

doi:10.6041/j.issn.1000-1298.2013.08.028

# 可见/近红外光谱结合软化指标快速判定沙果货架期\*

薛建新 张淑娟 孙海霞 周靖博

(山西农业大学工学院, 太谷 030801)

**摘要:** 基于可见/近红外光谱技术探讨利用软化指标对不同货架期沙果进行分类的可行性。以 SNV 为最优预处理结合 PLS 建立软化指标模型, 经优化后其校正和预测模型的决定系数  $R^2$  分别为 0.847 和 0.813。证实软化指标与光谱数据之间存在很高的相关性, 可以以软化为指标对不同货架期的沙果进行分类。分别采用非线性的 LS-SVM 和线性的 PLS-LDA 建立分类模型, 结果表明: 利用 LS-SVM 所建模型得到的效果要优于 PLS-LDA 模型, 其正确识别率和正确拒识率都达到了 94%。

**关键词:** 沙果 可见/近红外光谱 软化指标 货架期

**中图分类号:** S123; S661.3 **文献标识码:** A **文章编号:** 1000-1298(2013)08-0169-05

## Detection of Shelf Life of *Malus asiatica* Using Near-infrared Spectroscopy and Softening Index

Xue Jianxin Zhang Shujuan Sun Haixia Zhou Jingbo

(College of Engineering, Shanxi Agricultural University, Taigu 030801, China)

**Abstract:** The classified feasibility of *Malus asiatica* fruit by using near-infrared spectroscopy (VIS/NIR spectrometer) and softening index during different shelf life was analyzed. A mathematical model was established by combining partial least square (PLS) and standard normalized variate (SNV) which was regarded as the best pre-processing technology. The determination coefficients ( $R^2$ ) of calibration set and prediction set were 0.847 and 0.813 respectively, which illustrated that there was a high correlation between spectrum and the softening index. It showed that the softening index could be used to classify *Malus asiatica* samples during different shelf life. The LS-SVM and PLS-LDA were applied to build classification models, respectively. The results indicated that non-linear LS-SVM model was more suitable for classification of *Malus asiatica* samples than linear PLS-LDA model. The average correct recognition rate and average correct rejection rate were above 94%.

**Key words:** *Malus asiatica* Visible/near-infrared spectroscopy Softening index Shelf life

### 引言

沙果<sup>[1]</sup> (*Malus asiatica*) 学名花红, 蔷薇科苹果, 属于落叶小乔木, 是我国特有的一个苹果栽培种。沙果中含有多种维生素、矿物质、微量元素和糖等营养成分, 尤以硒、锌、铁及抗氧化因子的含量最为突

出, 具有治疗高血压、冠心病等功效, 加之经济实惠, 因此一直很受消费者青睐。

水果货架期<sup>[2~3]</sup> 是指水果结束储藏状态, 进入流通环节并最终被消费者食用的流通期。水果的成熟度是决定水果品质的主要因素, 对于沙果而言, 成熟期间最重要的变化就发生在货架期阶段, 因而检

收稿日期: 2013-03-11 修回日期: 2013-04-04

\* 国家自然科学基金资助项目(31271973)、高等学校博士学科点专项科研基金资助项目(20101403110003)和山西省自然科学基金资助项目(2012011030-3)

作者简介: 薛建新, 博士生, 主要从事农产品无损检测研究, E-mail: vickyxjx@126.com

通讯作者: 张淑娟, 教授, 博士生导师, 主要从事农产品无损检测研究, E-mail: zsujuan@263.net

测果实采收到消费的货架期过程非常重要。而软化程度<sup>[4]</sup>是评价水果货架期品质变化的一个重要的指标,在表征水果货架期的组织变化时,和硬度相比软化指标较少出现在文献资料中<sup>[5]</sup>。然而事实上,这两个属性从不同角度描述了水果相同的属性<sup>[6]</sup>。

由于近红外漫反射光谱技术具有无需预处理、速度快、无破坏、适于在线分析等优点,已经广泛应用于农产品品质及品种的鉴别研究<sup>[7~10]</sup>。

本文以沙果为实验材料,以硬度的变化为主要依据确定沙果在不同货架期的软化指标,使用 PLS 建立光谱与软化指标的相关模型;并以软化为分类指标,分别采用非线性的 LS-SVM 建模方式和线性的 PLS-LDA 建模方式建立不同货架期沙果的分类模型。

