

高光谱结合波长选择算法串联策略检测调理牛排新鲜度

孙宗保, 王天真, 刘小裕, 邹小波*, 梁黎明, 李君奎, 牛增, 高云龙

江苏大学食品与生物工程学院, 江苏 镇江 212013

摘要 生鲜调理牛排超过货架期时,其散发的腐败气味易被调料气味掩蔽,使消费者难以分辨。挥发性盐基氮(TVB-N)是表征肉品新鲜度的有效指标。由于测定 TVB-N 含量的化学方法繁琐耗时,利用高光谱对生鲜调理牛排中 TVB-N 含量进行预测,并讨论了不同波长选择算法所建模型的预测效果。分别在第 0, 2, 4, 6, 8 天取出制备好的生鲜调理牛排,进行高光谱数据采集和 TVB-N 含量测定。采用 1st Der, 2nd Der, MC, MSC, SG 和 SNVT 六种光谱预处理方法对光谱数据进行预处理,并建立偏最小二乘模型(PLS)优选出最佳预处理方法。采用竞争性自适应重加权算法(CARS)、变量组合集群分析法(VCPA)、间隔随机蛙跳(iRF)、iRF-CARS、iRF-VCPA 等方法对预处理后的光谱数据进行波长选择,建立特征波长下的预测模型。将 CARS 和 VCPA 重复运行 50 次考察其稳定性,并选择频次较高波长建模与单次运行比较。结果表明:在六种光谱预处理方法中,1st Der 为最佳预处理方法。CARS 和 VCPA 单次运行时分别选择了 21 和 11 个波长,其中 VCPA 选择波长建模预测效果更好,模型的 R_c 和 R_p 分别为 0.944 和 0.931, RMSECV 和 RMSEP 分别为 1.12 和 $1.28 \text{ mg} \cdot (100 \text{ g})^{-1}$ 。统计 CARS 和 VCPA 重复运行 50 次时各波长被选择频次,结果表明 VCPA 因其二进制矩阵采样法(BMS)为每个变量提供相同的采样机会而有更好的稳定性。同时发现两种方法有共同的高频次波长:694.9, 696.6, 761.8, 763.5, 811.5 和 813.3 nm 等。将波长被选频次降序排列,选择与单次运行数量相同的较高频次的波长建模,所得模型性能较差。将 iRF 分别与 CARS 和 VCPA 联用,其中 iRF-CARS 表现出较强的预测能力,选择了 24 个波长建模,模型的 R_c 和 R_p 分别为 0.966 和 0.938, RMSECV 和 RMSEP 分别为 0.91 和 $1.22 \text{ mg} \cdot (100 \text{ g})^{-1}$ 。这说明将波长区间选择和波长点选择联用可以实现它们的优势互补。高光谱技术结合波长选择方法可以很好地预测调理牛排中 TVB-N 含量,研究可为波长选择算法联用策略和调理牛排新鲜度快速检测提供理论参考。

关键词 高光谱成像; 生鲜调理牛排; 挥发性盐基氮; 竞争性自适应重加权采样; 变量组合集群分析

中图分类号: O657.3 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2020)10-3224-06

引言

调理牛排是以牛排为原料,加入适量调味料和食品添加剂,经过切分、滚揉腌制等过程加工而成的肉制品,因其滋味鲜美、食用方便等特点而深受消费者喜爱。冷冻和冷藏是调理牛排的两种主要储运方式。冷藏的牛排能够较好的保留原有的滋味和营养,但同时货架期较短、价格也较高。由于调理牛排中加入了调味料进行腌制,当牛排超过货架期变得不新鲜时,散发的腐败气味易被调料气味掩蔽,使得消费者难以分辨其新鲜度。市场上存在售卖超过货架期的调理牛排的现象,食用了这种牛排将损害人体健康,生鲜调理牛排因

其更容易变质而更可能产生这一问题。

挥发性盐基氮(total volatile basic nitrogen, TVB-N)是由于微生物分解蛋白质产生的具有挥发性的氨和胺类等碱性含氮物质的总称,可以有效地表征肉品新鲜度。测定 TVB-N 含量的化学方法操作复杂耗时,且对样本有破坏性。肉品发生腐败时产生的 TVB-N 将会引起一定波长下光谱吸收强度的变化,通过吸收强度的大小结合化学计量学方法可以定量分析 TVB-N 含量。高光谱成像因其获得信息全面,快速无损等优点,已广泛应用于肉品品质检测。Velásquez 等^[1]利用高光谱对牛肉大理石花纹进行分类。首先在 528 nm 波长的高光谱图像下分割背景,然后在 440 nm 下结合决策树进行分级,模型的误差仅有 0.08%。Zheng 等^[2]利用高光谱成

