

# 近红外光谱技术结合神经网络检测苹果咀嚼性 Detecting chewiness of apple by near infrared spectroscopy technology combined artificial neural network

曾祥媛 赵武奇 乔瑶瑶 尹艺然

ZENG Xiang-yuan ZHAO Wu-qi QIAO Yao-yao YIN Yi-ran

裴亚利 霍瑶瑶 郭玉蓉

PEI Ya-li HUO Yao-yao GUO Yu-rong

(陕西师范大学食品工程与营养科学学院, 陕西 西安 710119)

(Department of Food Engineering and Nutrition Science, Shaanxi Normal University, Xi'an, Shaanxi 710119, China)

**摘要:**以不同储藏期的 135 个苹果样品为研究对象,对其近红外光谱数据进行预处理并通过主成分分析法提取光谱特征,采用人工神经网络技术建立苹果咀嚼性的近红外光谱检测模型。结果表明,对苹果光谱咀嚼性的最佳光谱预处理方法是加权多元散射处理(WMSC)的光谱散射处理方法和“2441”的数学处理方法,通过主成分分析法提取 3 个主成分作为原始信息的特征变量,建立苹果咀嚼性检测的人工神经网络模型结构为 3—16—1,模型对验证集预测的决定系数为 0.992 4,均方根误差为 0.000 108 2。近红外光谱技术能对苹果咀嚼性进行快速、无损预测。

**关键词:**苹果;无损检测;质构;咀嚼性;近红外光谱;神经网络

**Abstract:** Near infrared spectroscopy data from 135 apple samples of different storage periods were determined, the character of them was extracted and analyzed using principal components analysis. Therefore an ANN model for detection of apple chewiness was established. Our results showed that the preprocessing of spectrum scattering was the weighted multiple scatter correction(WMSC) and mathematics processing was “2441”. The structure of the artificial neural network mode was 3—16—1, established after extracting 3 principle component as the characteristic variables of the original information. The decision coefficient of our model on validation is 0.992 4, and the root mean square error is 0.000 108 2. Our results confirmed that the near infrared spectroscopy technology can use to detect the chewiness of apple rapidly, without forecast destructive.

**基金项目:**农业部现代农业产业技术体系项目(编号:CARS-28)

**作者简介:**曾祥媛,女,陕西师范大学在读本科生。

**通讯作者:**赵武奇(1965—),男,陕西师范大学副教授,博士。

E-mail:zwq65@163.com

**收稿日期:**2016-01-09

**Keywords:** apple; nondestructive testing; texture; chewiness; near infrared spectroscopy; artificial neural network

苹果特性的判断和检测一直是企业和科研工作者关注的重要课题,其质构特性直接影响到苹果的品质和人们对苹果的接受程度<sup>[1]</sup>。咀嚼性是反映苹果质地特性的主要参数之一<sup>[2]</sup>,研究<sup>[3]</sup>表明,咀嚼性能反映牙齿咀嚼样品时所需要的能量,表示果肉在牙齿咀嚼的抵抗能力。压力硬度计和质构分析仪检测是目前苹果咀嚼性检测的常用方法,但有检测速度慢、检测后苹果失去商品价值等缺点<sup>[4]</sup>。因此,研究一种快速、可靠、简便的质构无损检测方法十分重要。

近红外光谱检测技术<sup>[5]</sup>是一种高效、无损、快速、绿色、准确的无损检测方法,已经较为广泛地应用在食品<sup>[6]</sup>、农业、制药和烟草等行业中。张静等<sup>[7]</sup>在 600~1 100 nm 的波长范围内利用近红外透射光谱法对苹果糖度进行检测建模,方法可行,具有简便、快速、非破坏和相对准确的优点,适用于对大批苹果进行分级处理;李桂峰等<sup>[8]</sup>在 1 408~2 355 nm 的波长范围内,对富士苹果硬度进行近红外光谱无损检测研究,所建模型有很好的预测效果,能满足生产中苹果硬度检测精度的要求,具有实际应用价值;Lammertyn 等<sup>[9]</sup>利用 380~1 650 nm 波长范围的近红外光谱对乔纳金苹果的硬度进行无损检测,建立了偏最小二乘法检测苹果硬度的模型;孙炳新等<sup>[10]</sup>对苹果脆度进行近红外无损检测,建立了稳定的苹果脆度模型;陆辉山等<sup>[11]</sup>采取偏最小二乘回归(PLSR)算法结合不同光谱预处理方法建立苹果内部的可溶性固形物含量(SSC)的定量模型,试验结果说明近红外光谱漫透射技术能够快速、无损地检测出苹果的可溶性固形物含量。

