

基于电子舌技术的鱼新鲜度定性、定量分析

韩方凯, 黄星奕, 穆丽君, 顾海洋

(江苏大学食品与生物工程学院, 江苏镇江 212013)

摘要: 为实现鱼新鲜度的快速定性、定量分析。采用电子舌技术对 4 °C 下不同冷藏天数的鲳鱼进行检测。同时测量鲳鱼体内挥发性盐基氮 (Total volatile basic nitrogen, TVB-N) 含量及细菌总数 (Total viable count, TVC)。对电子舌数据进行分析处理, 构建了 K 最近邻 (K-nearest neighbor, KNN) 判别模型和 BP 人工神经网络 (Back-propagation artificial neural network, BP-ANN) 模型定性评价鲳鱼新鲜度。结果显示, KNN 模型的训练集、测试集识别率分别为: 99.11% 和 98.21%; BP-ANN 模型的训练集、测试集识别率分别为: 92.86% 和 91.07%。构建了电子舌数据和 TVB-N 及 TVC 之间的支持向量机回归模型对鲳鱼新鲜度进行定量评价, 独立样本检验结果显示, 对 TVB-N 及 TVC 的预测, 支持向量机回归模型的预测值和实测值的相关系数分别为: 0.9727 和 0.9457, 预测均方根误差分别为 2.8×10^{-4} mg/g 和 0.052 log(CFU/g)。可见三种模型均能达到较好的效果。研究表明: 电子舌技术在鱼新鲜度的快速定性、定量评价中具有很大的潜力。

关键词: 电子舌; 鱼新鲜度; K 最近邻; BP 神经网络; 支持向量机回归

文章编号: 1673-9078(2014)7-247-251

Quantitative and Qualitative Analysis of Fish Freshness based on Electronic Tongue Technique

HAN Fang-kai, HUANG Xing-yi, MU Li-jun, GU Hai-yang

(School of Food and Biological Engineering, Jiangsu University, Zhenjiang 212013, China)

Abstract: In order to quantitatively and qualitatively evaluate the fish freshness, an electronic tongue was employed to detect the pomfret stored at 4 °C for different days. The total volatile basic nitrogen (TVB-N) and total viable count (TVC) of the fish samples were detected concurrently. K-nearest neighbor (KNN) model and back-propagation artificial neural network (BP-ANN) model were built to assess the freshness of the fish. Results showed that identification rate of training set and prediction set of KNN model were 99.11% and 98.21% respectively. While, the identification rate of training set and prediction set of BP-ANN model were 92.81% and 91.07% respectively. Support vector machine regression (SVR) model was established between the electronic tongue data and TVB-N as well as TVC for quantitative determination. The correlation coefficients between SVR predicted and measured TVB-N and TVC values were respectively 0.9727 and 0.9457, and root mean square error of prediction were 2.8×10^{-4} mg/g and 0.052 log(CFU/g), respectively. The overall results sufficiently demonstrate that the electronic tongue technique combined with appropriate pattern recognition method has a great potential to quantitative and qualitative evaluation of fish freshness rapidly.

Key words: electronic tongue; fish freshness; K-nearest neighbor; back-propagation artificial neural network; support vector machine regression

鱼肉味道鲜美, 营养丰富, 是深受消费者喜爱的动物性食品。随着生活水平的提高, 人们更加注重鱼

收稿日期: 2013-11-25

基金项目: 国家自然科学基金资助项目 (31071549); 公益性行业 (农业) 科技专项 (201003008-04); 江苏省高校优势学科建设工程资助项目; 江苏省普通高校研究生科研创新计划项目 (GXZZ13_0698)

