

## 专题研究专栏

**编者按:**随着国民经济的飞速发展,人们对农产品品质提出了更高的要求,因而高质量的农产品需求量愈来愈大;同时,目前中国农产品采后损耗率约在20%~30%,而发达国家则不足5%。农产品贮藏、保鲜和加工技术是现代农业生产中十分重要而迫切的研究课题。本期选择了蜜柚、白玉菇、樱桃番茄和石榴4个农产品品种保鲜与加工研究的4篇文章,分别从近红外透射光谱分级分选、中短波红外干燥、超声协同钙浸渍处理保鲜、超声处理提取多酚的角度,探讨了新技术、新工艺对上述农产品进行保鲜和加工的效果和可行性。希望此方面的研究能为降低农产品损耗,提高农产品品质,增加农产品附加值等方面工作提供有益借鉴。

(主持人:刘东红教授)

doi:10.12301/j.issn.2095-6002.2021.01.004

文章编号:2095-6002(2021)01-0037-09

引用格式:孙潇鹏,刘灿灿,陆华忠,等.基于近红外透射光谱与机器视觉的蜜柚汁胞粒化分级检测[J].食品科学技术学报,



2021,39(1):37-45.

SUN Xiaopeng, LIU Cancan, LU Huazhong, et al. Detection of honey pomelo in different granulation levels based on near-infrared transmittance spectroscopy combined with machine vision[J]. Journal of Food Science and Technology, 2021,39(1):37-45.

## 基于近红外透射光谱与机器视觉的蜜柚汁胞粒化分级检测

孙潇鹏<sup>1</sup>, 刘灿灿<sup>2</sup>, 陆华忠<sup>1,3</sup>, 徐 赛<sup>4,\*</sup>

(1. 华南农业大学 工程学院, 广东 广州 510642; 2. 广西大学 轻工与食品工程学院, 广西 南宁 530004; 3. 广东省农业科学院, 广东 广州 510642; 4. 广东省农业科学院 农产品公共监测中心, 广东 广州 510642)

**摘要:**汁胞粒化是一种柑橘类水果中汁液囊的生理失调现象,表现为汁液囊变硬、干燥等,对水果内部品质产生消极影响。蜜柚是一种厚皮的柑橘类水果,很难通过外部果皮及果形,鉴定果实内部的汁胞粒化程度。采用近红外透射光谱结合机器视觉技术的快速无损检测方法,对蜜柚汁胞粒化程度进行分级检测。采集600个不同生长期的蜜柚样本在900~1700 nm的光谱数据,按果实的汁胞粒化程度将其分为5级。结合化学计量学研究由汁胞粒化引起的内部品质的化学变化,而机器视觉技术可用于研究由汁胞粒化引起的外部特征的物理变化。因此,该方法相较于传统检测方法,分级模型的预测能力更好。尤其是,连续投影-K近邻算法预测模型的准确性、敏感性和特异性分别达到0.970 0、0.923 1和0.987 4以上。结果表明:该方法可用于汁胞粒化的鉴定与评估分级,且具有巨大潜力,以期对厚皮类水果在线分选及内部品质研究提供参考和理论依据。

**关键词:** 机器视觉; 近红外透射光谱; 蜜柚; 汁胞粒化; 分级模型

**中图分类号:** TS255.3; O657.61

**文献标志码:** A

收稿日期:2020-06-19

基金项目:国家自然科学基金资助项目(31901404);广东省农业科学院新兴学科团队建设项目(201802XX);广东省农业科学院院长基金面上项目(201920);广东省重点领域研发计划项目(2018B020240001);广州市科创委资助项目(201904010199)。

第一作者:孙潇鹏,男,博士研究生,研究方向为农产品品质无损检测技术与装备研发。

\*通信作者:徐 赛,男,副研究员,主要从事农产品品质无损检测技术方面的研究。

蜜柚(*Citrus maxima* L.)为芸香科柑橘属乔木,果圆球形或梨形,风味极佳,营养价值高。柑橘类水果经常遭受各种生理性疾病的困扰,如汁胞粒化(又称为硬粒化)等,对蜜柚的品质和口感产生极大的影响<sup>[1]</sup>。1934年,Bartholomew等<sup>[2]</sup>首次报道了汁胞粒化后,又有来自不同国家的研究人员进行了相关报道。汁胞粒化是柑橘类水果中汁液囊的一种生理失调,其汁液囊变硬、干燥膨大,呈灰色,几乎没有可提取的汁液<sup>[3]</sup>。随着汁胞粒化程度的增加,水果从果汁囊泡中损失了更多的水分,伴随着内部品质的迅速降低,使其缺乏食用价值。柑橘类水果的汁胞粒化通常与汁液囊泡中的次生壁形成和细胞壁增厚有关<sup>[4]</sup>。收获时间也对果实内部品质有重要影响,收获时间较晚可能会加重蜜柚的汁胞粒化程度<sup>[5]</sup>。由于造成汁胞粒化的原因很复杂,所以对汁胞粒化的相关研究非常有限。

近红外透射光谱是一种无损检测技术,现已广泛应用于水果的内部品质检测,如可溶性固形物含量(SSC)<sup>[6]</sup>、水果成熟度<sup>[7]</sup>或柑橘黄龙病<sup>[8]</sup>等检测。然而,有限的穿透深度限制了其对水果内部品质检测的潜力,尤其是可能会影响厚皮水果(如西瓜和蜜柚)内部品质检测的准确性<sup>[9]</sup>。由于样本的特殊性和汁胞粒化的复杂性,很难准确地分类不同粒化等级的蜜柚。近年来,机器视觉技术已广泛用于农业和食品行业,尤其是对水果特征的提取和计算(如果形提取、体积估算等)<sup>[10]</sup>。研究发现:体积的差异可用于水果品质检测,如柑橘类水果的汁胞粒化分级<sup>[11]</sup>。因此,本研究将近红外透射光谱与机器视觉技术相结合,对蜜柚的粒化等级进行分级检测和评估。

