

文章编号: 0253-2239(2008)03-0478-04

# 可见/近红外光谱法无损检测赣南脐橙可溶性固形物

刘燕德<sup>1</sup> 陈兴苗<sup>1</sup> 欧阳爱国<sup>2</sup>

(<sup>1</sup> 江西农业大学工学院, 江西 南昌 330045; <sup>2</sup> 江西蓝天学院汽车系, 江西 南昌 330098)

**摘要** 应用可见/近红外光谱法对赣南脐橙可溶性固形物进行了无损检测研究。通过主成分分析, 获取光谱的有效信息, 将其作为人工神经网络的输入变量进行非线性建模。90个建模样品训练结果是, 样品参考值与预测值之间的相关系数为0.9147, 训练均方差为0.5203; 38个未知样品预测结果是: 样品参考值与预测值之间的相关系数为0.9033, 预测均方差为0.6964, 相对预测偏差4.5709%。实验结果表明基于人工神经网络的可见/近红外光谱法无损检测赣南脐橙可溶性固形物是可行的。

**关键词** 医用光学与生物技术; 可见/近红外光谱; 无损检测; 人工神经网络; 主成分分析; 可溶性固形物; 赣南脐橙  
**中图分类号** O657.33; TN247 **文献标识码**: A

## Non-Destructive Measurement of Soluble Solid Content in Gannan Navel Oranges by Visible/Near-Infrared Spectroscopy

Liu Yande<sup>1</sup> Chen Xingmiao<sup>1</sup> Ouyang Aiguo<sup>2</sup>

(<sup>1</sup> *Engineering College, Jiangxi Agriculture University, Nanchang, Jiangxi 330045, China*)  
(<sup>2</sup> *Department of Automobile, Jiangxi Blue Sky College, Nanchang, Jiangxi 330098, China*)

**Abstract** Non-destructive measurement of soluble solid content in Gannan navel oranges was carried out by visible/near-infrared spectroscopy detection method. Effective information of spectra was obtained by principal component analysis, and was used as the input variables of artificial neural network for building the nonlinear model. The results, based on calibration for 90 samples, are 0.9147 and 0.5203 for calibration correlation coefficient and root mean square error of calibration. The results, based on prediction for 38 unknown samples, are 0.9033, 0.6964 and 4.5709% for prediction correlation coefficient, root mean square error of prediction, and relative standard deviation (RSD), respectively. Experimental results show that visible/near-infrared spectroscopy detection method, based on artificial neural network, for non-destructive measurement of soluble solid content in Gannan navel oranges is feasible.

**Key words** medical optics and biotechnology; visible/near-infrared spectroscopy; non-destructive measurement; artificial neural network; principal component analysis; soluble solid content; Gannan navel oranges

## 1 引 言

近红外光谱所含信息极其丰富, 主要是有机分子(含氢基团)的倍频与合频吸收光谱, 与中红外光谱一样, 能得到分子的结构、组成、状态的信息。而且从近红外反射光谱还能得到样品的密度、粒度、高分子物的聚合度及纤维的直径等物质的物理状态信息<sup>[1,2]</sup>。近年来, 国内外学者利用可见/近红外光谱检测技术对水果品质检测做了大量的工作<sup>[3~9]</sup>。

对多组份完整水果的可见/近红外光谱分析发现, 光谱参量与样品化学值含量之间具有一定的非线性关系, 特别是当样品的化学值含量范围较大时, 其非线性较明显。人工神经网络(Artificial neural network, ANN)可以模仿人脑处理信息过程, 具有分布并行处理、非线性映射、自适应学习和稳健容错等特性, 因此被广泛用于光谱检测中的非线性模型<sup>[10~14]</sup>。误差反向传播学习算法(Back

收稿日期: 2007-06-15; 收到修改稿日期: 2007-10-31

基金项目: 国家自然科学基金(60468002, 30560064)、教育部新世纪优秀人才资助计划(NCET-06-0575)和江西省青年科学家(井冈之星)培养对象资助课题。

作者简介: 刘燕德(1967—), 女, 江西泰和人, 教授, 博士生导师, 主要从事生物光电及应用等方面的研究。

E-mail: jxliuyd@163.com

propagation, BP)是一种有教师的学习算法,以样本的输出与期望输出的偏差极小为目标进行学习,利用梯度最速下降法,使权值误差沿误差函数的负梯度方向改变。学习的实质是不断修改权值以使误差趋于零。主成分分析-反向传输人工神经网络(PCA-BP)方法是将样品光谱用主成分算法压缩为几个变量,将其作为人工神经网络的输入结点。