作者简介: 韩方凯 (1989-), 男, 博士研究生, 研究方向为农产品质量检测与控制

通讯作者: 黄星奕 (1963-), 女, 博士生导师, 教授, 主要从事农产品无损检测技术研究

类产品的食用安全性、营养价值和风味品质。新鲜度是反映鱼品质的一项重要指标, 在产、销过程中鱼新鲜度的快速、准确评价有着重要的科学意义和实践价值。

目前, 世界各国主要采用感官评价和理化指标检测的方法对鱼新鲜度及食用安全性进行整体评价。感官评价实用性强, 但是需要专门培训的测评小组, 操作费时费力, 还受测评小组人员身体和心理状态的影响, 有较强的主观性, 一致性差。表征鱼新鲜度的理化指标主要有挥发性盐基氮、三甲胺、K 值等, 理化

指标检测就是利用理化分析的方法测定这些指标的相对含量。该方法客观可信,但操作复杂,耗时较长。

因此,迫切需要开发新型的鱼新鲜度检测方法。电子舌是一种分析、识别样品“味道”的新型智能感官仪器,主要由传感器阵列和模式识别系统构成。与常用的分析仪器相比,电子舌的输出结果并不是某一种组分的精确含量,而是对样品溶液理化信息的整体响应,现已被广泛应用于食品品质检测与监控^[1-3]、环境监测^[4]等领域。研究者们采用电子舌技术对鱼肉的新鲜度定性评价进行了初步研究^[5-6]。但这些研究均未构建明确的预测模型用于未知鱼样新鲜度的定性、定量分析。本研究采用电子舌技术对4℃下不同冷藏天数的鲳鱼进行检测,同时采用半微量定氮法和标准平板计数法测定样本的挥发性盐基氮含量(Total volatile basic nitrogen, TVB-N)和细菌总数(Total vital count, TVC)。分别采用了KNN线性分类算法和BP-ANN非线性分类算法对电子舌数据进行处理,构建判别模型对鲳鱼新鲜度进行定性分析。以TVB-N和TVC作为鲳鱼新鲜度的量化指标,采用支持向量机回归(Support vector machine regression, SVR)建立电子舌传感器数据和TVB-N及TVC之间的相关关系,构建预测模型对鲳鱼新鲜度进行定量分析。

1 材料与方法

1.1 样本准备与预处理

大型超市购买0.25 kg左右的鲳鱼168尾,加冰运回实验室,迅速擦干表面水渍、储藏于4℃冰箱内,待试验取用。

进行电子舌检测之前要对样品进行预处理,过程如下:取鱼背脊肉进行打浆,之后取10.00 g肉样加蒸馏水100 mL放入匀浆机进行匀浆2 min。匀浆液经高速离心机离心10 min,转速10000 r/min。离心之后取清液30 mL定容至100 mL容量瓶中,备用。

1.2 试验方法

1.2.1 电子舌检测

试验采用法国 α -AstreeII型电子舌检测装置,该电子舌系统主要由自动进样器、味觉传感器阵列、电子舌主机和计算机组成。味觉传感器将样品溶液中的化学物质信息转换成电信号,由电子舌主机采集信号数据并将数据传到计算机中进行分析处理。电子舌在测量前需要经过初始化、校准、诊断的过程,以确保数据的可靠性。试验中设置每个样品数据采集时间为120 s,每秒采集1组数据,取最后10 s测量值的平均

值作为该样品的电子舌数据。预实验结果表明,传感器响应强度在2~3次之后趋于稳定。因此,设置每个样品测量6次,计算后三次测量值的平均值作为该样品的一组数据。由于鲳鱼在4℃下7 d之后明显腐败,7 d之后不再进行试验,这样每天随机拿出24尾鲳鱼用于试验,连续7 d,最后得到168组数据。

1.2.2 TVB-N测定

TVB-N是评定鱼新鲜度的主要理化指标^[7],采用水产行业标准中规定的半微量定氮法^[8]测量。测量时,每个样品做三次平行实验,取其算术平均值作为该样品的TVB-N值。试验条件与电子舌检测试验保持一致,每天随机选取20尾鲳鱼用于检测,七天共得到140个样本的数据。