收稿日期: 2020-02-28, 修订日期: 2020-06-07

基金项目: 国家重点研发计划项目(2016YFD0401104), 江苏高校优势学科建设工程项目和镇江市重点研发项目(SH2019019)资助

作者简介: 孙宗保, 1976 年生, 江苏大学食品与生物工程学院副教授 e-mail: zongbaos@163.com

* 通讯联系人 e-mail: zou_xiaobo@ujs.edu.cn

像预测了碎羊肉中的鸭肉含量, 并实现掺假含量分布可视化。谢安国等^[3]利用高光谱成像构建了调理牛肉在煎制过程中的品质可视化模型, 预测调理牛肉的水分和剪切力相关系数分别为 0.908 和 0.763, 表明了高光谱成像技术具备检测混有调料的复杂肉品的能力。但少有高光谱技术检测调理牛排的报道。

由于高光谱数据包含的信息量巨大且冗杂, 如何去除无效信息、筛选特征波长从而提升模型效率和精度成为高光谱技术应用的关键。本工作利用高光谱技术对生鲜调理牛排贮藏期间 TVB-N 含量进行预测, 采用了竞争性自适应重加权算法(competitive adaptive reweighted sampling, CARS)、变量组合聚类分析法(variables combination population analysis, VCPA)、间隔随机蛙跳(interval random frog, iRF)、iRF-CARS 和 iRF-VCPA 等波长选择算法并分析了 CARS 和 VCPA 的稳定性, 为波长选择算法联用策略和调理牛排新鲜度快速检测提供参考。

1 实验部分

1.1 样本制备

试验用牛肉和辅料均购于镇江麦德龙超市。牛肉为西冷部位, 全程冷链运输。调理牛排制作工艺: 原料肉修整→切片→配制腌制液→滚揉腌制→托盘密封包装。共制备了 150 个样本, 置于 4℃ 的冰箱内冷藏, 分别在第 0, 2, 4, 6 和 8 天每天取出 30 个样本, 编号后先进行高光谱图像采集, 再进行 TVB-N 含量测定。

1.2 高光谱数据采集

采用的高光谱成像系统如图 1 所示, 硬件部分主要包括 CCD 摄像机、光纤卤素灯、三轴精密电控平移台和计算机等。光谱采集参数: 摄像机曝光时间为 45 ms, 光谱范围为 431~962 nm, 光谱波长间隔为 0.858 nm。对获取的原始图像进行黑白板校正, 以减少采集过程中噪音的影响。

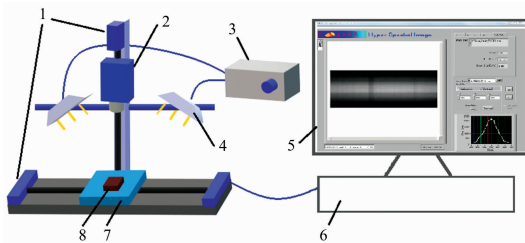


图 1 高光谱成像系统

- 1: 步进电机; 2: CCD 摄像机; 3: 光源系统; 4: 光纤卤素灯;
5: 计算机; 6: 三轴精密电控平移台; 7: 载物台; 8: 样品

Fig. 1 Hyperspectral imaging system

- 1: Stepper motor; 2: CCD camera; 3: Light source system;
4: Optical fiber halogen lamp; 5: Computer;
6: Three-axis precision electronically controlled translation stage;
7: Object stage; 8: Sample

1.3 TVB-N 含量测定

依照 GB 5009.228—2016 中的半微量定氮法对调理牛排

样本中 TVB-N 含量进行测定。

1.4 数据处理方法

1.4.1 光谱提取与预处理方法

利用 ENVI 软件选取每个样本中心附近 200 像素×200 像素的矩形区域作为感兴趣区域(region of interest, ROI), 计算光谱平均反射率作为光谱数据。采用一阶导数(first derivative, 1st Der)、二阶导数(second derivative, 2nd Der)、均值中心化(mean centering, MC)、多元散射校正(multiplicative scatter correction, MSC)、卷积平滑(Savitzky-Golay, SG)、标准正态变量变换(standard normal variate transformation, SNVT)六种预处理方法建立模型, 选取效果最佳的预处理方法进行后续数据处理。

