

应用可见/近红外光谱进行黄酒品种的判别

刘飞, 王莉, 何勇, 蒋益虹*

浙江大学生物系统工程与食品科学学院, 浙江 杭州 310029

摘要 为了实现对黄酒品种的快速判别, 采用可见/近红外光谱对不同品种的黄酒获取光谱曲线, 然后采用主成分分析方法对光谱数据进行聚类分析, 并将其提取的主成分作为BP神经网络的输入值, 建立了黄酒品种鉴别模型。该模型将前6个主成分作为神经网络的输入变量, 加速了神经网络的学习速度, 提高了模型的预测精度。随机选取每个品种的15个黄酒样本, 共45个样本组成预测集, 剩余的145个黄酒样本组成训练集建立训练模型, 并用预测集样本对其进行验证。将品种鉴别的偏差标准定为 ± 0.1 , 结果表明, 只有1个未知样本超出偏差范围, 该方法的品种鉴别正确率为97.78%, 获得了满意的结果。说明文章提出的方法具有很好的分类和鉴别作用, 为黄酒品种的快速鉴别提供了一种新方法。

关键词 可见/近红外光谱; 黄酒; 主成分分析; BP神经网络; 品种鉴别

中图分类号: O657.3; TS262.4 **文献标识码**: A **文章编号**: 1000-0593(2008)03-0586-04

引言

绍兴黄酒是我国黄酒的佼佼者, 在历史上久负盛名。自宋代以来就有所发展。针对目前市场上各种黄酒, 存在一些假冒伪劣产品, 只凭感官评定, 一方面需要经验, 另外受主观因素影响较大, 准确性难以保证^[1]。而常规的化学方法进行酒的不同品种的鉴别费时、费力, 还需要昂贵的分析仪器^[2]。因此, 研究一种简单、快速、无损的黄酒鉴别技术是很有必要的。

可见/近红外光谱法是一种快速简便无损的分析方法。它被广泛应用于食品、石油、化工、医药等行业^[3]。可见/近红外光谱技术在水果上的应用如苹果的品种鉴别^[4], 杨梅的内部品质检测^[5], 水果坚实度的检测^[6]等。在酒的行业, 近红外光谱技术也有所应用。例如利用红外光谱技术对假酒中的甲醇进行定性分析^[7], 利用近红外透射光谱法结合气相色谱建立茅台酒的指纹模型^[8], 对葡萄酒产地鉴别^[2]和化学成分的测定^[9], 啤酒的酒精度^[10]以及主要成分的近红外光谱分析^[11], 白酒酒精度的测定^[12]等。对黄酒不同品种的鉴别方面, 还少有相关报导。本研究通过对不同品种黄酒的可见/近红外透射光谱进行分析, 结合现代化学计量学的方法, 寻找一种鉴别黄酒品种的快速有效的方法。

1 材料与方法

1.1 仪器设备

实验使用美国ASD (Analytical Spectral Device) 公司的Handheld FieldSpec 光谱仪, 其光谱测定范围325~1 075 nm, 采样间隔为1.5 nm, 分辨率为3.5 nm, 探头视场角为20°, 光谱扫描次数设定为30次。光源采用14.5 V 卤素灯。分析软件采用ASD View Spec Pro, Unscrambler V9.6和DPS (data procession system for practical statistic)^[13]。

1.2 样品制备及光谱的扫描

从当地超市购得3个品牌的绍兴加饭酒, 分别为古越龙山, 黄中皇和会稽山。每个品牌都密封保存在恒温25℃的实验室内。在样品采集前, 先将酒摇匀, 分别制备了70、60和60个样本。光谱扫描采用透射法, 光程为2 mm, 每个样本扫描30次, 保存3条光谱, 以其平均光谱作为最终的透射光谱。从全部的190个黄酒样本中, 每个品种随机选择15个共45个样本作为预测集, 其余145个作为定标集。

1.3 光谱数据预处理

为了去除来自高频随机噪声、样本不均匀、基线漂移、光散射等影响, 需对原始光谱数据进行预处理^[14]。我们采用Savitzky-Golay平滑法, 选用平滑点数为3, 并进行SNV (standard normal variate) 处理, 再采用一阶Savitzky-Golay导数处理, 求导间隔为3。为消除光谱数据在采集时首端与