## 1 材料与方 法

### 1.1 实验材料

沙果样品采自山西省太谷县,当天运达实验室,并对果实进行筛选,以保证实验前果实的物性初始状态相对一致,降低果实之间个体差异对实验结果的影响。将精选出的 240 个沙果样品分别编号和标记,并在常温条件下储藏。采后的第 1、4、7 天从常温储藏的样品中每次随机抽取 80 个样品,其中随机选取 60 个样本作为校正集,其余 20 个样本作为预测集,分别对其进行光谱采集,再使用质构仪对其硬度进行测定。

### 1.2 实验方法

#### 1.2.1 光谱采集

光谱仪使用美国 ASD (Analytical Spectral Device) 公司的 FieldSpec 3 型光谱仪,光谱数据间隔为 1 nm,采样范围为 350 ~ 2 500 nm,扫描次数为 30,分辨率 3.5 nm,探头视场角 10°,采用漫反射方式进行样品光谱采样。在每个果实最大横径处等距离取 3 个点,扫描 3 次,使用 ASD View Spec Pro V5.0 软件求其 3 次扫描的平均光谱值。

#### 1.2.2 硬度测定

沙果硬度真实值测定采用 TMS-PRO 型食物物性分析仪(美国 FTC 公司),测试探头为直径 6.0 mm 的圆平头,测前和测后速度均为 3.0 mm/s,测中速度 1.5 mm/s,起始力 0.3 N,插入深度为 6.0 mm<sup>[11]</sup>。用食物物性分析仪测定沙果硬度时,探头插入点与光谱采集位点相对应。硬度的统计结果如表 1 所示。

#### 1.2.3 软化指标确立

沙果软化的变化程度主要由硬度和货架期的变化来表述,因此,通过数据拟合建立了一个非线性的

表 1 货架期间沙果样本硬度的变化及分析

Tab.1 Comparison of firmness of *Malus asiatica* during shelf life

货架期/d	最大值	最小值	平均值	标准偏差	方差值
	/N·mm <sup>-1</sup>	/N·mm <sup>-1</sup>	/N·mm <sup>-1</sup>		
1	10.20	6.50	6.69	1.343	1.805
4	6.30	4.50	5.25	0.991	0.983
7	4.60	2.30	3.84	0.879	0.772

混合效应模型来定义软化指标,该软化指标定义为

$$I = \left( \frac{F_i}{F_{\max}} \right)^{1 + \frac{t}{w+t}} \quad (1)$$

式中  $F_i$ ——样本硬度

$F_{\max}$ ——硬度最大值

$t$ ——货架期储藏时间

$w$ ——常数,取  $w = 1$

不同货架期沙果软化指标的统计结果如表 2 所示。由表 2 可见,沙果的软化指标和硬度的变化趋势相似,即随着货架期时间的延长而逐渐降低;但和硬度的统计结果相比可知,软化指标的标准偏差和方差值都较小,变化趋势较为平缓,同时消除了在不同货架期间硬度重叠的现象,能更好地区分不同货架期的沙果样本。

表 2 货架期间沙果软化指标的变化及分析

Tab.2 Comparison of softening index of *Malus asiatica* during shelf life

货架期/d	最大值	最小值	平均值	标准偏差	方差值
1	1.00	0.600	0.769	0.111	0.013
4	0.57	0.380	0.455	0.094	0.010
7	0.36	0.042	0.284	0.073	0.005

#### 1.2.4 数据处理

数据统计、光谱数据的预处理、主成分回归分析、偏最小二乘法建模和线性判别分析在 Unscrambler 10.2 (CAMO ASA, Trondheim, Norway) 中进行。使用 SPSS 19.0 对模型进行优化处理。偏最小二乘支持向量机在 Matlab 软件中实现。

## 2 实验结果与讨论

### 2.1 不同货架期沙果的光谱分析

图 1 为不同货架期沙果光谱数据平均后的原始近红外光谱图。

从图 1 可见,不同货架期的沙果光谱图都在波长为 688、986、1 193、1 442、1 663、1 963、2 253 nm 处出现了吸收峰。根据 Polesello 等<sup>[12]</sup>的研究结果,水的吸收峰出现在波长 970、1 190、1 450、1 780、1 940 nm 等处。其中,波长 1 190 nm 为 O—H 的合频吸收峰,波长 1 940 nm 是 O—H 的伸缩振动的二级倍频,波