1.4.2 特征波长选择方法

良好的波长选择方法可以简化模型, 提高模型预测性能。波长选择方法可以分为波长区间选择和波长点选择^[4-5]。波长区间选择方法以一组连续波长为基本单位进行波长筛选, 最终选择的变量为几组连续的波长区间组合, 所建立的模型有很好的解释性。波长点选择方法以单个波长为基本单位, 最终选择的变量呈离散分布, 具有选择变量少、建模预测效果好等优点。不同的波长选择方法各有其优缺点和适应性, 将这些波长选择方法联用, 可以结合它们的优势, 达到更优的效果。联用并非任意的排列组合, 需要基于一定的组合策略。通常情况下, 当两种波长选择方法联用时, 前一步方法进行粗略选择, 去除无信息变量, 同时不能漏掉关键变量; 后一步方法精准选择重要变量, 保留较少变量的同时, 取得较好的建模效果^[6]。波长区间选择和波长点选择符合上述特点, 具有一定互补性。基于以上分析, 本工作采用 CARS 和 VCPA 两种波长点选择方法, 并与波长区间选择方法 iRF 联用(iRF 在前), 比较它们的建模效果。

CARS 是一种基于偏最小二乘法(partial least square method, PLS)回归系数的波长选择方法, CARS 首先通过蒙特卡罗采样(Monte carlo sampling, MCS)随机抽取校正集的一部分样本建立 PLS 模型, 通过指数衰减函数(exponentially decreasing function, EDF)和自适应重加权采样(adaptive reweighted sampling, ARS)选取回归系数较大的波长, 选择建模时交互验证均方根误差(root mean square error of cross-validation, RMSECV)最小时对应波长^[7-8]。

VCPA 是一种基于模型集群分析(model population analysis, MPA)思想的新型变量选择方法。VCPA 采用二进制矩阵采样法(binary matrix sampling, BMS), 每个变量有同等被选中机会, 利用 EDF 确定保留波长的数量, 缩小变量空间, 计算出剩余变量间所有可能组合的 RMSECV, 选择 RMSECV 最小的变量组合作为最终建模数据^[9]。

iRF 是在随机蛙跳(random frog, RF)的基础上提出的一种新型波长区间选择方法。iRF 模拟一条服从稳态分布的马尔科夫链迭代计算每个波长变量被选择的频率, 对选择可能性最高的波长变量建立 PLS 模型, 选择模型误差最小的变量, 寻优搜索能力强^[10]。

2 结果与讨论

2.1 新鲜度指标测定结果

生鲜调理牛排贮藏期间 TVB-N 含量测定结果如图 2 所示。根据国标规定,肉制品 TVB-N 含量小于 $15 \text{ mg} \cdot (100 \text{ g})^{-1}$ 为新鲜状态。图中,TVB-N 含量第 0 天的平均值为 $8.13 \text{ mg} \cdot (100 \text{ g})^{-1}$,并随贮藏时间的延长不断增长。第 8 天的平均值为 $17.43 \text{ mg} \cdot (100 \text{ g})^{-1}$,已经超过标准规定。贮藏后期(4~8 d)的 TVB-N 含量增长速度大于贮藏前期(0~4 d),这是因为贮藏后期致腐微生物大量繁殖,加速了蛋白质的分解,产生的碱性含氮物质与有机酸结合成 TVB-N。与其他研究中^[11]相同包装条件下的牛肉 TVB-N 测定结果相比,调理牛排的 TVB-N 增长速度较慢。这是因为牛排在调理过程中加入了食用盐和复合磷酸盐,一定程度上抑制了致腐菌的生长。

