头放置于此位置,由瓜柄部位开始,与西瓜纵径垂直,每1cm 切去一层西瓜,并且采集光谱,直至剩余西瓜厚度为2cm 左右停止(图4、图5)。

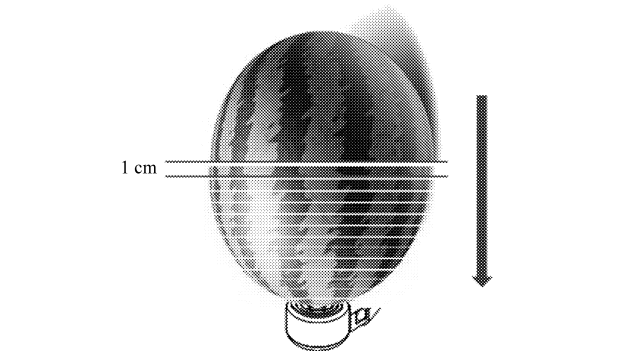


图4 可见/近红外光透过性测试图
Fig.4 Schematic of Vis/NIR permeability

如图5所示,随着目标西瓜厚度的减小,光谱的吸光度也相应减弱,在厚度为3cm 处时,光谱吸光度出现大幅度变化,特征差异明显,说明光已经透过了剩余西瓜。因此,可以认为近红外光透过品种为京秀的小型西瓜果实的深度至少为2~3 cm。图6为5个西瓜样品纵径上不同深度瓜瓤可溶性固形物含量平均值的变化趋势。自瓜顶开始,随着深度的

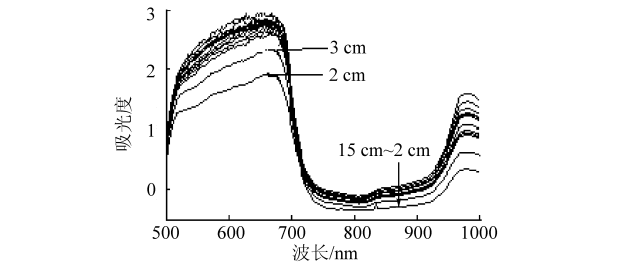


图5 可见/近红外光透过深度图
Fig.5 Depth of NIR permeability

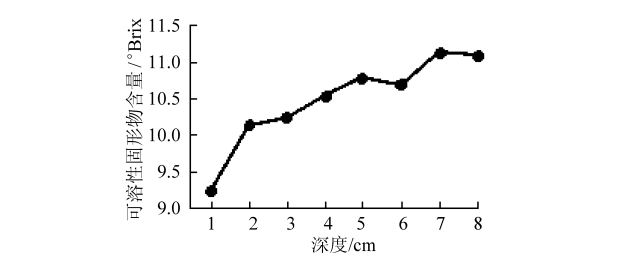


图6 瓜瓤可溶性固形物含量随深度变化趋势图
Fig.6 Changes of flesh SSC content with depth

增加,瓜瓤可溶性固形物含量(SSC)会有明显的变化,在2~6 cm 这段范围内,各处的SSC 比较接近,变化较小,相对稳定,而此范围正是前文所定义的中层信息区域,本区域包含了瓜子带,最能直观体现西瓜的实际成熟程度和品质状况。因此,选择2~3 cm 处为中层信息的化学标准值采集位置。

为了验证中层信息区域划分及明确化学信息采集位置的意义和必要性,以可溶性固形物含量为研究对象,对其进行建模分析比较,其中一组为深度2~3 cm 处取样,另一组为整个中层区域整体取样,结果如表1所示。通过以下比较可知,进行了化学信息采集位置确定之后样品的建模效果明显优于未区分的样品。

表1 深度划分结果比较
Tab.1 Results of depth division comparison

区域划分	光谱采集位置	基础信息采集深度	<i>R</i>	RMSEC /°Brix	RMSEP /°Brix
划分前	瓜顶	2~3 cm	0.952	0.530	0.666
	赤道	2~3 cm	0.929	0.518	0.632
划分后	瓜顶	未划分	0.922	0.621	0.668
	赤道	未划分	0.868	0.734	0.733