1.2.3 TVC测定

TVC是评定鱼新鲜度的另一主要指标^[7],采用国标中规定的标准平板计数法^[9]进行检测。测量时,每个样品做三次平行实验,取其算术平均值作为该样品的TVC值。试验样本同TVB-N测定的样本,7 d共得到140个样本的数据。

1.3 数据处理方法

研究首先对电子舌数据进行主成分分析(Principal Component Analysis, PCA),将PCA结果作为模型的输入,分别构建KNN和BP-ANN判别模型用于鱼新鲜度的定性分析;以TVB-N和TVC作为鲳鱼新鲜度的量化指标,采用SVR建立电子舌传感器数据和TVB-N及TVC之间的相关关系,以实现鲳鱼新鲜度的定量评价。其中,PCA、KNN、SVR在MATLAB 7.6软件上实现,BP-ANN在NeuroShell2软件上实现。

2 结果与分析

2.1 电子舌传感器数据的主成分提取

由电子舌的原理决定,传感器具有非专一性和弱选择性,而试验样本成分复杂,不同的传感器对样本中的组分具有交互敏感性,造成各传感器的响应值间存在一定相关性,从而使变量间产生冗余信息。这些冗余信息参与模型建立,会降低模型性能。主成分分析^[10]是模式识别中最常用的数据降维的方法,它是将原变量进行转换得到少数几个综合变量(即主成分),使这些综合变量能够概括原变量的大部分信息,这些综合变量是原变量的线性组合且互不相关,即正交。

电子舌传感器数据经过主成分分析后,前3个主成分的因子得分如图1所示。以前3个主成分因子得分向量分别作为X、Y、Z轴,建立三维视图。由图1

可以看出 7 d 中表征鲷鱼新鲜度信息的电子舌传感器数据的大致趋势。前 3 个主成分因子的贡献率分别为 86.41%、9.29%、2.25%，它们的累计贡献率为 97.95%，大于 85%。因此提取出来的这 3 个主成分能够表达原始数据的绝大部分信息。

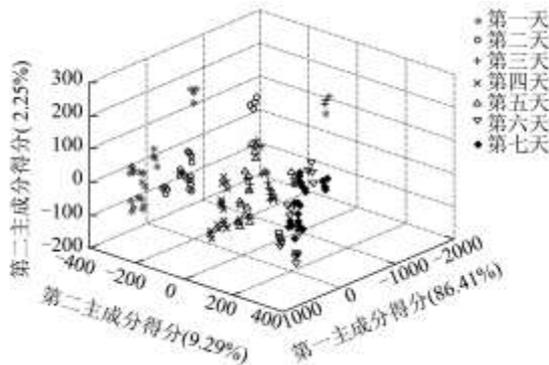


图 1 鲷鱼样本电子舌数据前三个主成分因子得分

Fig.1 Principal component of electronic tongue data of pomfret samples

2.2 鲷鱼新鲜度的 KNN 分类分析

最近邻分类算法依据最近距离的基本思想对待识样本进行分类，是最简单的机器学习算法之一。对于本研究中的 7 分类问题，假设有训练集样本：

$$\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$$

其中 x_i 是 7 维向量， $y_i \in \{1, 2, \dots, 7\}$ 是类别标志。对于一个待判样本 $x \in R^N$ ，若

$$\text{dist}(x, x_{i_0}) = \min \{ \text{dist}(x, x_1), \text{dist}(x, x_2), \dots, \text{dist}(x, x_n) \}$$

其中函数 $\text{dist}(y, z)$ ，表示矩阵 y 和 z 之间的欧氏距离。按照最近邻算法法则， x 当属 x_{i_0} 类。KNN 分类算法就是基于一个待判样本在特征空间中的 k 个最邻近的样本中的大多数属于某一个类别，则该样本也属于这个类别的思路进行分类的。KNN 算法中，最邻近的样本个数即 K 值并不是固定的，它的选择以模型识别率最高来进行。