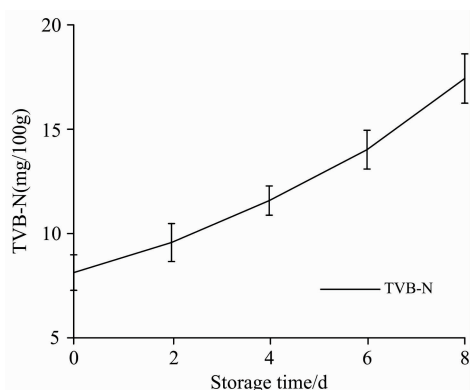


图 2 不同贮藏时间调理牛排 TVB-N 含量测定结果

Fig. 2 Results of TVB-N content in prepared steaks at different storage times

2.2 光谱数据预处理

图 3(a)和(b)分别为生鲜调理牛排的原始光谱图和平均光谱图。从平均光谱图可以看出,不同贮藏天数的样本光谱反射曲线趋势相近,且随着储藏时间的增加,光谱反射率增大。这是由于在贮藏过程中,调理牛排一些物理特征和内部化学成分发生了变化如水分流失、肉色变浅、蛋白质降解、脂肪氧化等,导致了光谱吸收的变化。光谱曲线在 550, 580, 760 和 960 nm 有明显吸收峰,550 和 580 nm 处主要与氧合肌红蛋白有关^[12],760 nm 处是肌红蛋白稳定吸收峰^[13],960 nm 附近主要与调理牛排水分中 O—H 伸缩振动的二级倍频吸收有关^[14]。

2.3 TVB-N 含量预测模型建立

2.3.1 全波段下 PLS 模型

将 150 个样本的预处理后的光谱和对应 TVB-N 数据采用随机分组的方式以 2 : 1 的比例划分为校正集和预测集。通过主成分分析法(principle component analysis, PCA)优选前十个主成分作为自变量,以 TVB-N 测量值作为因变量,建立 PLS 预测模型。模型的预测结果如表 1 所示,可以看出,光谱经 1st Der 处理后建立的模型预测效果最好,模型的

R_C 和 R_P 分别为 0.919 和 0.907, RMSECV 和 RMSEP 分别为 1.45 和 $1.52 \text{ mg} \cdot (100 \text{ g})^{-1}$ 。故对光谱数据进行 1st Der 处理后进行后续数据处理。

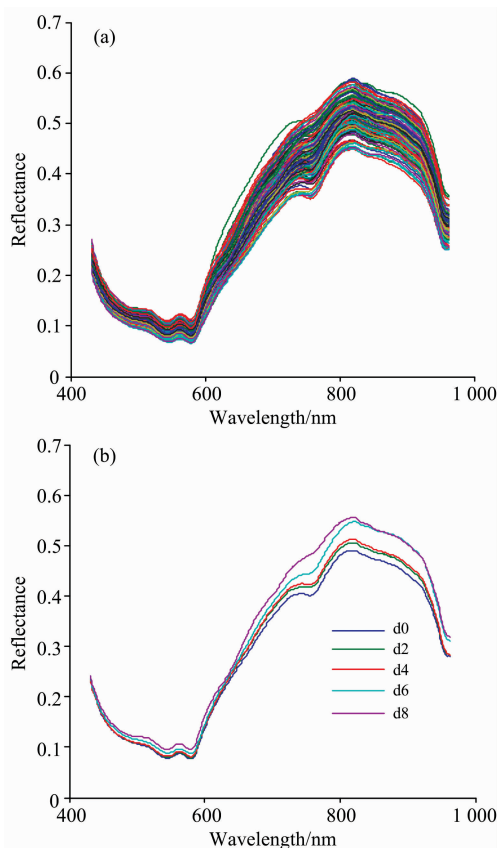


图 3 不同贮藏时间调理牛排原始光谱图(a)和平均光谱图(b)

Fig. 3 Raw (a) and average spectra (b) of prepared steaks at different storage times

表 1 不同预处理方法对 TVB-N 含量 [$\text{mg} \cdot (100 \text{ g})^{-1}$] 的预测结果

Table 1 Predictive results of TVB-N content [$\text{mg} \cdot (100 \text{ g})^{-1}$] based on different pre-processing methods

预处理方法	R_C	RMSECV	R_P	RMSEP
1 st Der	0.919	1.45	0.907	1.52
2 nd Der	0.901	1.29	0.882	1.43
MC	0.907	1.44	0.905	1.49
MSC	0.898	1.50	0.878	1.69
SG	0.911	1.44	0.896	1.46
SNVT	0.901	1.48	0.896	1.63

2.3.2 基于 CARS 的 PLS 模型

CARS 采样次数设置为 100 次,运行过程如图 4 所示。从图中可以看出,随着采样次数的增加,被选波长数量逐渐减少, RMSECV 先缓慢减小而后突然增加。采样次数为 59 时, RMSECV 最小,对应选择了 21 个特征波长。将选择的波长变量代替全波段模型的输入,建立 TVB-N 的 PLS 预测模型。模型的 R_C 和 R_P 分别为 0.932 和 0.927, RMSECV 和