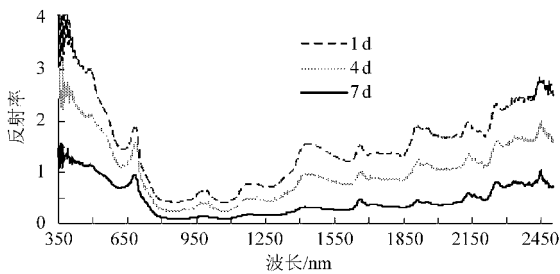


图1 不同货架期沙果样品的近红外光谱

Fig. 1 Near infrared absorbance spectra of *Malus asiatica* samples in different shelf life

长 1 450 nm 是 O—H 的伸缩振动的一级倍频。该研究结果与光谱图中波长 986、1 193、1 442、1 663、1 963 nm 处出现的吸收峰非常接近,由此证明:沙果在光谱图中的吸收峰主要是由于水的吸收所导致的。

果胶是水果成熟过程中一种和软化有很大关联的物质<sup>[13]</sup>。作为一种主要由甲氧基构成的高分子化合物,Boeriu 等<sup>[14]</sup>研究表明果胶的主要特征波长位于波长 2 248 nm 处,该处也正是 C—H 的伸缩振动的二级倍频。该研究结果与沙果光谱在波长 2 253 nm 出现的吸收峰非常接近。由此证明,对于软化这个物理指标,果胶对货架期沙果的光谱变化具有尤为重要的作用。

此外,从光谱图中可看出,处于不同货架期的沙果光谱图变化趋势是一致的,但是随着货架期的延长,沙果样本的内部化学成分和组织结构都发生了变化,导致光谱图的波峰强度有较大的差异。随着货架期的延长,沙果水分损失量不断增大,软化、粉质化现象严重,口感变差。当货架期为 7 d 时,沙果样本已经出现褐变和烂果现象,几乎失去了保鲜度。

## 2.2 基于软化指标建立模型

### 2.2.1 光谱预处理

为提高模型的预测能力,对原始光谱分别进行了以下预处理:9 点平滑处理(Savitzky-Golay, S-G)、多元散射校正(Multiplicative scatter correction, MSC)、变量标准化(Normalize, N)、变量的标准归一化(Standard normalized variate, SNV)、基线校正(Baseline correction, B)。

不同预处理方式下最佳模型选择标准是选择最少的因子数和最小的预测均方根误差(Root mean square error of prediction, RMSEP),因为更多的因子会导致过度拟合而降低模型的预测能力。本研究中基于沙果软化指标的各种数据预处理方法的最佳 PLS 模型如表 3 所示。

由表 3 可知,基于软化指标的近红外校正模型较优的光谱预处理方法为 SNV,此时该模型校正集

的决定系数  $R_c^2$  和校正均方根误差 RMSEC 分别为 0.814 和 0.619,预测集的决定系数  $R_p^2$  和预测均方根误差 RMSEP 分别为 0.800 和 0.799,优于其他模型。

表 3 不同光谱预处理方法下沙果软化指标的 PLS 模型校正结果

Tab. 3 Calibration results of PLS model about softening index based on different spectrum pre-treatment methods

方法	主成分数	$R_c^2$	RMSEC	$R_p^2$	RMSEP
原始光谱	6	0.807	0.702	0.764	0.844
S-G	9	0.899	0.521	0.778	0.818
MSC	4	0.803	0.718	0.765	0.852
N	5	0.805	0.705	0.767	0.855
SNV	6	0.814	0.619	0.800	0.799
B	8	0.802	0.488	0.753	0.856