本文采用厚皮水果赣南脐橙为研究对象,应用PCA-BP方法对其内部品质可溶性固形物进行无损检测实验研究。

## 2 材料与方法

实验选用美国 ASD 公司的 QualitySpec<sup>®</sup> Pro 光谱仪,其波长范围为 350~1800 nm,光谱采样间隔为 1 nm,扫描时间为 100 ms/次,扫描次数为 10 次,检测距离为 15 mm,光源为 12 V/30 W 钨卤灯。图 1 为本实验检测系统。采用漫反射方式,每个样品在赤道部位等间隔采集三次光谱,取平均值。

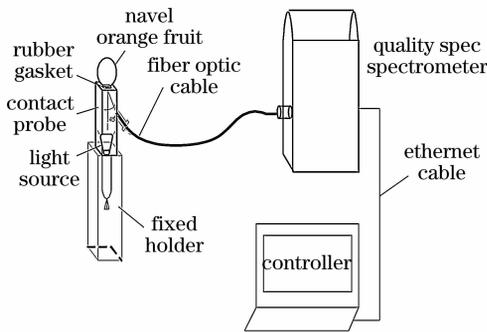


图 1 脐橙可见光/近红外光谱检测系统

Fig. 1 Schematic diagram of the setup for Vis/NIR measurement of navel orange fruit

实验样品来自江西省信丰县的三种不同品种(纽贺尔、奈佛宁娜、朋娜)。选择代表性样品,对 90 个建模样品进行训练,38 个未知样品进行外部验证。样品的参考化学值采用折射式数字糖度计(PR-101 $\alpha$ ,日本)进行测量。

光谱数据由美国 ASD 公司的 Indico v4.0 软件采集和转换,采用标准归一化方法(SNV)对光谱数据进行预处理,以消除固体颗粒大小、表面散射及光程变化对漫反射光谱的影响<sup>[15]</sup>,并且符合神经网络中输入值结点所要求的范围<sup>[16]</sup>。PCA-BP 模型由江苏大学和南京林业大学联合研制的 NIRSA V2.2 软件进行训练,其结构由输入层、隐含层和输出层组成。输入层结点数由主成分分析决定,隐含层结点数由实验训练获得,输出层结点数 1(可溶性固形

物化学值)。影响模型预测性能的目标误差、学习速率和动量因子也由实验训练获得。

数学模型采用外部验证对其性能进行评价,由相关系数( $r_p$ )、预测均方差(RMSEP)和相对预测偏差(RSD)评价。

## 3 结果与讨论

### 3.1 可见/近红外光谱分析

参量设置对实验结果有一定的影响,因此光谱采集必须保证在相同的实验条件下进行。图 2 是三种赣南脐橙品种的吸收光谱图。虽然不同品种间由于其内部组织的差异性,光谱图稍有变化,但其光谱形状相似,差异性不明显,因此可以将三品种样品放在一起进行建模分析。

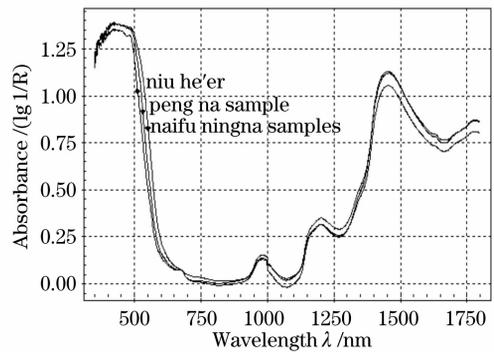


图 2 三种赣南脐橙品种的可见光/近红外光谱

Fig. 2 Vis/NIR spectra of three different kinds of Gannan navel oranges

### 3.2 光谱主成分分析

对 90 个训练样品的光谱进行主成分(PC)分析,取累计贡献率达到 99% 时的主成分数为 PCA-BP 模型的输出变量。图 3 是前两个主成分得分图(A1:纽贺尔样品、A2:奈佛宁娜样品、A3:朋娜样

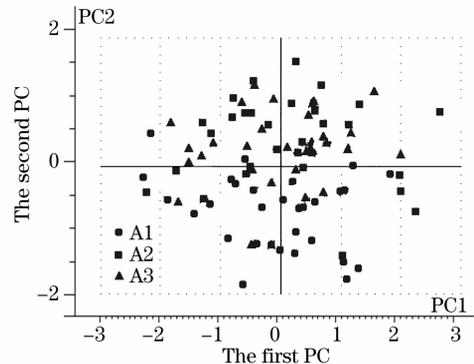


图 3 赣南脐橙前两个(PC1, PC2)主成分得分图

Fig. 3 Plot of the first two principal components for navel oranges

品),从图 3可知训练样品分布集中,不存在异常点。

表 1 是前 7 个主成分的特征值及累计贡献率。

表 1 前 7 个主成分的特征值及累计贡献率

Table 1 Eigenvalue and cumulative contribution rate of the first seven PCs

No. of PC	Eigenvalue	Variance /%	Cumulative contribution rate /%
PC1	710.61	48.97	48.97
PC2	382.42	26.36	75.33
PC3	152.49	10.51	85.84
PC4	104.41	7.20	93.04
PC5	50.71	3.49	96.53
PC6	26.93	1.86	98.39
PC7	9.42	0.65	99.04

