

二维相关光谱结合偏最小二乘法测定牛奶中的掺杂尿素

杨仁杰^{1,2}, 刘蓉^{1*}, 徐可欣¹

(1. 天津大学精密测试技术及仪器国家重点实验室, 天津 300072; 2. 天津农学院机电工程系, 天津 300384)

摘要: 为了检验牛奶中是否掺杂尿素并将其量化测定, 配置含有尿素质量浓度范围为 1~20 g/L 之间 40 个牛奶样品, 以掺杂物尿素浓度为外扰, 分别研究了掺杂尿素牛奶的二维相关(近红外-近红外, 中红外-中红外, 近红外-中红外)光谱特性, 在此基础上, 分别选择随浓度变化大的 4 200~4 800 cm^{-1} 和 1 400~1 704 cm^{-1} 为建模区间, 采用偏最小二乘法建立定量分析模型。研究表明: 4 200~4 800 cm^{-1} 建模分析效果优于 1 400~1 704 cm^{-1} 建模结果, 其交叉验证均方根误差为 0.266 g/L, 对未知样品集预测相关系数达到 0.999, 预测均方根误差为 0.219 g/L, 这表明所建模型具有较好的预测效果。该方法无需样品处理, 成本低, 为快速判别牛奶是否掺杂提供了一种新的可能的方法。

关键词: 红外光谱, 尿素, 模型, 偏最小二乘法, 掺杂牛奶

doi: 10.3969/j.issn.1002-6819.2012.06.042

中图分类号: O657.33

文献标志码: A

文章编号: 1002-6819(2012)-06-0259-05

杨仁杰, 刘蓉, 徐可欣. 二维相关光谱结合偏最小二乘法测定牛奶中的掺杂尿素[J]. 农业工程学报, 2012, 28(6): 259—263.
Yang Renjie, Liu Rong, Xu Kexin. Detection of urea in milk using two-dimensional correlation spectroscopy and partial least square method[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2012, 28(6): 259—263. (in Chinese with English abstract)

0 引言

牛奶被称作居民生活中的“白色血液”, 蛋白质是牛奶的重要营养物质, 所以蛋白质含量指标对于牛奶的品质是非常重要的。尿素是蛋白质代谢产物, 因此, 一些不法商人为了追求高额利润, 在蛋白质含量不合格的牛奶中掺入尿素, 以提高牛奶中表观蛋白质含量。而这种掺杂的牛奶严重危害了消费者的身体健康, 同时使乳品加工企业造成巨大的经济损失。

二维相关光谱方法凸显由外界扰动引起的光谱变化的细微特征^[1-4], 是强大而灵活的光谱分析技术。二维相关光谱分析相对于传统的红外光谱有以下显著优势: 二维光谱因为将光谱信号扩张到第二维上, 故具有较高的分辨率, 可以区分出在一维光谱上被覆盖的小峰和弱峰, 从而提高了光谱的分辨率; 通过谱线之间相关性的分析, 能够详细地研究不同分子间或分子内的相互作用, 通过同步交叉峰和异步交叉峰的相关性可以分析信息来源, 从而提高光谱的解释能力; 不同光谱数据的二维结果可以相互验证, 从而减小由于基线、噪音或吸收峰位置红移/蓝移所造成的二维分析误差, 大大提高二维分析的准确度; 二维相关红外光谱技术是一个很有发展前景的分