### 2.2.2 异常值剔除及模型优化

为进一步提高模型的精度,本文使用 SPSS 19.0 软件绘制箱体图排除异常值,对模型进行优化,结果如图 2 所示。

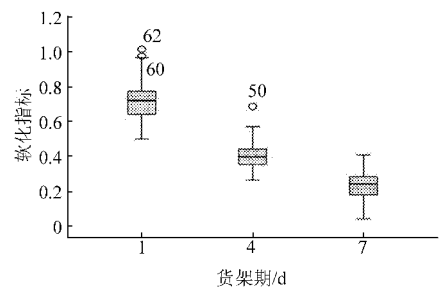


图 2 沙果样本软化指标中的异常点

Fig. 2 Outliers in softening index of *Malus asiatica* samples

由图 2 可知,在货架期为 1 d 和 4 d 时的样本中分别有 2 个和 1 个异常值被检出,应当在建模过程中予以剔除以免影响模型的准确性。

异常点剔除后该模型得到优化。其精度和准确性均有所提高,优化后模型的预测结果如图 3 所示。校正模型的决定系数  $R_c^2$  增加到 0.847,校正均方根误差增加为 0.069;其预测模型的决定系数  $R_p^2$  为 0.813,预测均方根误差为 0.085。证实软化指标与不同货架期的沙果光谱数据之间存在很高的相关性,可以利用软化指标对沙果的货架期进行分类。

## 2.3 不同货架期的沙果样本分类

将不同货架期的沙果根据软化指标并结合感官评价分为 3 大类,分类情况如表 4 所示。

由于沙果货架期较短,当货架期超过 7 d 以后其失重率将达到 5%,出现烂果坏果的现象,不可食用,已经失去研究意义。

为了便于分析,定义两个评价分类效果的指标<sup>[15]</sup>:正确识别率:指某一类(设第 1 类)的测试样

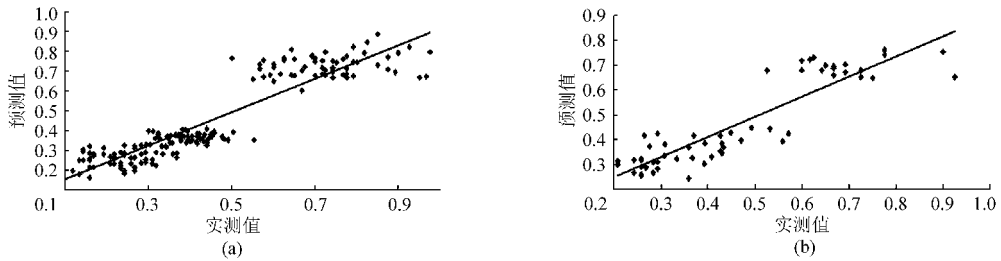


图3 沙果软化指标的校正集和预测集中实测值与预测值的关系

Fig. 3 Reference measured softening index and NIR predicted softening index in calibration set and validation set  
(a) 校正集 (b) 预测集

表4 不同货架期沙果的分类及感官评价

Tab. 4 Classification and sensory evaluation of *Malus asiatica* during shelf life

分类	货架期/d	软化指标	感官评价
第1类	1	0.60 ~ 1.00	新鲜多汁,无水分损失,质地较硬,无变软现象
第2类	2~4	0.38 ~ 0.57	出现轻微水分损失,开始出现质地变软的现象
第3类	5~7	0.042 ~ 0.36	果实脱水,表皮出现皱皮情况,有褐变出现,品质下降

本被分类为第1类样本的比率。正确拒识率:指非第1类样本被分类为非第1类样本的比率。

本文分别采用非线性(LS-SVM)和线性(PLS-LDA)两种方式建立分类模型。

### 2.3.1 LS-SVM

最小二乘支持向量机<sup>[16~17]</sup>(Least square support vector machines, LS-SVM)属于神经网络和非线性建模的范畴,在小样本回归、非线性、高维模式识别中均有优异的表现,同时避免了人工神经网络(ANN)中局部最优问题,求解速度快,因此本文采用该非线性建模方法分类,结果如表5所示。

表5 LS-SVM分类模型的校正和预测结果

Tab. 5 Results of calibration set and prediction set using LS-SVM

分类	主成分数	校正集			预测集		
		样本数	正确识别率/%	正确拒识率/%	样本数	正确识别率/%	正确拒识率/%
第1类	6	59	100	100	19	100	100
第2类	7	59	100	95.8	20	100	94.9
第3类	5	60	100	100	20	100	100