目前,有关苹果咀嚼性的研究尚未见报道。人工神经网络

络(ANN)技术以其具有辨识和逼近任意复杂非线性系统的能力,在农产品加工时间选择、水果品质预测、检测分类、危害分析和临界点控制中关键点的判断,以及在模式识别等研究中应用广泛<sup>[12]</sup>。因此,本试验拟在对苹果近红外光谱数据进行预处理,通过主成分分析法提取光谱特征的基础上,建立人工神经网络模型,以实现利用近红外光谱技术快速无损地检测苹果的咀嚼性质构特性,为研究苹果的品质评价方法提供参考依据。

## 1 材料与方法

### 1.1 材料与仪器

红富士苹果采自陕西乾县梁山官地村果园,采摘原则是大小颜色均匀、无病虫害、无机械伤,采摘后将苹果套网袋装入纸箱内,迅速运回实验室放入冷库中贮藏。每次试验前取出样品,置于常温下(25℃左右)2 h后进行测量。

多功能近红外分析仪:InfraXact型,瑞典福斯公司;

质构仪(物性测试仪):TA.XT.Plus型,英国 stable micro system 公司。

### 1.2 试验方法

1.2.1 苹果咀嚼性的测定 用 TA-XTplus 质构仪对苹果进行 TPA 质地测定。将苹果去皮切片,沿横向切成直径 5 cm 厚度 1.5 cm 的柱状果片,测试时选用 P50 探头,参数设置为:测前速率 2 mm/s,测试速率 1 mm/s,测后速率 1 mm/s,压缩程度 5 mm,停留间隔 5 s,数据采集速率 400 pps,触发值 5 g<sup>[13-14]</sup>。每个苹果样品重复 2 次,压下的位置与扫描近红外光谱的采样点相对应<sup>[15]</sup>,测定后保存所得质构参数的数据。

1.2.2 光谱的采集 InfraXact 多功能近红外分析仪对苹果进行扫描,试验间隔为 7 d,每次样品个数为 5 个苹果,试验时间从 2014 年 10 月份苹果采摘开始,到 2015 年 4 月下旬结束,共 135 个苹果样本。扫描方式是在苹果的赤道面上选取对应两点分别进行扫描,每个点扫描两次,采点间隔为 2 nm,波长范围为 570~1 850 nm。将采集的光谱按  $\lg[1/R]$  转换为吸光度, $R$  为反射光强与入射光强的比值,即样本的反射率,获得苹果样品光谱数据并存储。

1.2.3 异常光谱的剔除与样本集的划分 利用 WinISI III 分析软件,对原始光谱进行处理,计算马氏距离,马氏距离大于 3.0 即视为是异常光谱被剔除。样本集的划分是将剔除异常后的 130 个光谱数据升序排列,根据每四取一的原则,将样本分成 98 个训练集和 32 个验证集。

1.2.4 光谱预处理方法的确定 采用数学处理和光谱散射处理分别对苹果原始光谱数据进行处理,建立局部最小二乘法(MPLS)回归模型,根据校正标准误差(SEC)和交叉验证决定系数( $R_{cv}^2$ )的大小来确定最优预处理方法。光谱散射处理方法包括无散射预处理(None)、标准化处理(SNV)、去趋势化处理(Detrend)、标准化+去趋势化处理(SNV+Detrend)、标准化多元散射处理(SMSC)、加权多元散射处理(WMSC)、反向多元散射处理(IMSC),采用的数学处理方式有“0011”、“1441”、“2441”(4 位数字分别代表导数处理的阶

数、导数处理的数据点隔数、一次平滑处理点隔数、二次平滑处理点隔数)3 种方法。SEC 表示定标模型对训练集样品预测值与化学分析数值的标准差; $R_{cv}^2$  表示模型对训练集的交叉验证决定系数。进行光谱预处理时,SEC 越小, $R_{cv}^2$  越大,效果越好。

1.2.5 光谱特征的提取 按 1.2.4 确定的方法对苹果近红外光谱数据进行预处理后,苹果样品的近红外光谱数据中有 633 个吸光度,存在信息的重叠,若以此作为变量建模,具有训练速度缓慢、效率低等缺点<sup>[16]</sup>。为简化数据集,采用主成分分析,提取光谱特征。