析技术, 这种技术无需提取分离即可直接测试, 大大缩短了分析时间。

传统的奶制品质量检测主要是使用一些化学分析方法, 其结果准确, 但过程耗时, 费力, 一般需要试剂, 对样本具有破坏性, 不适合于原料奶收购中及时的质量控制^[5]。红外光谱分析技术具有无需对样品处理、成本低、对样品无损检测等特点^[6-13], 适合对乳制品中掺杂物质的检测。之前何彬等^[6,14]以温度为外扰, 研究了牛奶中掺杂尿素、三聚氰胺等物质的二维红外相关红外光谱特性, 指出通过二维相关谱可以定性识别牛奶中掺杂的目标物。但是由于纯牛奶中各组分与掺杂物的光谱信息都随着外扰温度发生变化, 而且其特征吸收峰相互重叠, 所以牛奶中微量的掺杂物光谱信息无法有效地提取。本文以牛奶中掺杂尿素浓度为外扰, 研究了掺杂尿素牛奶的二维相关光谱特性, 有效地提取了牛奶中掺杂尿素的特征光谱信息, 并选择随掺杂尿素浓度变化最敏感的光谱信息区间为建模区间, 采用偏最小二乘法建立定量分析牛奶中掺杂尿素的数学模型, 实现对牛奶中掺杂的尿素定性和定量分析。该研究对于维护乳制品市场稳定, 保护消费者利益具有重要的意义。

1 材料与方 法

1.1 仪器与分析条件

采用美国 PerkinElmer 公司的 Spectrum GX 傅立叶变换红外光谱仪, 近红外采用液氮冷却的 InGaAs 探测器; 中红外采用 ATR 附件, ZnSe 晶体池; FLUKO 公司的 FA25 型高剪切分散乳化机。

仪器参数: 近红外光源为仪器自带的卤素灯, 扫描范围为 4 000~10 000 cm^{-1} ; 中红外光源为稳压空气冷却线盘, 扫描范围为 800~4 000 cm^{-1} ; 两个波段的分辨率

收稿日期: 2011-05-08 修订日期: 2011-11-10

基金项目: 国家自然科学基金(60938002, 30900275)、高等学校博士学科点专项科研基金(20090032120064)

作者简介: 杨仁杰(1978—), 男, 山西运城人, 在职博士, 天津农学院机电工程系讲师, 研究方向为食品安全检测。天津 天津大学精密测试技术及仪器国家重点实验室, 300072。Email: rjyang1978@163.com

*通信作者: 刘蓉(1978—), 女, 副教授, 研究方向为组织光学与光谱应用。天津 天津大学精密测试技术及仪器国家重点实验室, 300072。

Email: rongliu@tju.edu.cn

都为 4 cm^{-1} ，扫描间隔为 8 cm^{-1} ，扫描次数为 8。

试验前，对所配置的掺杂尿素牛奶进行均质。为了减少随机误差，随机抽取样品，而不按浓度高低顺序进行测量；每次测量均扫描 8 次取平均以降低随机噪声。测量样品前，先测量蒸馏水的光谱作为背景光谱，用样品光谱扣除背景光谱，即得到样品的吸收光谱。

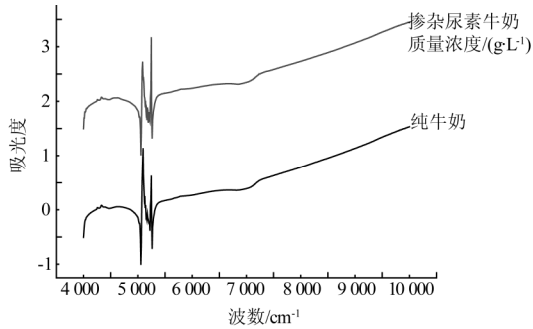
1.2 试剂和样品处理

试验中，所采用的牛奶包括：蒙牛全脂灭菌纯牛奶、蒙牛鲜牛奶、伊利全脂灭菌纯牛奶和海河全脂灭菌纯牛奶。在牛奶中分别加入不同质量的尿素（分析纯，中国医药集团上海化学试剂公司），分别配置含有不同质量浓度范围为 $1\sim 20\text{ g/L}$ 的牛奶样品共 40 个。样品的浓度分布按照低浓度分布密、高浓度分布稀疏的原则。用浓度梯度法从掺有尿素的牛奶样品中选出 27 个样品作为校正集，余下 13 个作为预测集。