3.2 基础信息采集深度

近红外光谱包含了与水分、可溶性固形物相关的O—H、C—H等化学键的振动吸收,这就为近红外技术应用于西瓜果实组织含水率和可溶性固形物含量测定提供了化学依据^[10~11]。

田海清^[3]利用与本研究相似的漫透射数据采集方式,600 W 的光源强度,以可溶性固形物含量这一化学指标对品种为“早春红玉”的小型西瓜进行

了分析研究,认为光谱采集位置不会对其检测效果带来明显的影响,故本文针对西瓜的赤道部位和瓜顶部位分别进行研究探讨,以探究在不同深度采集化学基础信息情况下,近红外光谱技术检测小型西瓜品质的差异性,以选出最佳的基础信息采集位置。

分别将瓜顶、赤道部位经预处理后的光谱矩阵和对应的浅层 SSC、中层 SSC、深层 SSC、浅层含水率、中层含水率、深层含水率等化学值矩阵导入 TQ Analyst V 8.0 (Thermo Fisher Nicolet Co., 美国)数据分析软件进行偏最小二乘法(PLS)数学建模分析。建立模型时,依据相关系数法、移动窗口偏最小二乘法、遗传算法结合实际效果,选取 750 ~ 950 nm 范围内的漫透射光谱作为目标光谱。最佳因子数采用交叉验证法所得的 PRESS 值确定,当其

值最小且小于 10 时一般认为是最佳因子数。西瓜样品按照 3:1 的原则分为校正集和预测集。以相关系数 R 和校正均方根误差 RMSEC (Root mean square errors of calibration) 作为模型相关性及稳定性的评价指标,预测均方根误差 RMSEP (Root mean square error of prediction) 作为衡量模型预测准确性的标准。

表 2 为试验西瓜样品的化学基础数据统计结果。如表 2 所示,由于数学建模时需要按照随机方式对样品进行建模集和验证集的划分,会导致同一样品集的样品,赤道部位和瓜顶部位的数据统计有所差别。整体而言,建模集的各项指标分布范围较宽,并且完全涵盖了预测集样品的信息范围区域,这样更有利于提高模型的稳定性和可靠性。

表 2 西瓜样品的化学值分布统计
Tab.2 Statistics of parameter measurements of watermelon samples

指标	校正集			验证集		
	范围	均值	标准偏差	范围	均值	标准偏差
赤道浅层 SSC/°Brix	5.4 ~ 9.2	7.2	0.91	5.8 ~ 8.9	7.0	0.64
赤道中层 SSC/°Brix	7.4 ~ 13.0	10.4	1.35	7.7 ~ 13.0	11.0	1.34
赤道浅层含水率/%	89.8 ~ 94.3	92.6	0.91	89.8 ~ 94.2	92.6	0.78
赤道中层含水率/%	87.9 ~ 93.6	90.5	1.13	87.9 ~ 92.8	90.4	0.93
瓜顶浅层 SSC/°Brix	5.3 ~ 9.3	7.1	1.36	5.5 ~ 9.3	6.8	0.86
瓜顶中层 SSC/°Brix	7.8 ~ 12.1	10.2	1.79	7.0 ~ 11.5	10.3	1.27
瓜顶浅层含水率/%	90.0 ~ 94.8	92.8	0.99	90.8 ~ 94.0	92.3	1.08
瓜顶中层含水率/%	88.6 ~ 95.2	91.1	1.23	87.1 ~ 91.8	89.0	1.19
深层 SSC/°Brix	7.9 ~ 14.5	11.6	1.17	8.9 ~ 13.1	11.9	0.86
深层含水率/%	86.4 ~ 95.4	89.4	1.32	86.5 ~ 95.4	89.3	0.96

如表 3、4 所示,将瓜顶和赤道部位所得的光谱和对应的建模结果进行分析比较可知,采用偏最小二乘法(PLS)结合 SG 卷积平滑、二阶导数等预处理方法进行不同深度信息的定量建模分析,可以得到良好的效果。

