

纯牛奶中还原奶的近红外检测判别分析

李凯歌¹, 韩东海², 孙 明¹

(中国农业大学 a 信息与电气工程学院; b 食品科学与营养工程学院, 北京 100083)

摘 要: 探讨了近红外光谱技术结合化学计量学算法对纯牛奶中是否掺入还原奶鉴别的可行性。考察了马氏距离判别分析和仿生模式识别方法在还原奶检测中的适用性。结果表明: 利用马氏距离判别分析方法鉴别, 正确判别率为 87.5%; 利用仿生模式识别方法鉴别, 正确判别率为 98.3%。由此方法说明: 近红外技术可以实现对纯牛奶中是否含有还原奶进行快速鉴别; 采用仿生模式识别比马氏距离判别分析法鉴别效果更好。

关键词: 近红外光谱; 纯牛奶; 还原奶; 马氏距离; 仿生模式识别

中图分类号: S37, O657.33

文献标识码: A

文章编号: 1003-188X(2008)08-0145-03

0 引言

随着人们生活水平的日益提高, 牛奶以其香甜可口、营养丰富和易于消化吸收等特点, 赢得了人们的喜爱, 其消费量不断增大。某些牛奶供应商为了谋取商业利益, 将掺有还原奶的牛奶充当纯牛奶销售, 损害了消费者利益。目前, 对掺假牛奶的鉴别, 主要是通过闻气味、看颜色和放置看沉淀物等方式。这些方法较简单, 但并不很准确。

近红外光谱分析是近年来快速发展起来的一种快速、无损、无公害、可多组分同时分析的现代技术, 广泛应用于农产品及食品质量检测中。目前, 近红外光谱分析技术在乳品行业已得到了广泛应用, 如利用近红外分析技术实现原料奶理化指标的快速检测^[1,2]、利用近红外技术实现原料奶中体细胞的检测^[3,4]等。韩东海等人采用近红外技术进行了生鲜乳中还原乳的鉴别^[5]。本文着重考察了采用近红外技术和马氏距离判别法与拓扑模式识别法相结合的方法, 对掺有还原奶的纯牛奶进行鉴别, 以得到更为可行和有效的牛奶掺假判别方法。

1 实验材料与方法

1.1 奶样的获取

原料奶取自中国农业大学中美奶业研究中心的 3 头荷斯坦奶牛。每天中午采样, 采集后放入 4℃ 便携冰箱中冷藏运回。

收稿日期: 2007-10-17

作者简介: 李凯歌 (1983-), 女, 河南平顶山人, 硕士研究生, (E-mail) kakaikg@163.com。

通讯作者: 孙 明 (1963-), 男, 沈阳人, 副教授, 博士, (E-mail) dmingsun@163.com。

不同梯度的还原奶是将奶粉 (伊利全脂鲜奶粉) 加水还原后, 按照不同梯度 (10%、20%、33%、50%、67%、80% 或 100%) 加入原料奶中, 然后在水浴锅中恒温至 40℃, 手摇均质后得到。

1.2 近红外光谱采集

收集的样品利用美国 Nicolet 傅立叶变换近红外光谱仪积分球附件进行积分球反射光谱采集。光谱采集范围为 12000~4000 cm⁻¹, 间隔 2 cm⁻¹, 采集前将样品恒温到 40℃。

1.3 数据处理方法

1.3.1 马氏距离判别分析

预先将样品分为确定的几类, 每张样品谱图形成一个点, 每类样品谱图形成一个簇, 簇的重心点为 x_0 。每类样品簇的重心位置不同。通过比较样品光谱点的位置与各类簇的重心点位置间的接近程度, 对样品进行分类。计算这种距离有多种方法, 在此采用马氏距离法。采用此方法计算距离时, 在不同方向上给出不同的权重。

近红外检测判别试验将采集的纯牛奶样品作为一类, 将不同梯度值的还原奶作为一类, 对纯牛奶和还原奶进行判别分析。光谱预处理及其判别分析利用 TQ Analyst V6.2 软件实现。

1.3.2 仿生模式识别

在高维空间中, 对同类样品进行闭区间最佳覆盖, 不同类样品在覆盖区间外, 以此将一类样品与其他类样品分开。在二维空间中拓扑模式识别如图 1 所示。

近红外检测判别试验将采集的纯牛奶样品分为两个集合: 一个作为建模集, 另一个和不同梯度值的还原奶一起作为预测集。对建模用的纯牛奶样品光

谱在高维空间中进行闭区间最佳覆盖,对预测集样品光谱值计算其到闭区间的距离 d 。整个闭区间为单个闭区间的并集。图 2为选取的单个闭区间模型,即超环与超球的并集。确定半径 r ,如果距离 $d < r$,在闭区间内则认为该样品为纯牛奶;反之,当 $d > r$ 时,在闭区间外则认为该样品为还原奶。光谱预处理及其判别在 M atlab中编程实现。