1.3 分析方法

1.3.1 二维红外谱的相关运算

使用 2D shige 软件（由日本 Kwansai Gakuin 大学



a. 近红外光谱

Daisuke Adachi 所编写) 进行相关运算，以平均谱作为参考谱，得到二维红外相关图。在相关谱图中，非阴影部分是正相关区域，阴影部分是负相关区域。

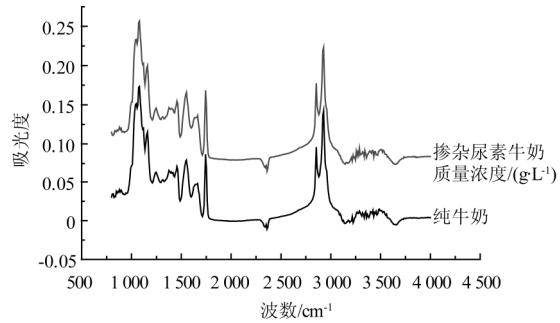
1.3.2 偏最小二乘 (PLS) 分析方法

采用 Unscrambler 分析程序中 PLS 建立数学模型。利用交叉验证 (Cross-Validation) 的方法来确定模型的最佳主成分数，利用测定值与模型计算值之间的建模相关系数 (R_v)，内部交叉验证均方根误差 (RMSECV)，相关系数 (R_p)，均方根误差 (RMSEP) 等指标评定模型的质量。

2 结果与分析

2.1 纯牛奶和掺杂尿素牛奶常规二维红外特性

图 1 分别是纯牛奶和掺杂尿素牛奶在 $4000\sim 10000\text{ cm}^{-1}$ 近红外和 $800\sim 4000\text{ cm}^{-1}$ 中红外范围的吸收谱。显然在整个光谱范围内纯牛奶样品与掺杂尿素牛奶样品的光谱形状，谱峰位置几乎相同，看不出关于掺杂目标物尿素任何信息，表明简单的通过光谱形状和谱峰位置很难判别牛奶中是否掺杂尿素。



b. 中红外光谱

图 1 纯牛奶和掺杂牛奶的近红外与中红外光谱

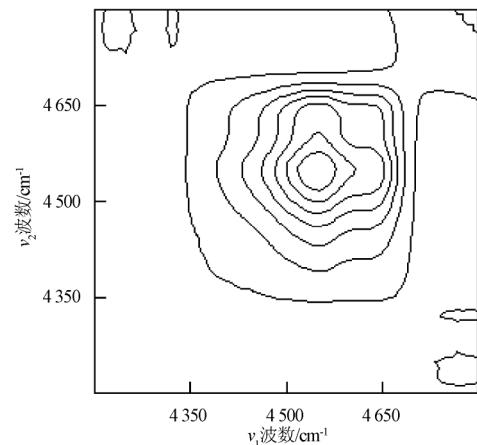
Fig.1 Near-infrared and mid-infrared spectrum of pure milk and adulterated milk

2.2 纯牛奶和掺杂牛奶二维相关红外特性

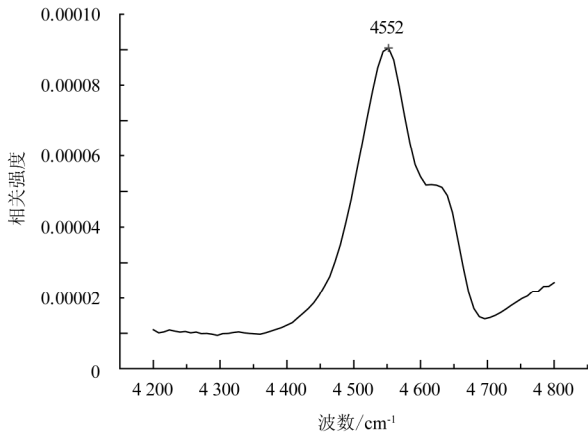
同步二维相关谱强度变化代表了在波数 ν_1 和 ν_2 测得的光谱强度同时或同步的变化，它是关于对角线对称的（即相应光谱的相同波数坐标 $\nu_1=\nu_2$ ）。相关峰可以出现在对角线上或对角线之外。同步相关图中对角线上 ($\nu_1=\nu_2$) 自相关峰的出现表明所对应的振动峰随着掺杂浓度增大的变化，其反映了同步图中自相关峰的峰高。对角线两侧交叉峰的出现表明其对应在 (ν_1, ν_2) 两个频率的一对基团振动峰强度变化在随尿素浓度变化过程中是彼此相关的，如果交叉峰为正，说明两个峰所对应的官能团吸收随牛奶中掺杂物浓度的增大变化相同，两个官能团可能来自同一物质。