通过比较发现,在瓜顶部位采集光谱,针对不同化学信息指标,在不同的信息深度区域,检测小型西瓜京秀果实品质的效果存在很大差异。就含水率而言,效果最好的信息区域为中层信息区 (R 为 0.953, RMSEP 为 0.510),浅层次之 (R 为 0.951, RMSEP 为 0.580),最内部的深层信息模型相关性及预测性最差 (R 为 0.882, RMSEP 为 0.980)。

对可溶性固形物含量而言,中层信息区为最佳区域 (R 为 0.952, RMSEP 为 0.666),浅层信息区次之 (R 为 0.934, RMSEP 为 0.669),深层区域效果最差 (R 为 0.811, RMSEP 为 0.802)。

通过以上对比分析可知,以瓜顶部位采集漫透射光谱,进行近红外检测的最佳的信息区域为中层信息区。

表 3 不同位置组织含水率建模结果
Tab.3 Modeling results of tissue water content at different locations

位置	信息深度	主因子数	R	RMSEC/%	RMSEP/%
赤道	浅层	6	0.902	0.470	0.604
	中层	8	0.924	0.556	0.766
	深层	8	0.906	0.645	0.966
瓜顶	浅层	8	0.951	0.365	0.580
	中层	7	0.953	0.454	0.510
	深层	6	0.882	0.748	0.980

表 4 不同位置可溶性固形物含量 (SSC) 建模结果
Tab.4 Modeling results of SSC at different locations

位置	信息深度	主因子数	R	RMSEC /°Brix	RMSEP /°Brix
赤道	浅层	8	0.903	0.518	0.668
	中层	9	0.929	0.518	0.732
	深层	9	0.900	0.536	0.698
瓜顶	浅层	7	0.934	0.482	0.669
	中层	7	0.952	0.530	0.666
	深层	6	0.811	0.662	0.802

而以赤道部位采集光谱,分别对各层次信息进行建模分析比较可以看出,在不同的采集深度,对基础数据的建模分析结果仍然存在着一一定的规律性。就含水率而言,中层信息效果最好(R 为0.924, RMSEP为0.766),其次是浅层信息(R 为0.902, RMSEP为0.604)和深层信息(R 为0.906, RMSEP为0.966)。虽然浅层信息模型的相关系数 R 略小于深层信息,但是其RMSEP则明显小于后者,说明在模型的预测性上,浅层信息要明显优于深层信息,故可认为浅层区域效果优于深层区域。

对可溶性固形物含量而言,中层信息效果最佳(R 为0.929, RMSEP为0.732),浅层次之(R 为0.903, RMSEP为0.668),深层较差(R 为0.900, RMSEP为0.698)。综上可知,无论是在瓜顶部位还是赤道部位采集近红外光谱,对组织含水率和可溶性固形物含量进行比较分析,都可证明中层信息区域是近红外检测小型西瓜品质的最佳基础信息采集区域。分析其原因可能有以下两点:

(1) 光谱信息量不同。由于采用漫透射的光谱采集方式,环形光源覆盖面积更大,且瓜顶部位瓜皮较薄,近红外光更易透过,采集的信息多为浅层和中层信息。如图5所示,可见/近红外光透过西瓜深度至少是2~3 cm,而京秀小型西瓜果实的绿色硬质瓜皮及不可食部位厚度平均为0.5 cm左右,与漫反射相比,漫透射采集方式穿透性更强,使得光谱的变化可以更多地表征西瓜果实的中层信息。

(2) 基础信息的差异性。基础信息的准确性及稳定性会对校正模型的效果产生很大的影响。西瓜果实的瓜皮表面细胞排列紧密,具有比较复杂的结构。最表层为硬的角质层和排列紧密的表皮细胞,下面是8~10层细胞的叶绿素带或无色细胞(外果皮),其内是由几层厚壁木质化的细胞组成的机械组织。再向内是中果皮,即习惯上所称的果皮,由肉质薄壁细胞组成,较紧实。因此,建立近红外模型时会造成所采集瓜皮可溶性固形物含量和含水率的化学基础信息与光谱信息不能较好对应的问题,导致模型的预测效果降低,不只是不同样品间瓜皮可溶性固形物含量和组织含水率会存在较大的差异,同一样品的不同部位差异性也很明显。而瓜瓢主要由大的薄壁细胞组成,细胞间隙大,其间充满汁液,各部位的含量相对较均匀,所以基础信息的采集也更加准确,建立的模型效果更好。