收稿日期: 2006-11-28, 修订日期: 2007-02-26

基金项目: 国家自然科学基金项目(30671213)和高等学校优秀青年教师教学科研奖励计划(02411)资助

作者简介: 刘飞, 1983年生, 浙江大学生物系统工程与食品科学学院博士生 *通讯联系人 email: yhjiang@zju.edu.cn

末端产生的部分噪声,截取 350 ~ 1 050 nm 波段的光谱数据进行分析。

1.4 鉴别模型的建立

主成分分析(PCA)是一种多元统计分析方法,能够在不丢失主要光谱信息的前提下提取出较少的新的特征变量,解决了因近红外谱带的重叠而难以分析的困难。本研究对原始光谱数据进行预处理,通过主成分分析提取有效的特征变量,作为 BP 神经网络的输入值,经过调整网络参数,建立了黄酒品种鉴别的模型。

2 试验结果与分析

2.1 不同黄酒品种的可见/近红外光谱图分析

三种黄酒的可见/近红外光谱图如图 1 所示。每个品种的黄酒随意选取 6 条光谱。图中横坐标为光谱波长,范围为 350 ~ 1 050 nm,纵坐标为吸光度。从图 1 中可以看出,三种黄酒的原始图谱的曲线趋势基本相同,并没有太明显的差别,因此需要进一步结合现代化学计量学方法对光谱进行分析处理,以便建立黄酒品种的鉴别模型。

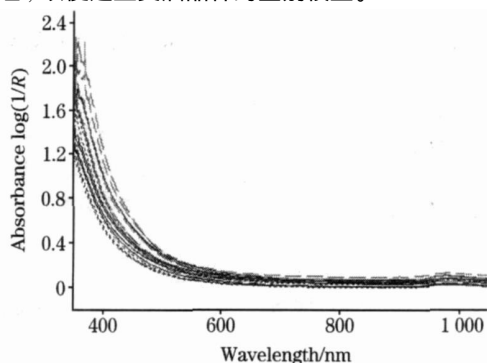


Fig. 1 Vis/NIR absorbance spectra of three different varieties of yellow wines

—: 古越龙山; ---: 黄中皇;: 会稽山

Table 1 First seven principal components and accumulative reliabilities

主成分数	1	2	3	4	5	6	7
累计贡献率/%	55.698	82.332	90.028	92.923	94.790	96.263	97.145

2.3 基于 BP 神经网络的黄酒品种鉴别模型

通过主成分分析,将提取的前 6 个主成分作为 BP 神经网络的输入。通过调整网络参数得到最优的网络结构,从而得到最优的预测模型^[16]。经过反复试验得到的最优网络结构为:输入层为 6 个输入单元,隐含层为 8 个隐含单元和 1 个输出层单元的三层 BP 神经网络模型。通过对 145 个建模样本进行学习训练,限定最大迭代次数为 2 000,得到最终的拟合残差为 1.004×10^{-3} ,对未知的 45 个样本进行预测,预测结果如表 2 所示。设定品种真实值“1”代表古越龙山,“2”代表黄中皇,“3”代表会稽山。设定品种预测结果偏差在 ± 0.1 以内,为品种区分的界限。结果表明,只有 1 个未知样本超出偏差范围,得到对未知样本预测的正确辨别率为

2.2 主成分分析得到新的特征变量

主成分分析是把多个指标化为少数几个综合指标的一种统计分析方法。其主要目的是用来降维,主成分分析找出几个综合因子即特征变量来代表原来众多的变量,使这些特征变量尽可能的反映原来变量的信息量,而且彼此之间互不相关^[13]。图 2 是三个不同品种黄酒的第 1 和第 2 主成分的得分图。图中横坐标表示每个样本的第 1 主成分的得分值,纵坐标表示每个样本第 2 主成分的得分值。从图 2 中可以看出,古越龙山和会稽山的聚类效果较好,并且古越龙山主要分布在纵坐标的正半轴区域,相反会稽山主要分布在纵坐标的负半轴,并且两者的聚合度都比较好。相对而言,黄中皇的分布比较分散,但主要分布在横坐标的负半轴。