以掺杂物尿素浓度为外扰，在 $4000\sim 10000\text{ cm}^{-1}$ 和 $800\sim 4000\text{ cm}^{-1}$ 区间分别构建同步二维相关红外谱。由于随着牛奶中掺杂尿素浓度的增大，其光谱信息变化主要发生在 $4200\sim 4800\text{ cm}^{-1}$ 和 $1400\sim 1704\text{ cm}^{-1}$ 区间，所以选取这两个波段来研究掺杂尿素牛奶的二维相关光谱特性。就图 2 和图 3 分别是在两波段各自的同步二维相关谱及其对应的自相关谱，图 4 是 $4200\sim 4800\text{ cm}^{-1}$ 与 $1400\sim 1704\text{ cm}^{-1}$ 区间的异

谱 (NIR-IR) 相关同步谱，可以看出很多在常规二维谱中的重叠峰变得可以清晰的分辨出来。显然，在图 2 近红外同步相关谱图中，在对角线上出现了一个较强的自相关峰，其所对应的基团振动峰的位置在 4552 cm^{-1} 处（在常规二维谱中并未出现），表明该吸收峰所对应的基团随着牛奶中掺杂尿素浓度的增大变化较明显，该峰是尿素 N-H 反对称伸缩振动与 N-H 弯曲振动合频产生的^[15-17]。



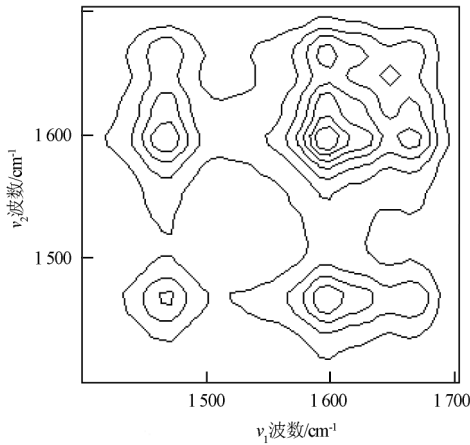
a. 二维相关近红外同步谱



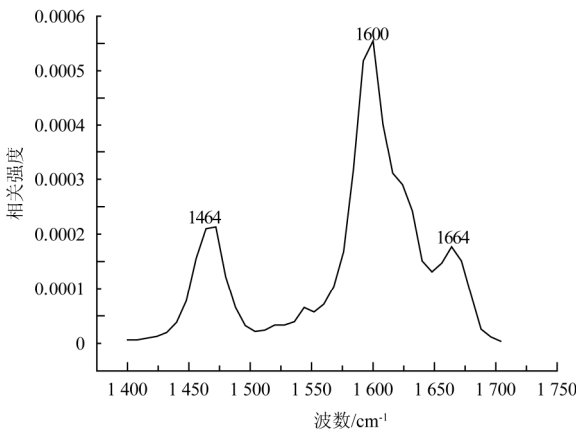
b. 自相关谱

图 2 掺杂尿素牛奶二维近红外同步谱与自相关谱

Fig.2 2D NIR synchronous map and autocorrelation spectrum of milk adulterated by urea



a. 二维相关中红外同步谱



b. 自相关谱

图 3 掺杂尿素牛奶二维中红外同步谱与自相关谱

Fig.3 2D IR synchronous map and autocorrelation spectrum of milk adulterated by urea

在同步二维相关中红外谱(图3),在对角线上出现3个较强的自相关峰,它们所对应的基团振动峰的位置分别在1464、1600和1664 cm⁻¹处,这一波段的吸收峰主要来自尿素的C-N伸缩振动和酰胺III带C=O伸缩振动^[18-21]。在近红外(4200~4800 cm⁻¹)-中红外(1400~1704 cm⁻¹)异谱二维相关图中(图4),出现3个较强正相关交叉峰,