1.2.6 神经网络模型的建立 以提取的主成分作为神经网络的输入,以苹果咀嚼性为输出;以 TRAINRP 函数为传递函数,以 LEARNGDM 为权值和阈值的 BP 学习算法,对训练集进行训练,以网络性能函数 MSE(神经网络的均方误差)值为指标,研究不同隐含层数的建模效果,以确定神经网络的最佳隐含层数,建立近红外光谱技术结合神经网络检测苹果咀嚼性的神经网络模型。

### 1.3 模型的验证

用 1.2.6 建立的神经网络模型对验证集进行预测,以预测值和实测值的预测决定系数( $R^2$ )来评价模型的准确性。

## 2 结果与分析

### 2.1 苹果咀嚼性的数据范围统计

样品评价参数覆盖的范围在很大程度上决定建模的质量。试验从 2014 年 10 月份苹果采摘开始,到 2015 年 4 月下旬结束。在试验过程中,苹果的质地特性会发生变化,因此试验选择的样品覆盖了样品质地变化的整个过程,得到的数据丰富且跨度广,建立的模型具有广泛适用性。表 1 是试验苹果咀嚼性的数据范围统计。通过观察可以得到,苹果咀嚼性参数的数据范围广,代表性比较强。

### 2.2 剔除异常光谱与样本集划分

图 1 是原始光谱谱线图像,图 2 是剔除异常处理后的图像,计算马氏距离得出原始光谱中有 5 个明显的异常光谱。

表 1 苹果咀嚼性分析结果

咀嚼性/N	4 以下	4~8	8~12	12~16	16 以上
个数/个	22	28	38	35	12

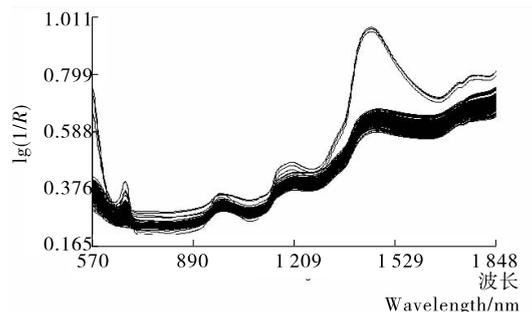


图 1 苹果原始光谱图

Figure 1 Apple's original spectra

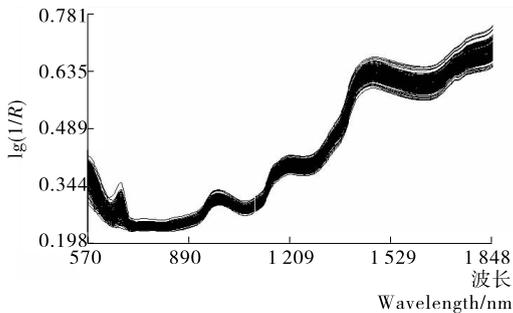


图 2 剔除异常后的光谱图

Figure 2 Spectrum after eliminating abnormal

对比观察,剔除异常后光谱轮廓变化更加明显,谱图对称的分布在基线周围,减缓了背景干扰,走向趋势统一,比原光谱有更高的灵敏度。

表 2 是样本集划分结果,从表中可看出,训练集包括整个数据范围的最大值和最小值,验证集咀嚼性数据的平均值及标准差与训练集相似,且验证集包含在训练集内,表明验证集较合理,可用于训练集所建模型的验证。

表 2 样本集划分结果

Table 2 Sample set classification results

样本	样品数	最小值	最大值	平均值	标准差
训练集	98	0.000 4	22.955 6	9.678 6	5.407 5
验证集	32	0.005 3	21.084 0	9.938 7	5.146 6

### 2.3 光谱预处理结果

仪器采集的原始光谱中包含与样品结构组成有关的信息,此外还受测试条件、外界温度、仪器状态等因素的影响<sup>[10]</sup>。因此通常采用数学处理和光谱散射处理这两类方法来消除杂音,可提升光谱的精确性<sup>[17]</sup>。表 3 是不同光谱预处理方法的建模效果。由表 3 可知,对苹果光谱数据的最佳预处理方法为:散射处理 Weighted MSC、数学处理 2441(即进行 2 阶导数处理,导数处理间隔点数为 4,一次平滑处理的间隔点数为 4,不进行二次平滑处理),该预处理效果最好,其  $R_{cv}^2$  的值最大,SEC 的值较小。

### 2.4 主成分分析

主成分分析结果见图 3。图 3 中的 3 个主成分对原始信息的累计贡献率达到了 95% 以上,包含了绝大部分的光谱数据信息,因此选 3 个主成分作为原始信息的特征变量,作为神经网络的输入。