RMSEP 分别为 1.24 和 1.29 $\text{mg} \cdot (100 \text{ g})^{-1}$ 。相比全波段光谱建模, CARS-PLS 大大减少了建模所需变量, 提高了模型预测效果。

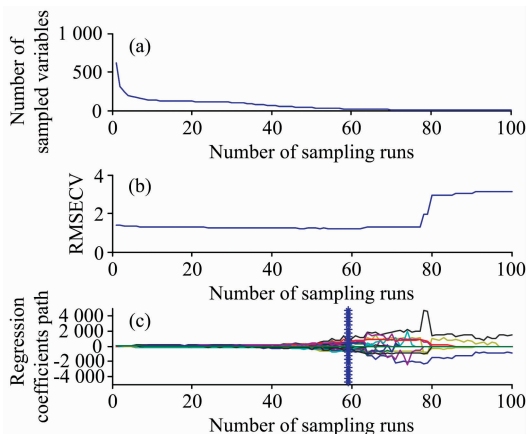


图 4 CARS 选择特征波长过程

Fig. 4 Process of selecting characteristic wavelengths by CARS

2.3.3 基于 VCPA 的 PLS 模型

VCPA 运行参数设置如下: 最佳子集占比 0.1, BMS 运行的次数设为 1 000, EDF 运行的次数设为 50, 剩余变量数目设为 14。先通过 BMS 对校正集数据进行 1 000 次采样得到 1 000 组变量组合, 在此过程中, 每个变量被选择的概率相同。对这些变量组合建立 PLS 模型, 取 RMSECV 最小的 100 组变量组合进行下一步 EDF 运算。在 EDF 运行过程中, RMSECV 变化如图 5 所示。可以看出, 随着 EDF 的重复运行, 特征空间缩小, RMSECV 整体呈下降趋势, 此时与 TVB-N 含量相关性不大的波长变量被删除, 剩下的变量被添加到最佳子集中。EDF 运行结束后, 计算选择的 14 个波长变量所有可能组合的 RMSECV, 并选取 RMSECV 最小的组合。最终选择了 11 个特征波长, 所建模型的 R_c 和 R_p 分别为 0.944 和 0.931, RMSECV 和 RMSEP 分别为 1.12 和 1.28 $\text{mg} \cdot (100 \text{ g})^{-1}$ 。相比于 CARS, VCPA 选择的变量更少, 同时模型的预测效果更优。

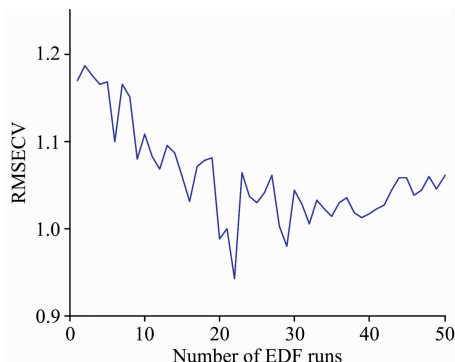


图 5 RMSECV 随 EDF 运行次数变化趋势

Fig. 5 Changes in RMSECV with the number of EDF runs

2.3.4 CARS 与 VCPA 的比较和稳定性分析

从前三小节的结果可知, 全光谱数据中包含很多与调理

牛排中 TVB-N 无关波长, 导致建立的 PLS 模型预测性能削弱。CARS 和 VCPA 都能在减少变量的同时提升模型预测效果。且相比于 CARS, VCPA 选择的变量更少, 预测效果更好。这是因为 VCPA 可以通过 EDF 压缩变量空间, 消除无关变量, 并能够考虑所有的变量组合效应。EDF 迭代运行过程中, MPA 思想保留了最佳的 10% 的通过 BMS 采样获得的子模型, 整个迭代过程类似于“优胜劣汰”的进化论。而保留变量少是因为 VCPA 最终选择变量从 EDF 运行后剩余的 14 个变量的所有组合中确定。