其位置分别在(4552, 1464)、(4552, 1600)和(4552, 1664) cm⁻¹,这说明随着牛奶中掺杂尿素浓度的增大,4552 cm⁻¹所对应的基团吸收峰强度变化与1464、1600和1664 cm⁻¹处所对应的基团吸收峰强度变化相同,这就进一步的说明,这些吸收峰都来自于牛奶中的掺杂尿素。

根据上述对掺杂尿素牛奶二维相关红外谱特性的研究,可以定性的判定牛奶中是否掺杂尿素,可以实现定性分析。

2.3 牛奶中尿素含量定量检测模型的建立

光谱范围对建立模型的性能和预测准确性有很大的影响,所以最理想建模区间应该选在测量目标物特征信息波段,且该波段不受样品中其它物质的干扰,或者说选择随着待测目标物浓度变化大的波数区间。在上述研究掺杂尿素牛奶二维相关谱特性的基础上,选择4200~4800 cm⁻¹和1400~1704 cm⁻¹分别为建模区间,将光谱数据作为输入变量,牛奶中掺杂的尿素浓度作为预测值,建立PLS模型。通过测定未知样品的光谱,应用相应的数学校正模型即可预测其成分浓度,其预测结果如表1所示。显然,无论从预测相关系数,还是内部交叉验证均方根误差和预测均方根误差来说,4200~4800 cm⁻¹建模分析效果优于1400~1704 cm⁻¹建模结果。

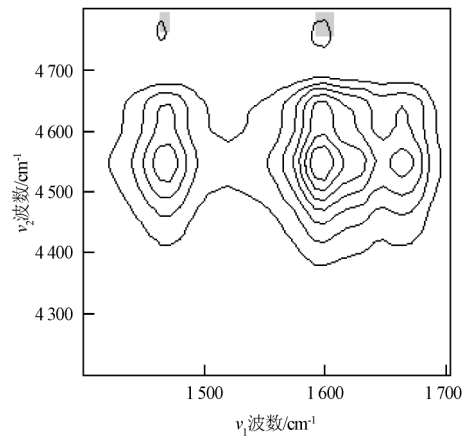


图 4 掺杂尿素牛奶二维近红外-红外同步谱图

Fig.4 2D NIR-IR synchronous map of milk adulterated by urea

表 1 不同波段 PLS 建模结果比较

Table 1 Comparisons of parameters of PLS models within different wavelengths

波数范围/(cm ⁻¹)	主成分数	校正集		验证集	
		建模相关系数	交叉验证均方根误差/(g·L ⁻¹)	相关系数	均方根误差/(g·L ⁻¹)
4200~4800	2	0.999	0.266	0.999	0.219
1400~1704	2	0.993	0.669	0.997	0.462

从表中可以看出:4200~4800 cm⁻¹波段所建模型,其内部交叉验证均方根误差为0.266 g/L,相关系数R_v为0.999,说明所建模型的稳定性和拟合效果非常好。利用所建的模型对预测集进行预测,采用外部验证方法对所建模型实际预测效果进行评价。图5是采用模型对独立的预测集样品的预测结果(C_{预测})与实测结果(C_{实测})之

间关系的线性拟合,其拟合方程为 $C_{\text{预测}} = -0.052 + 1.009C_{\text{实测}}$, 相关系数 R_p 为 0.999, 预测均方根误差为 0.219 g/L。

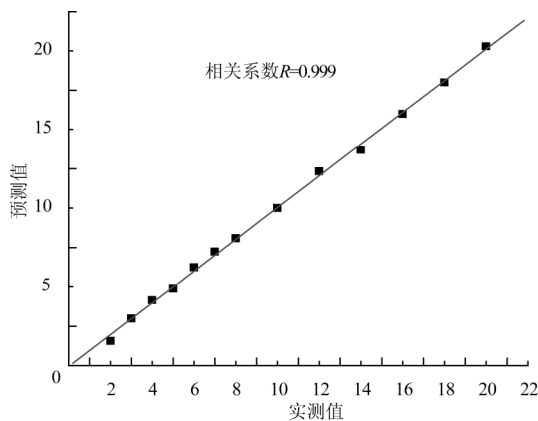


图5 预测集样本预测值和实测值之间关系

Fig.5 Relation between true concentration and predicted concentration of urea in prediction set