由表5可以看出,第1类和第3类样本的正确识别率和正确拒识率都达到了100%,只有第2类样本的正确拒识率较低。这可能是由于第2类沙果样本的货架期处于2~4d,该区间的沙果样本新鲜程度还较高,其失水情况和软化程度的变化不大,在分类时容易与第1类样本混淆,所以易出现错误拒识的情况。

### 2.3.2 PLS-LDA

PLS-LDA<sup>[18]</sup>(Partial least squares linear discriminant analysis)通过PLS提取主成分,再结合LDA进行线性判别分析,比其他线性判别方法往往

具有更好的模式识别能力,因此本文采用该线性建模方法分类,结果如表6所示。

表6 PLS-LDA分类模型的校正和预测结果

Tab. 6 Results of calibration set and prediction set using PLS-LDA

分类	主成分数	校正集			预测集		
		样本数	正确识别率/%	正确拒识率/%	样本数	正确识别率/%	正确拒识率/%
第1类	6	59	100	100	19	100	100
第2类	8	59	100	92.4	20	90.0	87.2
第3类	4	60	100	98.0	20	100	89.7

由表6可以看出,除第1类样本的正确识别率和正确拒识率达到100%之外,第2、3类样本的识别率和拒识率都较低。表明使用PLS-LDA对不同货架期的沙果样本进行分类时没有得到最优的分类效果。由此说明,线性分类模型不适用于对货架期样本的分类,这可能和本身数据存在的非线性或者和沙果中的化学物质组成有关。

## 3 结论

(1) 选用变量标准化(SNV)作为沙果样本的预处理方式建立的PLS模型,其校正模型的决定系数 $R_c^2$ 为0.847,校正均方根误差RMSEC为0.069;其预测模型的决定系数 $R_p^2$ 为0.813,预测均方根误差RMSEP为0.085。证实软化指标与不同货架期的沙果光谱数据之间存在很高的相关性,可以利用软化指标对沙果的货架期进行分类。

(2) 分别采用非线性的LS-SVM方式和线性的PLS-LDA方式建立不同货架期沙果的分类模型,结果表明,利用非线性的LS-SVM建立分类模



型得到的效果要优于线性的 PLS - LDA 分类模型, 达到了 94%, 说明对于不同沙果货架期的分类, 非校正集和预测集样品的正确识别率和正确拒识率都 线性的建模方式更为适用。