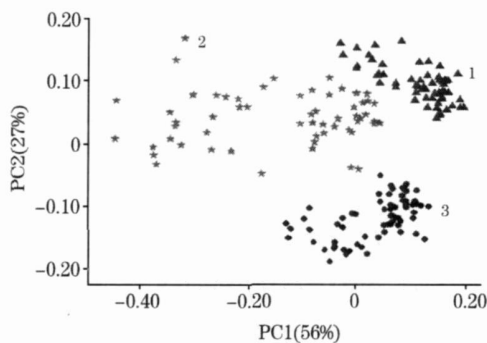


Fig. 2 Scatter plot (PC1 x PC2) of 190 yellow wine samples
1: 古越龙山; 2: 黄中皇; 3: 会稽山

通过主成分分析,得到 190 个黄酒样本的主成分的累计贡献率,如表 1 所示。其中前 6 个主成分的累计贡献率已达 96.263%,并且前 6 个主成分交互验证的累计贡献率也达到 95.662%。而前 7 个主成分的累计贡献率为 97.145%,只增加了 0.882%,说明前 6 个主成分已经足够代表 190 个样本的重要信息,所以选择前 6 个主成分作为 BP 神经网络的输入值^[15]。

97.78%,获得了较理想的预测结果。

3 结论

应用可见/近红外光谱结合主成分分析和 BP 神经网络对黄酒品种进行了初步的判别分析,通过提取样本的主成分作为神经网络的输入,进而得到了黄酒品种鉴别的模型,该模型的识别率达到 97.78%,获得了理想的预测精度。说明利用可见/近红外光谱技术结合现代化学计量学的方法对黄酒品种进行快速鉴别是可行的。为黄酒品种的快速检测提供了一种新方法。

Table 2 Prediction results of 45 unknown yellow wine samples by BP-ANN model

预测样本序号	真实值	预测值	预测样本序号	真实值	预测值	预测样本序号	真实值	预测值
(1)	1	1.000 00	(16)	2	1.997 27	(31)	3	2.995 87
(2)	1	1.000 22	(17)	2	1.999 94	(32)	3	2.997 88
(3)	1	1.000 00	(18)	2	1.957 36	(33)	3	2.998 67
(4)	1	1.029 56	(19)	2	1.905 49	(34)	3	2.997 63
(5)	1	1.000 00	(20)	2	2.012 64	(35)	3	2.995 64
(6)	1	1.000 00	(21)	2	1.939 53	(36)	3	2.991 98
(7)	1	1.000 00	(22)	2	1.922 61	(37)	3	2.995 41
(8)	1	1.000 00	(23)	2	1.954 61	(38)	3	2.995 93
(9)	1	1.000 12	(24)	2	1.939 04	(39)	3	2.998 26
(10)	1	1.000 00	(25)	2	2.123 62	(40)	3	2.998 29
(11)	1	1.000 47	(26)	2	1.982 90	(41)	3	2.998 87
(12)	1	1.000 00	(27)	2	2.049 93	(42)	3	2.998 27
(13)	1	1.000 00	(28)	2	1.994 37	(43)	3	2.998 72
(14)	1	1.000 00	(29)	2	2.003 71	(44)	3	2.998 14
(15)	1	1.000 00	(30)	2	2.020 30	(45)	3	2.937 47