将 CARS 和 VCPA 重复运行 50 次, 以评价其稳定性。每个波长变量被选择频次如图 6 所示。从图中可以看出, CARS 选择的变量较分散, 有较多杂峰, 而 VCPA 选择变量相对集中, 稳定性更好。这是因为 CARS 的采样策略 MCS 随机选择变量, 每个变量被选择的机会是不确定的, 一些不重要变量可能会因为这个随机过程而被最终保留。而 VCPA 的 BMS 采样策略为每个变量提供相同的采样机会, 重要变量被最终选择的几率相对更大。对比两种方法选择频率较高的波长, 发现有一些共同的重要波长变量: 694.9, 696.6, 761.8, 763.5, 811.5 和 813.3 nm 等。760nm 处是能够反映肉品新鲜度的肌红蛋白的吸收峰^[13], 810 nm 是蛋白质中 C—H 键伸缩振动倍频吸收^[15], 而 TVB-N 正是由于肉中蛋白质被微生物分解而产生的^[16]。

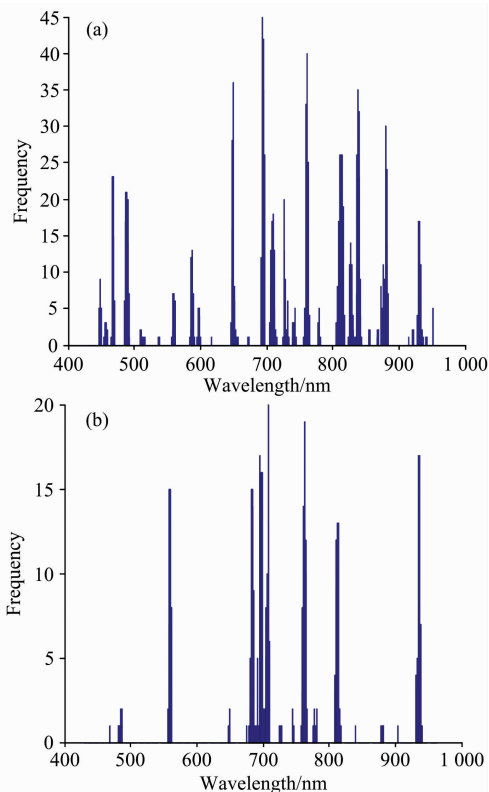


图 6 CARS (a) 和 VCPA (b) 重复运行中被选波长频次图

Fig. 6 Frequency of selected wavelengths by CAPS (a) and VCPA (b) in repeated runs

尝试选择 50 次重复运行过程中被选择频率较高的波长

进行建模,并与单次运行效果对比。在之前的模型中,CARS和VCPA分别选择了21和11个波长变量。按照频次排名选择同样数量的波长进行建模,得到的CARS-PLS模型的 R_C 和 R_P 分别为0.851和0.821, RMSECV和RMSEP分别为1.78和2.09 $\text{mg} \cdot (100 \text{ g})^{-1}$; VCPA-PLS模型的 R_C 和 R_P 分别为0.918和0.899, RMSECV和RMSEP分别为1.41和1.40 $\text{mg} \cdot (100 \text{ g})^{-1}$,模型效果远低于单次运行建模。考虑是因为虽然被选频次高的波长变量一定包含了与TVB-N含量有关的重要信息,但这些频次高的变量之间可能存在共线性,信息重复,而被舍弃的频次稍低的一些变量也包含了一些相关信息。而算法单次运行时基于一定的优化策略,虽然每次运行结果不一定相同,但最终选择的变量组合都能够取得较好的建模效果。

2.3.5 基于iRF-CARS和iRF-VCPA的PLS模型

将波长区间的粗选和波长点的精选联用,有望达到更好的建模效果。首先对全波段光谱数据进行iRF筛选。iRF运行参数设置如下:迭代次数为10 000,移动窗口宽度为20,初始变量集个数为50,节点处抽取变量个数为3。迭代运行后将599个间隔按被选概率排序,计算排名前1~10至1~599的组合的RMSECV,共计599个RMSECV,选择RMSECV最小时的波长间隔组合。最终选择的波长如图7所示。