3 结论

采用二维相关红外光谱技术,可以实现牛奶中掺杂微量的定性分析,选择随掺杂物尿素浓度变化最敏感的 $4\ 200\sim 4\ 800\ \text{cm}^{-1}$ 光谱区间为建模区间,进而利用 PLS 建立定量分析模型,具有较好的预测效果,预测值非常接近实测值,其相关系数都超过了 0.99,预测均方根误差为 0.219 g/L。该方法简单、快速,可为实现快速检测牛奶中掺杂物提供一种新的方法。

[参考文献]

- [1] Noda, I. Advances in two-dimensional correlation spectroscopy[J]. *Vibrational Spectroscopy*, 2004, 36(2): 143–165.
- [2] Liu Hongxia, Sun Suqin, Lü Guanghua, et al. Study on angelica and its different extracts by Fourier transform infrared spectroscopy and two-dimensional correlation IR spectroscopy[J]. *Spectrochimica Acta Part A*, 2006, 64(2): 321–326.
- [3] Czarnik-Matusiewicz B, Murayama K, Tsenkova R, et al. Analysis of near-infrared spectra of complicated biological fluids by two-dimensional correlation spectroscopy: protein and fat concentration-dependent spectral changes of milk[J]. *Applied Spectroscopy*, 1999, 53(12): 1582–1594.
- [4] Sasic S, Ozaki Y. Wavelength–wavelength and sample–sample two-dimensional correlation analyses of short-wave near-infrared spectra of raw milk[J]. *Applied Spectroscopy*, 2001, 55(2): 163–172.
- [5] 唐玉莲. 近红外光谱在乳制品成分快速检测方面的应用研究[J]. *乳业科学与技术*, 2009, 32(4): 190–194. Tang Yulian. Study on near infrared spectroscopy technique for composition detection of dairy production[J]. *Journal of Dairy Science and Technology*, 2009, 32(4): 190–194. (in Chinese with English abstract)
- [6] 何彬. 基于红外光谱技术的牛奶掺杂判别方法的研究[D]. 天津, 天津大学, 2010. He Bin. Research on Discriminant Analysis of Milk Adulteration by Infrared Spectroscopy[D]. Tianjin University, 2010. (in Chinese with English abstract)
- [7] He Bin, Liu Rong, Yang Renjie, et al. Adulteration detection in milk using infrared spectroscopy combined with two-dimensional correlation analysis[C]// *Proceedings of SPIE*, 2010, 7572: 75720P1-7.
- [8] 屠振华, 朱大洲, 籍保平, 等. 基于近红外光谱技术的蜂蜜掺假识别[J]. *农业工程学报*, 2011, 27(11): 382–387. Tu Zhenhua, Zhu Dazhou, Ji Baoping, et al. Adulteration detection of honey based on near-infrared spectroscopy[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2011, 27(11): 382–387. (in Chinese with English abstract)
- [9] Lu Chenghui, Xing Bingren, Hao Gang, et al. Rapid detection of melamine in milk powder by near infrared spectroscopy[J]. *J. Near infrared Spectroscopy*, 2009, 17(2): 59–67.
- [10] 李水芳, 单杨, 朱向荣, 等. 近红外光谱结合化学计量学方法检测蜂蜜产地[J]. *农业工程学报*, 2011, 27(8): 350–354. Li Shuifang, Shan Yang, Zhu Xiangrong, et al. Detection of geographical origin of honey using near-infrared spectroscopy and chemometrics[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2011, 27(8): 350–354. (in Chinese with English abstract)
- [11] Mauer L J, Chernyshova A A., Hiatt A, et al. Melamine detection in infant formula powder using near- and mid-infrared spectroscopy[J]. *J. Agric. Food Chem*, 2009, 57(10): 3974–3980.
- [12] 陶琳, 武中臣, 张鹏彦, 等. 近红外光谱法快速鉴定干海参产地[J]. *农业工程学报*, 2011, 27(5): 364–366. Tao Lin, Wu Zhongchen, Zhang Pengyan, et al. Rapid identification of dried sea cucumber products from different geographical areas by near-infrared spectroscopy[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2011, 27(5): 364–366. (in Chinese with English abstract)
- [13] Kasemsumaran S, Thanapase W, Kiatsoonthon A. Feasibility of near infrared spectroscopy to detect and to quantify adulterants in cow milk[J]. *Analytical Sciences*, 2007, 23(7): 907–910.
- [14] Sivakesava S, Irudayaraj J. Rapid determination of tetracycline in milk by FT-MIR and FT-NIR spectroscopy[J]. *J. Dairy Sci*, 2002, 85(3): 487–493.
- [15] Eddy C V, Flanigan M, Arnold M A. Near-infrared spectroscopic measurement of urea in dialysate samples collected during hemodialysis treatments[J]. *Applied Spectroscopy*, 2003, 57(10): 1230–1235.
- [16] Olesberg J T, Arnold M A, Flanigan M J. Online measurement of urea concentration in spent dialysate during hemodialysis[J]. *Clinical Chemistry*, 2004, 50(1): 175–181.
- [17] Cho D S, Olesberg J T, Flanigan M J, et al. On-line near-infrared spectrometer to monitor urea removal in real time during hemodialysis[J]. *Applied Spectroscopy*, 2008, 62(8): 866–872.
- [18] Czarnik-Matusiewicz B, Kim S B, Jung Y M. A study of urea-dependent denaturation of β -Lactoglobulin by principal