Note: 1 - 古越龙山, 2 - 黄中皇, 3 - 会稽山

参 考 文 献

- [1] YIN Yong, QIU Ming, LIU Yun-hong, et al(殷 勇, 邱 明, 刘云宏, 等). Transactions of the Chinese Society of Agricultural Machinery(农业机械学报), 2003, 34(6): 104.
- [2] Liu L, Cozzolino D, Cynkar W U, et al. J. Agric. Food Chem., 2006, 54: 6754.
- [3] YAN Yan-lu, ZHAO Long-lian, HAN Dong-hai, et al(严衍禄, 赵龙莲, 韩东海, 等). The Foundation of the Near Infrared Spectral Analysis and Its Application(近红外光谱分析基础与应用). Beijing: China Light Industry Press(北京: 中国轻工业出版社), 2005.
- [4] HE Yong, LI Xiao-li, SHAO Yong-ni(何 勇, 李晓丽, 邵咏妮). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2006, 26(5): 850.
- [5] Yongni Shao, Yong He. Journal of Food Engineering, 2007, 79: 1015.
- [6] FU Xia-ping, YING Yi-bin, LIU Yan-de, et al(傅霞萍, 应义斌, 刘燕德, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2006, 26(6): 1038.
- [7] YAO Jie, YANG Qian, SUN Cai-yun, et al(姚 杰, 杨 倩, 孙彩云, 等). Chinese Journal of Spectroscopy Laboratory(光谱实验室), 2000, 17(1): 35.
- [8] WANG Li, WANG Di-qiang, WANG Hua, et al(王 莉, 汪地强, 汪 华, 等). Liquor Making(酿酒), 2005, 32(4): 18.
- [9] Urbano-Cuadrado M, Luque de Castro M D, Perez-Juan P M, et al. Anal. Chim. Acta, 2004, 527: 81.
- [10] CHEN Bin, WANG Hao, LIN Song, et al(陈 斌, 王 毫, 林 松, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2005, 21(7): 99.
- [11] LI Dai-xi, WU Zhi-yong, XU Duan-jun, et al(李代禧, 吴智勇, 徐端钧, 等). Chinese Journal of Analytical Chemistry(分析化学), 2004, 32(8): 1070.
- [12] WANG Zhen-zuo, LU Jia-hui, GUO Wei-liang, et al(王贞佐, 逯家辉, 郭伟良, 等). Science and Technology of Food Industry(食品工业科技), 2005, 26(11): 163.
- [13] TANG Qi-yi, FENG Ming-guang(唐启义, 冯明光). DPS Data Processing System for Practical Statistics(实用统计分析及其 DPS 数据处理系统). Beijing: Science Press(北京: 科学出版社), 2002.
- [14] CHU Xiao-li, YUAN Hong-fu, LU Wan-zhen(褚小立, 袁洪福, 陆婉珍). Progress in Chemistry(化学进展), 2004, 16(4): 528.
- [15] Fernando A I ón, Salvador Garrigues, Miguel de la Guardia. Analytica Chimica Acta, 2006, 571: 172.
- [16] LUO Yi-fan, GUO Zhen-fei, ZHU Zhen-yu, et al(罗一帆, 郭振飞, 朱振宇, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2005, 25(8): 1231.

Discrimination of Varieties of Yellow Wines Using Vis/ NIR Spectroscopy

LIU Fei , WANG Li , HE Yong , JIANG Yi-hong *

College of Biosystems Engineering and Food Science , Zhejiang University , Hangzhou 310029 , China

Abstract In order to achieve the rapid discrimination of the varieties of yellow wines , the spectral curves of yellow wines were obtained by Vis/ NIR spectroscopy , and the principal component analysis (PCA) was applied to perform the clustering analysis. The principal components (PCs) extracted by PCA were employed as the inputs of the BP neural networks , and then a discrimination model was built. The first 6 PCs were regarded as the new eigenvectors to accelerate the training speed and to improve the precision of the model. Fifteen samples from each variety and a total of 45 samples were selected randomly as the prediction sets. The remaining 145 samples were used as the training sets to build the training model which is validated by the samples of the prediction sets. The result error was set to be ± 0.1 , and the results indicated that only one sample exceeded the threshold value , therefore the recognition rate of 97.78 % and an excellent precision were achieved. So the discrimination method studied in the present paper played a good role in the classification and discrimination , and offered a new approach to the rapid discrimination of the varieties of yellow wines.

Keywords Vis/ NIR Spectra ; Yellow wine ; Principal component analysis (PCA) ; BP neural networks ; Variety discrimination

(Received Nov. 28 , 2006 ; accepted Feb. 26 , 2007)

* Corresponding author

www.cnki.net