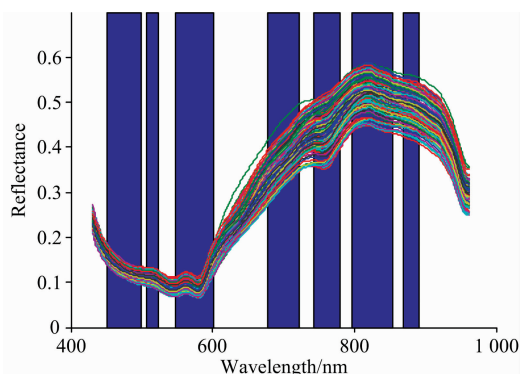


图7 iRF选择波长示意图

Fig. 7 Wavelengths selected by iRF

将iRF选择的变量作为新变量进行CARS和VCPA筛选。CARS和VCPA分别选择了24和7个波长。将所有波长选择后建立模型预测结果汇总至表2。从表中可以看出,相

比于全光谱PLS,iRF-PLS模型效果得到提升,但波长变量数过多。iRF-VCPA-PLS模型效果与VCPA-PLS相当,但波长数量减少到了7个。iRF-CARS-PLS模型效果最佳,模型的 R_C 和 R_P 分别为0.966和0.938, RMSECV和RMSEP分别为0.91和1.22 $\text{mg} \cdot (100 \text{ g})^{-1}$ 。串联策略取得较好效果,它充分利用了波长区间选择的强解释性,减少无效波长点对波长精选算法的干扰,帮助波长精选在较小的区间中进行最优选择。在本次建模结果中,相比于VCPA,CARS展现了更好的与iRF串联效果。不同波长选择下的模型的预测效果一定程度上依赖于样本集,因此这些波长选择联用方法需要在更多应用中检测其适用性。

表2 各波长选择方法对TVB-N含量 $[\text{mg} \cdot (100 \text{ g})^{-1}]$ 的预测结果

Table 2 Predictive results of TVB-N content $[\text{mg} \cdot (100 \text{ g})^{-1}]$ based on different wavelength selection methods

特征波长选择方法	波长数量	R_C	RMSECV	R_P	RMSEP
CARS	21	0.932	1.24	0.927	1.29
VCPA	11	0.944	1.12	0.931	1.28
iRF	329	0.936	1.21	0.925	1.40
iRF-CARS	24	0.966	0.91	0.938	1.22
iRF-VCPA	7	0.943	1.15	0.932	1.23

3 结论

利用高光谱成像对生鲜调理牛排贮藏过程中TVB-N含量进行预测,重点探究了波长选择方法对预测结果的影响。研究表明:(1)在所用六种光谱预处理方法中, 1^{st} Der为最佳预处理方法,CARS和VCPA单次运行时,VCPA所选波长数量更少,同时建模预测效果更好。(2)将CARS和VCPA重复运行50次,发现VCPA的稳定性更好。利用被选择频率较高的波长进行建模,模型性能较差。(3)将iRF分别与CARS和VCPA联用,CARS表现出更好的串联效果,iRF-CARS选择了24个波长,模型的 R_C 和 R_P 分别为0.966和0.938, RMSECV和RMSEP分别为0.91和1.22 $\text{mg} \cdot (100 \text{ g})^{-1}$ 。研究可为波长选择算法联用策略和高光谱检测调理牛排新鲜度提供理论参考。

References

- [1] Velásquez L, Cruz-Tirado J P, Siche R, et al. Meat Science, 2017, 133: 43.
- [2] Zheng X C, Li Y Y, Wei W S, et al. Meat Science, 2019, 149: 55.
- [3] XIE An-guo, KANG Huai-bin, WANG Fei-xiang, et al(谢安国, 康怀彬, 王飞翔, 等). Food & Machinery(食品与机械), 2018, 34(11): 20.
- [4] Yun Y H, Li H D, Deng B C, et al. TrAC Trends in Analytical Chemistry, 2019, 113: 102.
- [5] Guo Z M, Wang M M, Wu J Z, et al. Food Chemistry, 2019, 286: 282.
- [6] Yu H D, Yun Y H, Zhang W M, et al. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2020, 224: 117376.
- [7] Jiang H, Xu W D, Ding Y H, et al. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2020, 228: 117781.
- [8] Guo Z M, Wang M M, Akwasi A A, et al. Journal of Food Engineering, 2020, 279: 109955.