### 3.3 PCA-BP 校正数学模型建立

通过主成分分析,前面 7 个主成分已经包含了光谱 99.04% 的有效信息。因此,将这 7 个特征变量作为 PCA-BP 的输入变量,即输入层结点数为 7,经多次训练确定隐含层结点数为 11,输出层结点数为 1。以预测均方差最小为依据,经反复训练确定目标误差、学习速率和动量因子分别为 0.004,0.05 和 0.70。90 个建模样品的训练结果为:参考值与预测值之间的相关系数为 0.9147,训练均方差为 0.5203。

表 2 应用 PCA-BP 模型对 38 个未知样品的预测结果

Table 2 Prediction results for 38 unknown samples by PCA-BP model

Sample number	Predicted value	Reference value	Bias	RSD /%
01-A1	13.5763	12.60	0.9763	7.7484
02-A1	14.5545	14.40	0.1545	1.0729
03-A1	13.7515	12.60	1.1515	9.1389
04-A1	12.8409	12.30	0.5409	4.3976
05-A1	13.4555	12.30	1.1555	9.3943
06-A1	13.1131	14.00	-0.8869	6.3350
07-A1	13.1193	12.40	0.7193	5.8008
08-A1	13.4823	13.50	-0.0177	0.1311
09-A1	12.7058	12.90	-0.1942	1.5054
10-A1	13.7310	14.70	-0.9690	6.5918
11-A1	12.6052	12.30	0.3052	2.4813
12-A2	15.6767	15.90	-0.2233	1.4044
13-A2	15.2621	15.60	-0.3379	2.1660
14-A2	13.7387	13.30	0.4387	3.2985
15-A2	14.2574	14.10	0.1574	1.1163
16-A2	14.4183	13.80	0.6183	4.4804
17-A2	13.4473	12.50	0.9473	7.5784
18-A2	15.5217	16.30	-0.7783	4.7748
19-A2	14.3573	15.30	-0.9427	6.1614
20-A2	15.2822	15.70	-0.4178	2.6611

(续表)

Sample number	Predicted value	Reference value	Bias	RSD /%
21-A2	14.7505	13.60	1.1505	8.4596
22-A2	14.5119	14.40	0.1119	0.7771
23-A2	15.0076	14.80	0.2076	1.4027
24-A2	15.1288	14.80	0.3288	2.2216
25-A2	13.7056	12.50	1.2056	9.6448
26-A2	13.3908	13.80	-0.4092	2.9652
27-A2	13.8298	14.00	-0.1702	1.2157
28-A3	10.7110	10.20	0.5110	5.0098
29-A3	11.7690	12.00	-0.2310	1.9250
30-A3	11.8702	12.70	-0.8298	6.5339
31-A3	12.0377	12.00	0.0377	0.3142
32-A3	12.8623	11.40	1.4623	12.8272
33-A3	11.7409	11.20	0.5409	4.8295
34-A3	11.5405	11.30	0.2405	2.1283
35-A3	11.7341	12.60	-0.8659	6.8722
36-A3	11.1588	10.30	0.8588	8.3379
37-A3	12.2167	11.80	0.4167	3.5314
38-A3	11.7105	11.00	0.7105	6.4591

Note: A1: Niuhe'er samples; A2: Naifu ningna samples; A3: Pengna samples

### 3.4 PCA-BP 数学模型可靠性评价

用已建好的 PCA-BP 模型对 38 个未知样品进行预测分析,预测结果见表 2。预测样品中最大偏差及相对预测偏差为 1.4623 和 12.8272%。预测样品参考值与预测值之间的相关系数为 0.9033,预测均方差为 0.6964,相对预测偏差 4.5709%。由此可见所建的 PCA-BP 模型具有较好的预测效果。

## 4 结 论

本文对赣南脐橙样品光谱进行了主成分分析,提取光谱中的有效信息作为 PCA-BP 数学模型的输入结点,使网络训练迭代时间显著缩短,同时防止过拟合现象的产生。在 PCA-BP 进行训练过程中,隐含层结点数、目标误差、学习速率和动量因子等影响因素必须通过反复训练对比,获取较优参量以使预测性能达到最优。实验结果表明应用可见/近红外光谱检测技术对赣南脐橙可溶性固形物的无损检测是可行的。

### 参 考 文 献

- 1 Lu Wanzhen, Yuan Hongfu, Xu Guangtong *et al.*. *The Modern Analysis Technique for Near-infrared Spectra* [M]. Beijing: China Petrochemical Press, 2000. 19~26
- 2 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. 19~26
- 3 Yan Yanlu, Zhao Longlian, Han Donghai *et al.*. *The Foundation*