### 2.5 隐含层数的选择

在神经网络中,隐含层数的选择对神经网络效果有很大影响。若隐含层数较少,可能会使局部极小值过多,不能达到所要的效果;若隐含层数较多,使拟合函数太难,易发生过拟合现象,影响验证效果。表 4 为不同隐含层数时的训练结果,可以看出,隐含层数为 16 时,均方差 MSE 为最小值

表 4 不同隐含层数时的训练结果

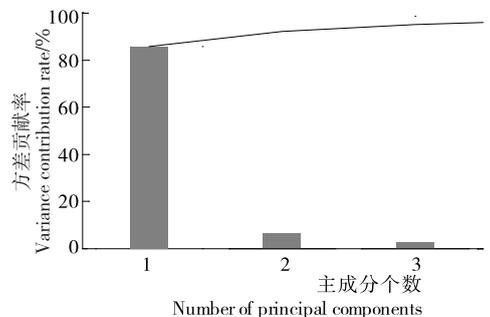
Table 4 The training of the different hidden layer

隐含层数	11	12	13	14	15	16	17	18
均方差 MSE	0.041 709	0.040 909	0.030 582	0.025 134	0.021 565	0.017 707	0.035 732	0.036 273

表 3 不同散射校正和数学处理方法的效果

Table 3 Effect of different scattering correction and mathematical processing method

数学处理方法	光谱处理方法	交叉验证决定系数 $R_{cv}^2$	校正标准误差 SEC
0011	None	0.880 0	0.507 0
	SNV and Detrend	0.220 0	0.502 2
	SNV only	0.400 0	0.516 6
	Detrend only	0.560 0	0.510 3
	Standard MSC	0.880 0	0.482 1
	Weighted MSC	0.690 0	0.473 4
	Inverse MSC	0.680 0	0.501 1
1441	None	0.490 0	0.506 1
	SNV and Detrend	0.640 0	0.497 9
	SNV only	0.710 0	0.494 6
	Detrend only	0.880 0	0.509 3
	Standard MSC	0.690 0	0.484 1
	Weighted MSC	0.510 0	0.508 9
	Inverse MSC	0.530 0	0.492 4
2441	None	0.960 0	0.478 7
	SNV and Detrend	0.590 0	0.489 2
	SNV only	0.610 0	0.489 2
	Detrend only	0.880 0	0.478 1
	Standard MSC	0.770 0	0.492 3
	Weighted MSC	0.980 0	0.475 6
	Inverse MSC	0.610 0	0.494 1



图中曲线表示累计贡献率

图 3 主成分分析

Figure 3 Principal component analysis

(0.017 707), 因此,建立的神经网络模型结构为 3—16—1。

### 2.6 神经网络模型验证

用建立的人工神经网络模型对验证集进行预测,模型对验证集的预测值与仪器测定值之间回归效果如图 4 所示,验证集各点分布在趋势线附近,回归决定系数  $R^2$  为 0.992 4,预测标准偏差值为 5.015 17,模型预测能力良好,能满足苹果咀嚼性的检测。

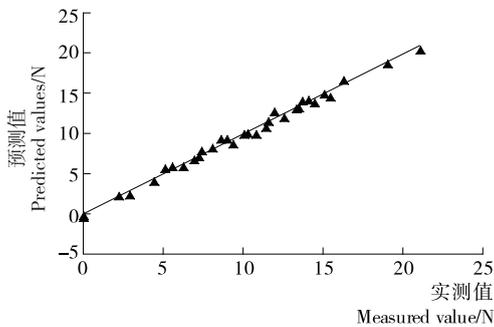


图4 预测值与仪器测定值的相关图

Figure 4 Correlation diagram of the predicted values and the instrument measurements

### 3 结论

建立苹果咀嚼性检测的3—16—1人工神经网络模型结构对验证集预测决定系数 $R^2$ 为0.9924,预测标准偏差值为5.01517,预测性能良好。对苹果的近红外光谱数据进行预处理和通过主成分分析法提取光谱特征,减少建模变量数后,建立人工神经网络模型,实现了对苹果咀嚼性的定量无损检测。该方法与常规的化学分析方法相比,样品处理简单,结果可靠有效,费用经济低廉,方法绿色环保,是一种非常实用的快速分析方法。虽然本研究所用材料具有较好的代表性,但样品咀嚼性的变化范围尚不够宽泛。因此,应在现有研究的基础上,增加一些有代表性的样品,建立通用性强的检测模型需要进一步的研究。