- [9] Ouyang Q, Yang Y C, Wu J Z, et al. *LWT*, 2020, 118: 108768.
- [10] Yun Y H, Li H D, Wood L R E, et al. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2013, 111: 31.
- [11] PAN Xiao-qian, ZHAO Yan, ZHANG Shun-liang, et al(潘晓倩, 赵燕, 张顺亮, 等). *Meat Research(肉类研究)*, 2016, 30(3): 15.
- [12] Van Beers R, Kokawa M, Aernouts B, et al. *Meat Science*, 2018, 136: 50.
- [13] HUANG Chang-ping, ZHU Xin-ran, ZHANG Chen-lu, et al(黄长平, 朱欣然, 张辰璐, 等). *Spectroscopy and Spectral Analysis(光谱学与光谱分析)*, 2018, 38(2): 552.
- [14] Talens P, Mora L, Morsy N, et al. *Journal of Food Engineering*, 2013, 117(3): 272.
- [15] Dai Q, Cheng J H, Sun D W, et al. *Food Chemistry*, 2016, 197: 257.
- [16] YANG Bin, CAO Yin-juan, YU Qun-li, et al(杨斌, 曹银娟, 余群力, 等). *Food Science(食品科学)*, 2019, 40(23): 199.

Detection of Prepared Steaks Freshness Using Hyperspectral Technology Combined With Wavelengths Selection Methods Combination Strategy

SUN Zong-bao, WANG Tian-zhen, LIU Xiao-yu, ZOU Xiao-bo*, LIANG Li-ming, LI Jun-kui, NIU Zeng, GAO Yun-long
School of Food and Biological Engineering, Jiangsu University, Zhenjiang 212013, China

Abstract When the freshprepared steaks exceed the shelf life, its putrid odour is easily masked by the seasoning odour, making it difficult for consumers to distinguish. Total volatile basic nitrogen (TVB-N) is an effective indicator of meat freshness. Due to the chemical method of TVB-N content determination was cumbersome and time-consuming, TVB-N content in freshlyprepared steaks were predicted by hyperspectral imaging technique in this study. Furthermore, the influence of different wavelengths selection algorithms on the prediction effect of the model was discussed. Fresh prepared steaks were taken out on day 0, 2, 4, 6, 8 respectively for hyperspectral data collection and TVB-N content determination. The spectral data were pretreated by six methods, first derivative (1st Der), second derivative (2nd Der), mean centering (MC), multiplicative scatter correction (MSC), Savitzky-Golay (SG), standard normal variate transformation (SNVT), which established the Partial least squares model of TVB-N content and obtained the optimum pretreatment method by comparison. After the optimum pre-processing method, the characteristic wavelengths were selected by competitive adaptive reweighted sampling (CARS), variables combination population analysis(VCPA), interval random frog (iRF), iRF-CARS, iRF-VCPA, the prediction effects of models were evaluated on different selection wavelengths methods. CARS and VCPA were run 50 times repeatedly to assess their stability, and higher frequency wavelengths were selected to establish the PLS model to compare with a single run. The results suggested that 1st Der was the best spectral pretreatment method among the six methods. 21 and 11 wavelengths were selected by CARS and VCPA single run respectively, and the prediction model by VCPA wavelengths selection method had better effect, with the R_c and R_p at 0.944 and 0.931 respectively, while RMSECV and RMSEP at 1.12 and $1.28 \text{ mg} \cdot (100 \text{ g})^{-1}$ respectively. The selected frequency for each wavelength when CARS and VCPA were run 50 times repeatedly indicated that VCPA had better stability because its binary matrix sampling method provided the same sampling opportunity for each variable. At the same time, it was found that the two methods have important common wavelengths: 694.9, 696.6, 761.8, 763.5, 811.5 and 813.3 nm, etc. The performance of the model established by higher frequency wavelengths was poor. IRF was combined with CARS and VCPA respectively for wavelengths selection, and iRF-CARS showed good predictive performance. The prediction model by iRF-CARS was established using 24 wavelengths, with the R_c and R_p at 0.966 and 0.938 respectively, while RMSECV and RMSEP at 0.91 and $1.22 \text{ mg} \cdot (100 \text{ g})^{-1}$ respectively. The results suggested that the combination of wavelengths interval selection and wavelengths point selection could realize their complementary advantages. The research suggested that hyperspectral imaging technique combined with wavelengths selection method can predict TVB-N content in prepared steaks well, which provides a theoretical reference for hybrid variable selection strategy and rapid detection of the freshness of prepared steaks.

Keywords Hyperspectral imaging; Freshprepared steaks; Total volatile basic nitrogen; Competitive adaptive reweighted sampling; Variables combination population analysis

* Corresponding author

(Received Feb. 28, 2020; accepted Jun. 7, 2020)