在进行 KNN 分类分析时，随机选取样本总数的 2/3 即 112 个样本作为训练集，其余 56 个样本作为预测集，目标分类为七类即同一天检测的样本数据为一类。建模时，考虑了不同主成分数和最近邻样本个数对模型的影响，采用留一法交互验证的方式对这两个参数进行优化，以期得到最佳模型。图 2 显示主成分因子数 (principal components, PCs) 分别为 2 和 3，最近邻样本数分别为 1、3、5、7、

9 时的训练集正确识别率的比较。由图可以看出当 PCs=3, K=1 时的 KNN 判别模型训练集的正确识别率达到最高，为 99.11%，此时测试集的正确识别率达到 98.21%，模型可达到理想的判别效果。

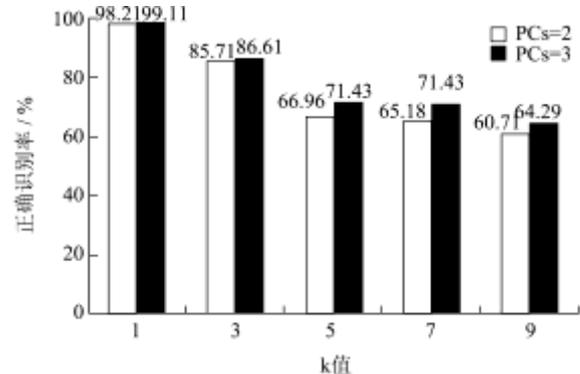


图 2 不同 PCs 和 K 值的 KNN 模型训练集识别率

Fig.2 Discrimination rates of training set of KNN model based on different PCs and K values

2.3 鲷鱼新鲜度的 BP-ANN 识别

人工神经网络是模拟人脑进行工作的，其中以误差反传为数学模型的前向多层神经网络即 BP-ANN 在模式识别和分类、非线性映射、特征提取等领域获得了成功应用^[4]。假设有 n 个输入的学习样本 p_1, p_2, \dots, p_n ，其对应的希望输出为 T_1, T_2, \dots, T_n ，BP-ANN 就是以网络的实际输出 q_1, q_2, \dots, q_n 与 T_1, T_2, \dots, T_n 之间的误差沿反向逐层传递并修正各层之间的连接权值，使网络输出层的误差平方和达到设定要求。

表 1 不同主成分因子下 BP-ANN 模型判别结果

Table 1 Discrimination result of BP-ANN with different PCs

主成分数	训练集			测试集		
	样本数	正确分类数	正确分类率/%	样本数	正确分类数	正确分类率/%
2	112	96	85.71	56	48	85.71
3	112	104	92.86	56	51	91.07
4	112	110	98.21	56	54	96.43
5	112	99	88.39	56	48	85.71
6	112	110	98.21	56	51	91.07

BP-ANN 建模时，从所有的样本中随机抽取训练样本 112 个，其余 56 个组成预测样本，目标分类为 7 类即同一天检测的样本数据为一类。研究采用 3 层 BP-ANN 网络拓扑结构：输入层、隐含层、输出层。对电子舌传感器数据主成分分析之后，以得到的主成分得分作为 BP-ANN 模型的输入，则输入层神经元数为主成分个数，输出层神经元数为 1 即类别。输入层到隐含层的传递函数采用正切 S 形函数，隐含层到输出层的传递函数采用 logistic 函数，初始化网络权值为

0.3, 学习因子为 0.1, 动态因子为 0.1, 目标误差为 0.001, 最大训练次数为 2000。用上述参数训练网络模型, 采用留一法交互验证的方式, 通过改变主成分数来优化模型。表 1 给出了主成分个数为 2~6 时对模型的正确识别率的影响, 从表中可以看出当主成分数为 4 时, 此时模型的训练集和预测集正确率分别为 98.21% 和 96.43%, 可达理想效果。