#### 参考文献

- [1] 王轩. 不同产地红富士苹果品质评价及加工适宜性研究[D]. 北京: 中国农业科学院, 2013: 1-3.
- [2] 杨玲, 肖龙, 王强, 等. 质地多面分析(TPA)法测定苹果果肉质地特性[J]. 果树学报, 2014, 31(5): 977-985.
- [3] 张鹏, 李江阔, 陈绍慧. 苹果质地的近红外光谱无损检测模型研

- 究[J]. 食品工业科技, 2015, 36(4): 79-83.
- [4] 李桂峰, 赵国建, 王向东, 等. 苹果质地品质近红外无损检测和指纹分析[J]. 农业工程学报, 2008, 24(6): 169-173.
- [5] Wedding B B, Wright C, Graul S, et al. Effects of seasonal variability on FT-NIR prediction of dry matter content for whole Hass avocado fruit[J]. Postharvest Biology and Technology, 2013, 75(1): 9-16.
- [6] 安泉鑫, 陈莉, 庞林江, 等. 近红外光谱技术在食品中的应用进展[J]. 食品与机械, 2012, 28(5): 239-242.
- [7] 张静, 程玉来, 重滕和明. 利用近红外透射光谱技术测定苹果糖度的研究[J]. 食品科技, 2007(2): 245-247.
- [8] 李桂峰, 赵国建, 刘兴华, 等. 苹果硬度的傅里叶变换近红外光谱无损检测[J]. 农业机械学报, 2009, 40(1): 120-123.
- [9] Lammertyn J, Nicolai B I, Ooms K, et al. Non-destructive measurement of acidity, soluble solids, and firmness of Jonagold apples using NIR-spectroscopy[J]. Transactions of the ASAE, 1998, 41(4): 1 089-1 094.
- [10] 孙炳新, 匡立学, 徐方旭, 等. 苹果脆度的近红外无损检测[J]. 食品与发酵工业, 2013, 39(5): 185-189.
- [11] 陆辉山, 陈鹏强, 闫宏伟, 等. 基于近红外光谱漫透射技术的苹果可溶性固形物含量在线检测[J]. 食品与机械, 2015, 31(3): 40-42.
- [12] 章德宾, 徐家鹏, 许建军, 等. 基于监测数据和BP神经网络的食物安全预警模型[J]. 农业工程学报, 2010, 26(1): 221-226.
- [13] 姜松, 王海鸥. TPA质构分析及测试条件对苹果TPA质构分析的影响[J]. 食品科学, 2004, 25(12): 68-71.
- [14] 张佰清, 关悦乐. 基于神经网络算法用质构特性指标预测寒富苹果贮藏品质[J]. 食品科技, 2012, 37(6): 290-293, 298.
- [15] 赵丽丽. 果品类内部品质近红外无损检测技术的研究[D]. 北京: 中国农业大学, 2003: 33-38.
- [16] 赵武奇, 石珂心, 谷月, 等. 近红外光谱技术检测石榴汁酸度的研究[J]. 食品工业科技, 2014, 35(16): 68-70, 75.
- [17] 夏柏杨, 任芊. 近红外光谱分析技术的一些数据处理方法的讨论[J]. 光谱实验室, 2005, 22(3): 629-634.

#### 信息窗

## 新型“气泡饮品”有望提高癌症化疗效果

新华社电 英国牛津大学6月8日宣布,牛津大学与北爱尔兰大学正合作开发一种含有大量氧气泡的“气泡饮品”,用以提高化疗对癌症的治疗效果。

先前研究发现,细胞缺氧是一些癌症扩散、治疗效果不理想的重要原因。这是随着肿瘤的生长,血管会不断被扭曲,传输能力变弱,氧气无法正常输送,化疗药物很难抵达肿瘤的中心区域。

为了供氧,目前的通行做法是让病人吸入纯氧或将含氧液体直接注射到肿瘤生长部位等,这些做法虽然有效但会带

来很大副作用,包括给肺部表面以及神经系统带来损伤。

动物试验中,牛津大学研究人员给患胰腺癌的试验鼠喝下富含氧气泡的液体,结果发现试验鼠癌肿的缺氧状态得到明显改善,化疗效果有显著提升。

研究人员表示,胰腺癌是一种典型的“缺氧”型癌症,死亡率高、治疗效果不理想。如果通过饮品的形式让病人将氧喝进体内,有望提升化疗药物的效果,而且风险更小、成本更低,他们计划下一步开展临床试验。

(来源:www.foodmate.net)