4 结论

(1) 在对小型西瓜进行内部品质分析比较时,应当对其内部结构进行分层处理,这样可以减小误差,提高无损检测模型的精确度和稳定性。

(2) 对于品种为京秀的小型西瓜而言,利用环形漫透射方式采集光谱时由中层信息建立模型的相关性及预测性最好,浅层区域次之,说明中层和浅层信息为漫透射可见/近红外光谱预测的主要信息,中层区域为最佳基础信息采集区域。

参 考 文 献

- 1 陆婉珍. 现代近红外光谱分析技术[M]. 北京:中国石化出版社,2000:129~130.
- 2 韩东海,刘新鑫,鲁超. 苹果内部褐变的无损检测研究[J]. 农业机械学报,2006,37(6):86~88.
Han Donghai, Liu Xinxin, Lu Chao. Study on optical-nondestructive detection of breakdown apples[J]. Transactions of the Chinese Society for Agricultural Machinery, 2006,37(6):86~88. (in Chinese)
- 3 田海清,应文斌,徐惠荣,等. 西瓜可溶性固形物含量近红外透射检测技术[J]. 农业机械学报,2007,38(5):111~113.
Tian Haiqing, Ying Wubin, Xu Huirong, et al. Near-infrared transmittance measuring technique for soluble solids content of watermelon[J]. Transactions of the Chinese Society for Agricultural Machinery, 2007,38(5):111~113. (in Chinese)
- 4 曹建康,姜微波,赵玉梅. 果蔬采后生理生化实验指导[M]. 北京:中国轻工业出版社,2007:9.
- 5 宁正祥. 食品成分分析手册[M]. 北京:中国轻工业出版社,1998.
- 6 郑咏梅,张军,陈星旦,等. 短波近红外光谱的整粒小麦蛋白质 PLS 方法的定量分析[J]. 光谱学与光谱分析,2004,24(9):1047~1049.
Zheng Yongmei, Zhang Jun, Chen Xingdan, et al. The PLS analysis with short wave near-infrared of whole wheat for protein[J]. Spectroscopy and Spectral Analysis, 2004, 24(9):1047~1049. (in Chinese)
- 7 王桂芳,王彬,陈庆先,等. 短波近红外漫反射光谱法对氨基西林粉末药品的非破坏定量分析[J]. 生命科学仪器,2007,5(12):37~39.
- 8 逯家辉,滕利荣,蒋富明,等. 短波近红外光谱法分析酒中乙醇含量[J]. 吉林大学学报:理学版,2003,41(2):245~247.
Lu Jiahui, Teng Lirong, Jiang Fuming, et al. Rapid determination of ethanol in wine by short-wavelength near-infrared spectroscopy[J]. Journal of Jilin University: Science Edition, 2003,41(2):245~247. (in Chinese)
- 9 饶震红,李民赞,吉海彦. 短波近红外光谱法同时定量分析水溶液中葡萄糖、果糖和蔗糖[J]. 光谱学与光谱分析,2006,26(4):633~635.
Rao Zhenhong, Li Minzan, Ji Haiyan. Simultaneous determination of glucose, fructose and sucrose in aqueous solution by short wavelength near infrared spectroscopy[J]. Spectroscopy and Spectral Analysis, 2006, 26(4):633~635. (in Chinese)
- 10 李民赞,韩东海,王秀. 光谱分析技术及其应用[M]. 北京:科学出版社,2008:7.
- 11 张楠,程玉,李东华,等. 近红外透射光谱测定水晶梨糖度的初步研究[J]. 分析检测,2007,28(3):215~217.