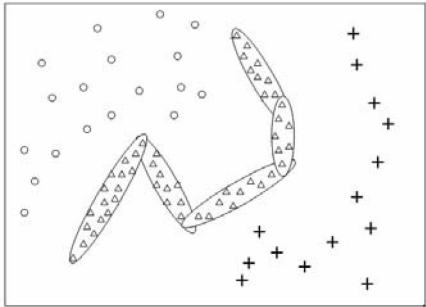


图 1 二维空间中拓扑模式识别

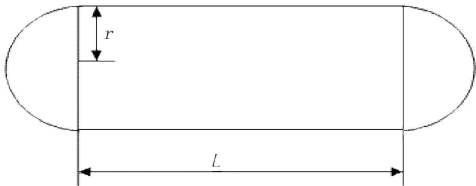


图 2 单个闭区间模型

2 结果与讨论

2.1 不同梯度还原奶的光谱分析

纯牛奶及配制的不同梯度还原奶的光谱图,如图 3所示。从图 3中可以看出:随着还原奶梯度的增加,光谱图有向下漂移的趋势,其与纯牛奶的光谱差异也增大。

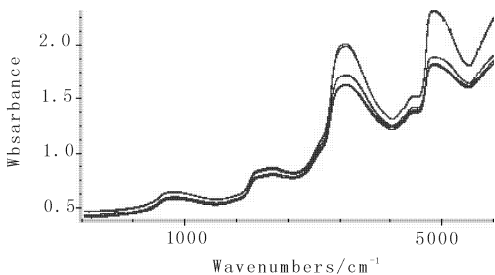


图 3 不同梯度还原奶的光谱图

2.2 马氏距离判别分析对掺有还原奶的原料奶鉴别

图 4为纯牛奶与 10% 浓度值还原奶的判别分析图。样品共 30个,其中有 15个掺入 10% 还原奶的样品,15个为纯牛奶样品。共有 3个误判,正确判别率为 90%。图 5为纯牛奶与 20% 浓度值的还原奶的判别分析图,样品共 30个,其中 15个掺入 20% 还原奶的样品,15个纯牛奶样品。没有误判,正确判别率为 100%。图 6为纯牛奶与不同梯度值的还原奶的判别分析图,样品