2.4 KNN 和 BP-ANN 比较

从 KNN 模型和 BP-ANN 模型的训练集和测试集的判别准确率可以看出, 两个模型的预测效果均较好, 可用于未知鱼样新鲜度的定性判别。同时也可以看出, 作为典型线性分类算法的 KNN 和具有良好非线性处理能力的 BP-ANN 在处理电子舌数据时性能同样优越, 甚至稍好。这主要是由于对电子舌传感器数据进行主成分分析后, 以主成分得分作为模型的输入, 极大地降低了原始变量之间的非线性关系且剔除了大量的冗余信息。

2.5 鲳鱼新鲜度的 SVR 定量分析

采用电子舌传感器数据为自变量, TVB-N 及 TVC 为因变量分别建立 SVR 模型。随机选取样本数的 3/5 即 84 个样本为训练集, 其余 56 个样本组成测试集。研究采用径向基核函数作为所建 SVR 模型的核函数。径向基函数为:

$$K(x, x_i) = \exp(-g \|x - x_i\|^2), g > 0 \quad (1)$$

支持向量机中, 为了解决低维空间线性不可分的问题经过核函数映射之后变成高维空间近似线性可分的问题而引入了松弛变量。松弛变量在目标函数中的系数称为惩罚因子, 用 c 来表示。 c 用于控制模型复杂度和逼近误差的折中, c 越大则对数据的拟合程度越高, 但泛化能力将降低。由 SVR 理论可知, 径向基核函数中的参数 g 以及惩罚因子 c 的选择对模型的性能有很大的影响。因此采用网格划分和留一法交互验证的方式, 以均方误差 (Mean square error, MSE) 最小为目标进行参数寻优。在寻找到的最优参数下进行建模。

在进行参数寻优之前, 先对自变量和因变量进行列归一化, 将全部的原始数据规整到 [0, 1] 之间。以自变量电子舌数据为例, 归一化方法如下: 设输入的特征向量为

$$x = (x_1, x_2, \dots, x_m), m=7$$

第 j 个样本的特征向量为

$$x(j) = (x_1(j), x_2(j), \dots, x_m(j))$$

其中 $j = 1, 2, \dots, n, n=168$ 。

归一化公式为: $y_i(j) = \frac{x_i(j) - x_i(\min)}{x_i(\max) - x_i(\min)}$, $i = 1, 2, \dots, m; j = 1, 2, \dots, n$ 。

$$y_i(j) = \frac{x_i(j) - x_i(\min)}{x_i(\max) - x_i(\min)}$$

在进行参数寻优的过程中, 设置 c 和 g 的范围均为 $[2^{-8}, 2^8]$, 步长为 0.5。结果显示, 对于 TVB-N 含量的拟合与预测, c 和 g 最优值为 1 和 0.7011, 此时的 MSE 为 6.1×10^{-4} mg/g; 对于 TVC 的拟合与预测, c 和 g 最优值为 4 和 0.3536, 此时的 MSE 为 0.046 log(CFU/g)。在最优的参数条件下, 构建用于 TVB-N 和 TVC 拟合与预测的 SVR 模型, 结果如图 3 和图 4 所示。