### 参 考 文 献

- 1 李艳梅, 赵福顺, 赵丽华. 沙果 123 苹果沈农 2 号苹果营养成分分析[J]. 农产品加工·学刊, 2009(2): 79 ~ 81.  
Li Yanmei, Zhao Fushun, Zhao Lihua, et al. Analysis of nutrients shaguo and 123 apple and 2-shennong apple[J]. Academic Periodical of Farm Products Processing, 2009(2): 79 ~ 81. (in Chinese)
- 2 余亚英, 袁唯. 食品货架期概述及其预测[J]. 中国食品添加剂, 2007(5): 77 ~ 79, 76.  
Yu Yaying, Yuan Wei. Summarize and forecast food shelf-life[J]. China Food Additives, 2007(5): 77 ~ 79, 76. (in Chinese)
- 3 邓云, 吴颖, 李云飞. 葡萄在贮藏和货架期间品质的预测模型[J]. 农业机械学报, 2006, 37(8): 93 ~ 97.  
Deng Yun, Wu Ying, Li Yunfei. Mathematical model of mass and firmness loss of grapes during storage and shelf life periods[J]. Transactions of the Chinese Society for Agricultural Machinery, 2006, 37(8): 93 ~ 97. (in Chinese)
- 4 Zerbini E P, Vanoli M, Grassi M, et al. A model for the softening of nectarines based on sorting fruit at harvest by time-resolved reflectance spectroscopy[J]. Postharvest Biology and Technology, 2006, 39(3): 223 ~ 232.
- 5 王加华, 陈卓, 李振茹, 等. 洋梨硬度的便携式可见/近红外漫透射检测技术[J]. 农业机械学报, 2010, 41(11): 129 ~ 133.  
Wang Jiahua, Chen Zhuo, Li Zhenru, et al. Evaluation of European pear (*Pyrus communis* L.) firmness based on portable vis/NIR transmittance technique[J]. Transactions of the Chinese Society for Agricultural Machinery, 2010, 41(11): 129 ~ 133. (in Chinese)
- 6 Infante R, Contador L, Rubio P, et al. Non-destructive monitoring of flesh softening in the black-skinned Japanese plums 'Angelino' and 'Autumn beaut' on-tree and postharvest[J]. Postharvest Biology and Technology, 2011, 61(1): 35 ~ 40.
- 7 杨芳. 近红外光谱分析技术应用[J]. 中国当代医药, 2011, 18(1): 15 ~ 16.
- 8 刘辉军, 孙斌, 陈华才. 基于近红外光谱的不同产地苹果货架期鉴别方法[J]. 光电工程, 2011, 38(5): 86 ~ 91.  
Liu Huijun, Sun Bin, Chen Huacai. Shelf-life identification method of different origin apples based on near infrared spectroscopy[J]. Opto-electronic Engineering, 2011, 38(5): 86 ~ 91. (in Chinese)
- 9 Sanchez M T, Perez-Marin D, Flores-Rojas K, et al. Use of near-infrared reflectance spectroscopy for shelf-life discrimination of green asparagus stored in a cool room under controlled atmosphere[J]. Talanta, 2009, 78(2): 530 ~ 536.
- 10 Xing J, Baerdemaeker D J. Fresh bruise detection by predicting softening index of apple tissue using VIS/NIR spectroscopy[J]. Postharvest Biology and Technology, 2007, 45(2): 176 ~ 183.
- 11 马庆华, 王贵禧, 梁丽松. 质构仪穿刺试验检测冬枣质地品质方法的建立[J]. 中国农业科学, 2011, 44(6): 1 210 ~ 1 217.  
Ma Qinghua, Wang Guixi, Liang Lisong. Establishment of the detecting method on the fruit texture of dongzao by puncture test[J]. Scientia Agricultura Sinica, 2011, 44(6): 1 210 ~ 1 217. (in Chinese)
- 12 Polesello A, Giangiacoimo R, Dull G G. Application of near infrared spectrophotometry to the nondestructive analysis of foods: a review of experimental results[J]. Critical Reviews in Food Science and Nutrition, 1983, 18(3): 203 ~ 230.
- 13 刘卉, 郭文川, 岳绒. 猕猴桃硬度近红外漫反射光谱无损检测[J]. 农业机械学报, 2011, 42(3): 145 ~ 149.  
Liu Hui, Guo Wenchuan, Yue Rong. Non-destructive detection of kiwifruit firmness based on near-infrared diffused spectroscopy[J]. Transactions of the Chinese Society for Agricultural Machinery, 2011, 42(3): 145 ~ 149. (in Chinese)
- 14 Boeriu C G, Stolle-Smits T, Van Dijk C. Characterisation of cell wall pectins by vibrational spectroscopy[J]. Food Science and Nutrition, 1998, 48(7): 257 ~ 266.
- 15 王徽蓉, 李卫军, 刘扬阳, 等. 基于遗传算法与线性鉴别的近红外光谱玉米品种鉴别研究[J]. 光谱学与光谱分析, 2011, 31(3): 669 ~ 672.  
Wang Huirong, Li Weijun, Liu Yangyang, et al. Study on discrimination of varieties of corn using near-infrared spectroscopy based on GA and LDA[J]. Spectroscopy and Spectral Analysis, 2011, 31(3): 669 ~ 672. (in Chinese)
- 16 张淑娟, 张海红, 赵艳茹, 等. 鲜枣可溶性固形物可见/近红外光谱检测建模方法比较[J]. 农业机械学报, 2012, 43(3): 108 ~ 112.  
Zhang Shujuan, Zhang Haihong, Zhao Yanru, et al. Comparison of modeling methods of fresh Jujube soluble solids measurement by NIR spectroscopy[J]. Transactions of the Chinese Society for Agricultural Machinery, 2012, 43(3): 108 ~ 112. (in Chinese)
- 17 Niu X Y, Zhao Z L, Jia K J, et al. A feasibility study on quantitative analysis of glucose and fructose in lotus root powder by FT-NIR spectroscopy and chemometrics[J]. Food Chemistry, 2012, 133(2): 592 ~ 597.
- 18 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 CSAE, 2011, 27(8): 350 ~ 354. (in Chinese)