- and Application of Near-Infrared Spectroscopy Analysis* [M]. Beijing: China Light Industry Press, 2005. 8~9
- 严衍祿, 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. 8~9
- 3 V. Andraw McGlone, Robert B. Jordan, Paul J. Martinsen. Vis/NIR estimation at harvest of pre- and post-storage quality indices for 'Royal Gala' apple[J]. *Postharvest Biology and Technology*, 2002, **25**(2): 135~144
- 4 Yande Liu, Yibin Ying. Use of FT-NIR spectrometry in non-invasive measurement of internal quality of 'Fuji' apples[J]. *Postharvest Biology and Technology*, 2005, **37**(1): 65~71
- 5 Zhao Jiewen, Zhang Haidong, Liu Muhua. Preprocessing methods of near-infrared spectra for simplifying prediction model of sugar content of apples[J]. *Acta Optica Sinica*, 2006, **26**(1): 136~140
- 赵杰文, 张海东, 刘木华. 简化苹果糖度预测模型的近红外光谱预处理方法[J]. *光学学报*, 2006, **26**(1): 136~140
- 6 Y. D. Liu, Y. B. Ying. Rapid determination of maturity in apple using outlier detection and calibration model optimization [J]. *Trans. of ASAE*, 2006, **49**(1): 91~95
- 7 Yongni Shao, Yong He. Nondestructive measurement of the internal quality of bayberry juice using Vis/NIR spectroscopy[J]. *J. Food Engineering*, 2007, **79**(3): 1015~1019
- 8 Liu Yande, Luo Ji, Ouyang Aiguo. Visible transmittance techniques for rapid analysis of sugar content of fresh navel orange juices[J]. *Spectroscopy and Spectral Analysis*, 2007, **27**(3): 569~572
- 刘燕德, 罗吉, 欧阳爱国. 可见光光谱检测赣南脐橙糖度的研究[J]. *光谱学与光谱分析*, 2007, **27**(3): 569~572
- 9 Zou Xiaobo, Zhao Jiewen. Methods of characteristic wavelength region and wavelength selection based on genetic algorithm[J]. *Acta Optica Sinica*, 2007, **27**(7): 1316~1321
- 邹小波, 赵杰文. 用遗传算法快速提取近红外光谱特征区域和特征波长[J]. *光学学报*, 2007, **27**(7): 1316~1321
- 10 Yixin Chen, Shilpa S. Thosar, Reba A. Forbess *et al.*. Prediction of drug content and hardness of intact tablets using artificial neural network and near-infrared spectroscopy[J]. *Drug Development and Industrial Pharmacy*, 2001, **27**(7): 623~631
- 11 Yongni Shao, Yong He, Yanyan Wang. A new approach to discriminate varieties of tobacco using vis/near infrared spectra [J]. *Eur. Food Res. Technol.*, 2007, **224**(5): 591~596
- 12 Yiqun Huang, Lars J. Kangas, Barbara A. Rasco. Applications of artificial neural networks (ANNs) in food science[J]. *Critical Reviews in Food Science and Nutrition*, 2007, **47**(2): 113~126
- 13 Qi Xiaoming, Zhang Luda, Du Xiaolin *et al.*. Quantitative analysis using NIR by building PLS-BP model[J]. *Spectroscopy and Spectral Analysis*, 2003, **23**(5): 870~872
- 齐小明, 张录达, 杜晓林 等. PLS-BP 法近红外光谱定量分析研究 [J]. *光谱学与光谱分析*, 2003, **23**(5): 870~872
- 14 He Yong, Li Xiaoli, Shao Yongni. Discrimination of varieties of apple using near infrared spectra based on principal component analysis and artificial neural network model[J]. *Spectroscopy and Spectral Analysis*, 2006, **26**(5): 850~853
- 何勇, 李晓丽, 邵咏妮. 基于主成分分析和神经网络的近红外光谱苹果品种的鉴别方法研究[J]. *光谱学与光谱分析*, 2006, **26**(5): 850~853
- 15 Zhu Xiaoli, Yuan Hongfu, Lu Wanzhen. Progress and application of spectral data pretreatment and wavelength selection methods in NIR analytical technique[J]. *Progress in Chemistry*, 2004, **16**(4): 528~542
- 褚小力, 袁洪福, 陆婉珍. 近红外分析中光谱预处理及波长选择方法进展与应用[J]. *化学进展*, 2004, **16**(4): 528~542
- 16 Xu Lu, Shao Xueguang. *Methods of Chemometrics* [M]. Beijing: Science Press, 2004. 284~285
- 许祿, 邵学广. *化学计量学方法* [M]. 北京: 科学出版社, 2004. 284~285