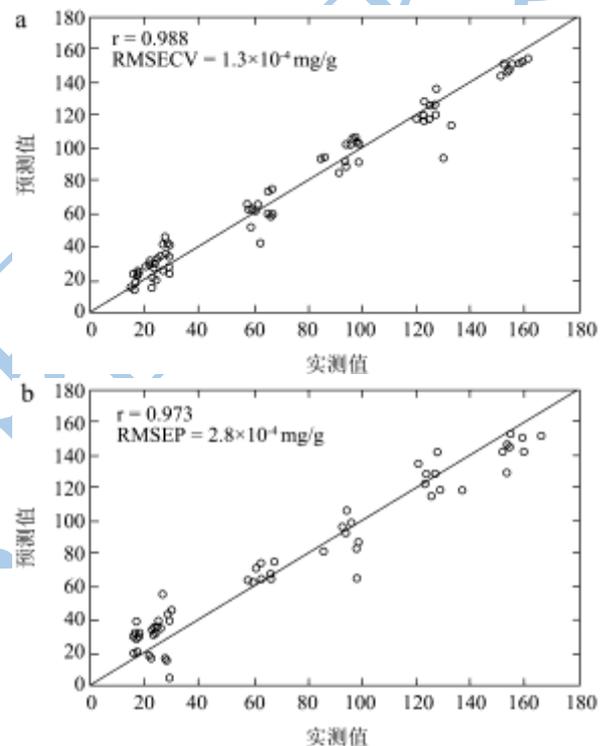


图 3 训练集 (a) 和预测集 (b) TVB-N 含量的 SVR 模型预测值与实测值的相关曲线

Fig.3 Correlation between measured TVB-N values and prediction value by SVR in training set (a) and prediction set (b)

注: a: 训练集, b: 预测集。

所建的 SVR 模型可以看出, 电子舌技术可以用于鱼体内 TVB-N 含量及 TVC 的预测, 从而实现鱼新鲜度的定量分析。这主要是由于, 随着鱼贮藏时间的延长, 鱼体组织 (脂肪、蛋白质、碳水化合物等) 在微生物和酶的作用下逐渐的分解为小分子的化合物^[12]。这些化合物中的氨, 小分子的胺类、酮类、醛类, 有机酸类, 醇类, 都有很好的水溶性。故选择蒸馏水和鱼肉一起匀浆, 匀浆液中包含鱼腐败过程中产生的大部分物质, 这些物质和鱼体内 TVB-N 及 TVC 存在一

定量的关系^[13-14], 这些物质在匀浆液的中种类和含量(化学信息)可以表征鱼的腐败程度。而电子舌传感器敏感薄膜上的敏感材料可以和匀浆液组分分子之间产生相互作用(氢键作用力、范德华作用力等)使膜电势与参比电极电势之差发生变化并作为最终输出信号, 实现对试液组分的分析。因此, 电子舌可以用于鱼体内 TVB-N 含量及 TVC 的预测。

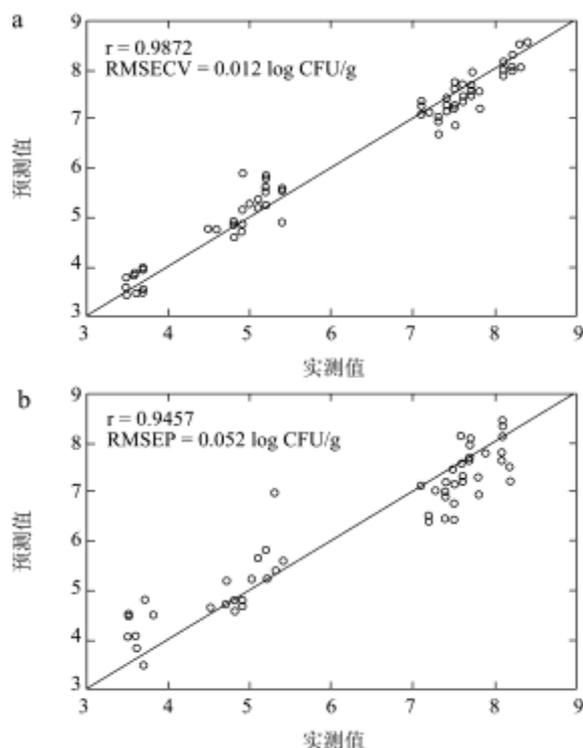


图4 训练集(a)和预测集(b) TVC 的 SVR 模型预测值与实测值的相关曲线

Fig.4 Correlation between measured TVC values and prediction values by SVR in training set (a) and prediction set (b)