- component analysis and two-dimensional correlation spectroscopy [J]. *J. Phys. Chem. B*, 2009, 113(2), 559–566.
- [19] Jung Y M, Czarnik-Matusewicz B, Kim S B. Characterization of concentration-dependent infrared spectral variations of urea aqueous solutions by principal component analysis and two-dimensional correlation spectroscopy[J]. *J. Phys. Chem. B*, 2004, 108(34): 13008–13014.
- [20] 杨仁杰, 刘蓉, 徐可欣. 基于中红外光谱检测牛奶中掺杂尿素[J]. *光谱学与光谱分析*, 2011, 31(9): 2383–2385.
- Yang Renjie, Liu Rong, Xu Kexin. Adulteration detection of urea in milk by mid-infrared spectroscopy[J]. *Spectroscopy and Spectral Analysis*, 2011, 31(9): 2383–2385. (in Chinese with English abstract)
- [21] Jung Y M. Two-dimensional correlation spectroscopy in analyzing the concentration-dependent IR spectra of urea aqueous solution[J]. *Bull. Korean Chem. Soc*, 2003, 24(9): 1243–1244.

Detection of urea in milk using two-dimensional correlation spectroscopy and partial least square method

Yang Renjie^{1,2}, Liu Rong^{1*}, Xu Kexin¹

(1. *State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin 300072, China;*

2. *Department of Electromechanical Engineering, Tianjin Agricultural University, Tianjin 300384, China*)

Abstract: For the detection and quantification of urea in milk, pure milk samples and 40 adulterated milk samples added different contents of urea were prepared. Then 2D correlation (NIR-NIR, IR-IR, NIR-IR) spectroscopy under the perturbation of adulteration concentration was calculated and the spectra in the range of 4 200-4 800 cm^{-1} and 1 400-1 704 cm^{-1} were selected to construct the partial least square (PLS) calibration model, respectively. The PLS calibration model showed 4 200-4 800 cm^{-1} was the better range for calibration performance and the root mean square errors of cross validation (RMSECV) of the model was 0.266 g/L. When using this model for predicting the urea contents in prediction set, the root mean square errors of prediction (RMSEP) was 0.219 g/L and the coefficient correlation of actual values and predicted values was 0.999, which means the model has good prediction ability. The method can be used for a correct discrimination on whether the milk is adulterated and provides a new and cost-effective alternative to test the adulteration of milk.

Key words: infrared spectroscopy, urea, models, partial least square, adulterated milk