为纯牛奶与不同梯度值的还原奶均为 15个,共 120个,其中有 15个误判,正确判别率为 87.5%。

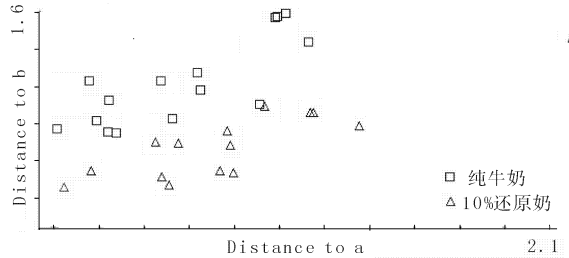


图 4 纯牛奶与 10% 浓度值的还原奶的判别分析图

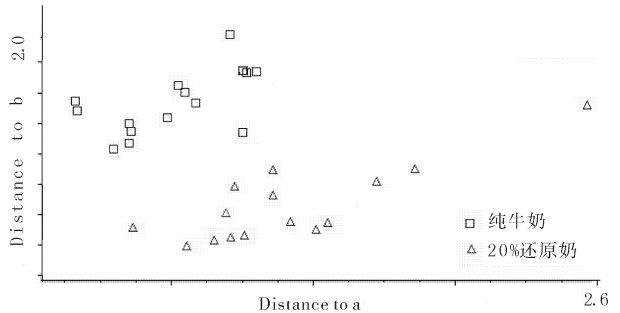


图 5 纯牛奶与 20% 浓度值的还原奶的判别分析图

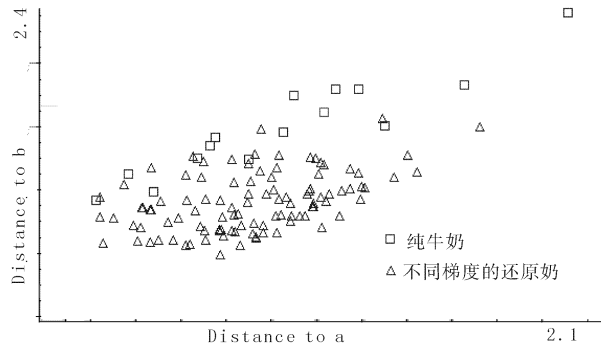


图 6 纯牛奶与不同梯度值的还原奶的判别分析图

2.3 仿生模式识别对掺有还原奶的原料奶鉴别

30个纯牛奶样品作为建模集,15个纯牛奶和 105个不同梯度的还原奶共 120样品作为预测集。图 7为 120个预测集样品到闭区间距离图。

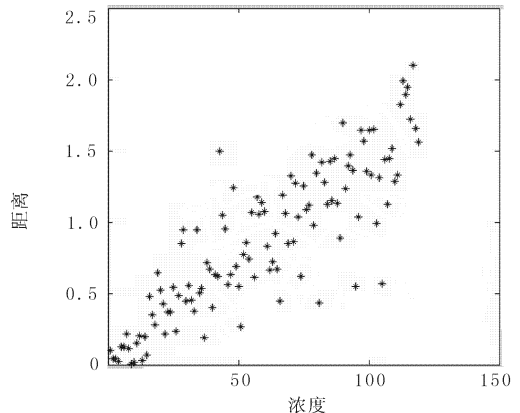


图 7 预测样品到闭区间的距离

预测集样品按还原奶梯度值升序排列。由图 7可以看出,从大体范围上说,随着还原奶浓度值的增加,到

闭区间距离逐渐增大。同时表明,还原奶浓度值越大,与纯牛奶差异越大,越容易区分。表 1为采用仿生模式识别法对掺有还原奶的原料奶的鉴别结果。当选取闭区间半径 $r=0.21$ 时,原料奶误判 1个,掺有还原奶的样品误判 1个,共误判 2个,正确判别率为 98.3%;当

选取不同的 r 值时,原料奶和还原奶误判数将改变;当 r 值增大时,原料奶误判减少,还原奶误判增大;当 r 值减小时,原料奶误判增大,还原奶误判减小。根据不同的需要确定合理的 r 值。从图 7和表 1可以看出,原料奶和还原奶得到了很好的鉴别。

表 1 纯牛奶与不同梯度值的还原奶的判别结果

选取 $r=0.21$	预测样品数	最小距离	最大距离	标准偏差	误判率 /%	正确判别率 /%
原料奶	15	0.017	0.217	0.072	6.7	98.3
还原奶	105	0.192	2.101	0.461	1.0	

2.4 两种判别方法比较

采用马氏距离判别分析方法,对纯牛奶与单一浓度的还原奶可以进行很好的鉴别,但当与不同梯度的还原奶一起建模定性判别时,就会增大误判率。这主要是由于采用马氏距离判别分析时,将纯牛奶与掺有不同梯度的还原奶作为两类建模,鉴别结果主要在于寻找两类样本的差异,达到两类的最优划分。当不同梯度的还原奶作为一类一起建模时,此类样品范围变宽,其重心发生偏移,对低浓度的还原奶就会增大误判率。采用仿生模式识别法鉴别时,对建模集纯牛奶样本合理做闭区间覆盖,不同梯度的还原奶并不参加建模,预测时互不影响,因此对纯牛奶与单一浓度的还原奶或是不同梯度还原奶鉴别的结果没有改变。

还原奶与原料奶进行鉴别时,正确判别率为 98.3%。在现实中,对原料奶中掺入的是不同浓度的还原奶,采用拓扑模式识别判别分析比马氏距离判别分析效果更好。又由于采用拓扑模式识别判别时,掺假牛奶并不参与建模,不影响判别效果,因此鉴别掺假牛奶更有广泛的适用性。

参考文献:

[1] 王 云,徐可欣,常 敏.近红外光谱技术检测牛奶中脂肪及蛋白质含量校正模型的建立[J].光学仪器,2006,28(3): 3-7

[2] 王丽杰,徐可欣,郭建英.采用近红外光谱技术检测牛奶中脂肪、蛋白质及乳糖含量[J].光子·激光,2004,15(4): 468-471.

[3] R. Tsenkova, S. Atanassova, Y. Ozaki, et al. Near infrared spectroscopy for biomonitoring influence of somatic cell count on cow's milk composition analysis[J]. International Dairy Journal, 2001(11): 779-783.

[4] V. Pravdova, B. Walczak, D. L. Massart, et al. Calibration of somatic cell count in milk based on near-infrared spectroscopy[J]. Analytica Chimica Acta, 2001, 450(12): 131-141.

[5] 韩东海,鲁 超,刘 毅,等.生鲜乳中还原乳的近红外光谱法鉴别[J].光谱学与光谱分析,2007,27(3): 465-468

3 结论

- 1) 利用马氏距离判别分析,对单一浓度的还原奶与原料奶的鉴别时,当掺入还原奶为 10% 时,正确判别率为 90%;当掺入还原奶为 20% 及以上时,正确判别率为 100%;当不同梯度的还原奶一起与原料奶进行鉴别时,正确判别率为 87.5%。
- 2) 利用拓扑模式识别法判别分析时,不同梯度的

Identification of Reconstructed Milk in Raw Milk Using Near Infrared Spectroscopy

LI Kai-ge¹, HAN Dong-hai², SUN Ming¹

(China Agricultural University a College of Information and Electrical Engineering b College of Food Science and Nutritional Engineering Beijing 100083 China)

Abstract Feasibility of reconstructed milk identification in raw milk was investigated using near infrared spectroscopy combining with chemometrics algorithms. And the applicability of Mahalanobis distance discrimination analysis and biometric pattern recognition for reconstructed milk identification was discussed. The result indicated that the accuracy of detection is 87.5% using Mahalanobis distance discrimination analysis while 98.3% using biometric pattern recognition. All of these suggested that NIRS has good potential to detect adulteration of raw milk with reconstructed milk. And biometric pattern recognition is better than Mahalanobis distance discrimination analysis in the identification.

Key words near infrared spectroscopy; raw milk; reconstructed milk; mahalanobis distance; biometric pattern recognition