注: a: 训练集, b: 预测集。

3 结论

电子舌作为一种新型的现代化智能仿生系统, 在食品检测等领域中的应用越来越多。研究采用电子舌技术对 4℃ 下不同冷藏天数的鲳鱼进行检测, 分别构建了 KNN 判别模型和 BP-ANN 预测模型对鲳鱼新鲜度进行定性判别。从模型结果来看, 两个模型都比较稳定, 效果较好。同时建立了电子舌传感器数据和鲳鱼体内 TVB-N 含量及 TVC 之间的 SVR 模型。结果表明, SVR 模型具有较好的预测效果, 能够对鲳鱼 TVB-N 含量及 TVC 进行预测, 从而实现鲳鱼新鲜度的定量分析。研究采用电子舌仪器对鲳鱼进行快速检测, 并结合适当的模式识别方法, 建立数学模型, 对不同新鲜度的鲳鱼样本进行定性、定量分析, 比较客

观、准确, 这为电子舌在鱼品质控制方面进行了探索性研究, 具有重要的参考价值。

参考文献

- [1] Kamalika Tiwari, Bipan Tudu, Rajib Bandyopadhyay, et al. Identification of monofloral honey using voltammetric electronic tongue [J]. Journal of Food Engineering, 2013, 117: 205-210
- [2] L Nuñez, X Cetó, M I Pividori, et al. Development and application of an electronic tongue for detection and monitoring of nitrate, nitrite and ammonium levels in waters [J]. Microchemical Journal, 2013, 110: 273-279
- [3] 孙红梅,张春晖,李侠,等.鸡骨素及其酶解液 Maillard 反应产物滋味成分研究[J].现代食品科技,2013,29(8):1872-1877
SUN Hong-mei, ZHANG Chun-hui, LI Xia, et al. Taste compounds in Maillard reaction products of chicken bone extract and its enzymatic hydrolysate [J]. Modern Food Science and Technology, 2013, 29(8): 1872-1877
- [4] Inmaculada Campos, Miguel Alcañiz, Daniel Aguadod, et al. A voltammetric electronic tongue as tool for water quality monitoring in wastewater treatment plants [J]. Water Research, 2012, 46(8): 2605-2614
- [5] 韩剑众,黄丽娟,顾振宇,等.基于电子舌的鱼肉品质及新鲜度评价[J].农业工程学报,2008,24(12):142-144
HAN Jian-zhong, HUANG Li-juan, GU Zhen-yu, et al. Evaluation of fish quality and freshness based on the electronic tongue [J]. Transactions of the CSAE, 2008, 24(12): 141-144
- [6] Luis Gil, José M. Barat, Isabel Escriche, et al. An electronic tongue for fish freshness analysis using a thick-film array of electrodes [J]. Microchim Acta, 2008, 163: 121-129
- [7] G Olafsdóttir, E Martinsdóttir, J Oehlenschläger, et al. Methods to evaluate fish freshness in research and industry [J]. Trends in Food Science & Technology, 1997, 8(8): 258-265
- [8] SC/T 3032-2007,水产品中挥发性盐基氮的测定[S]
SC/T 3032-2007, Determination of the total volatile basic nitrogen in fishery products [S]
- [9] GB 4789.2-2010,食品微生物学检验 菌落总数的测定[S]
GB 4789.2-2010, Food microbiological examination: Aerobic plate count [S]
- [10] Hongmei Chen, Yucai Ning, Xudong Sun. Production safety evaluation model based on principal component analysis [J]. Procedia Engineering, 2011, 26: 1949-1955
- [11] Hongfei Xiao, Yunli Tian. Prediction of mine coal layer

- spontaneous combustion danger based on genetic algorithm and BP neural networks [J]. *Procedia Engineering*, 2011, 26: 139-146
- [12] Alberto Pedrosa-Menabrito, Joe M regenstein. Shelf-life extension of fresh fish-a review part I-Spoilage of fish [J]. *Journal of Food Quality*, 1988, 11: 117-127
- [13] Owen Fraser, Sam Sumar. Compositional changes and spoilage in fish-an introduction [J]. *Nutrition & Food Science*, 1998, 5: 275-279
- [14] Owen Fraser, Sam Sumar. Compositional changes and spoilage in fish (part II)-microbiological induced deterioration [J]. *Nutrition & Food Science*, 1998, 6: 325-329

现代食品科技