## 1 材料与方法

### 1.1 材料与试剂

蜜柚采摘自梅州市蜜柚果园,采摘后立即运至实验室,剔除外部破损和畸形的蜜柚,擦拭蜜柚外部污渍,自然晾干,在室温(19~21℃)下静置24h后,进行实验。根据蜜柚生长期,每周采集20个样本进行实验,实验样本共计600个。

NaOH溶液、酚酞指示剂,广州和为医药科技有限公司。

### 1.2 仪器与设备

PAL-Grape Must型数字折光仪,ATAGO(爱拓)中国分公司;游标卡尺(测量范围0~300mm),

上海申韩量具有限公司;DHG-9030A型电热鼓风干燥箱,上海一恒科学仪器有限公司;酸碱中和滴定常用仪器(酸式滴定管、碱式滴定管、滴定管夹、铁架台、锥形瓶、玻璃棒等),江苏博雅教学设备有限公司。

近红外透射光谱检测系统:计算机、NIR-QUEST型光谱仪(测量波长900~1700nm,光谱分辨率0.25nm,信噪比15000:1,积分时间1ms~10s,美国海洋光学公司)、实验箱体、光纤、积分球和实验托盘等<sup>[12-13]</sup>。光谱采集软件,广州标旗光电科技发展有限公司。

### 1.3 实验方法

#### 1.3.1 近红外透射光谱采集及预处理

启动近红外透射光检测系统,预热15min使设备达到稳定的工作状态。在光谱采集软件中,积分时间设置2000ms,平均次数设置为1。将托盘置于光谱平台内,分别采用标准白板和标准黑板贴住积分球校正光谱仪后,进行样本光谱采集。将柚果放置在托盘上,每旋转90°采集1次,并存储光谱数据,取4次光谱的平均值作为该样本的光谱数据。在近红外透射光谱中,将样本的相对透射率作为光谱信息的研究参数,计算见式(1)。光谱数据中除有效光谱信息外,也含有大量无关或冗余信息,且存在背景噪声。光谱预处理可削弱或剔除干扰及无效信息,保留有效信息,降低模型的复杂度。研究采用Savitzky-Golay多项式平滑对原始光谱进行预处理<sup>[14]</sup>。

$$T = \frac{I_s - I_b}{I_w - I_b} \times 100\% \quad (1)$$

式(1)中, $T$ ,样本的相对透射率,%; $I_s$ ,样本的透射光强度,cd; $I_w$ ,标准白板的透射光强度,cd; $I_b$ ,标准黑板的透射光强度,cd。

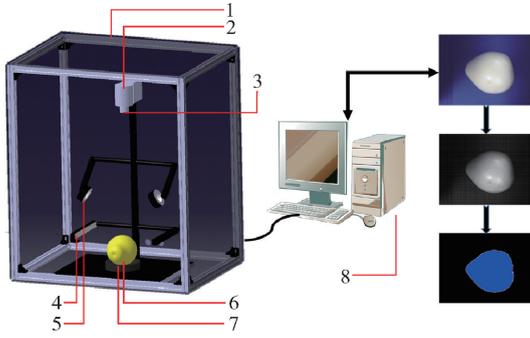
#### 1.3.2 光谱特征提取

连续投影算法(successive projections algorithm, SPA)是一种无监督的变量选择算法,可用于近红外透射光谱的特征变量提取。SPA从光谱数据中提取有效信息,依据均方根误差相对较低的值,提取符合要求的特征变量<sup>[15]</sup>。该算法能够有效剔除变量之间的共线性,最大程度避免信息的重复,使得变量之间的信息冗余度最低。经SPA预处理后,光谱数据以少数波长点来概括样品中大部分光谱信息,从而能够提升模型数据处理的速率<sup>[16]</sup>。

#### 1.3.3 样本特征的获取与估算

基于机器视觉技术,自搭建图像获取系统,如图1。图像获取系统配套图片捕获软件(IC capture)和

图像处理软件(Matlab R2018a)共同使用。



1. 实验箱体;2. GigE 相机(DFK 33GP006 型);3. 相机镜头(M0814-MP2, F1.4, f8 mm 2/3);4. 条形光源(24 V,580 mA);5. 环形光源(24 V,580 mA);6. 蜜柚样本;7. 蜜柚托盘;8. 计算机。

图1 图像信息获取系统

Fig.1 Acquisition system of image information

基于机器视觉技术,通过图像信息获取系统,可快速获取并估算样本的特征数据。该系统采用 IC capture 软件获取样本图像信息,110 mm × 110 mm 标定板和 Matlab 软件工具箱(computer vision system toolbox 8.1)进行相机标定。Matlab 软件工具箱(image processing toolbox 10.2)可快速获取横径和纵径的像素值,通过像素-尺度转换<sup>[17]</sup>,得到横径和纵径的估算值。

估算体积时,可将蜜柚默认为椭球体。椭球体的长、宽、高分别用  $L$ 、 $W$  和  $H$  表示。将椭球体划分为多个圆台,切面默认为标准圆,即  $W \approx H$ ,圆台的上顶面和下底面分别由  $H_{ij}$  和  $H_{ij}$  表示,椭球体的  $H$  和  $L$  可分别代表横径和纵径,如图2。因此,样本的体积可认为是多个圆台体积的累加,圆台体积( $V_j$ )和样本的估算体积( $V$ ),可由式(2)和式(3)计算得到。

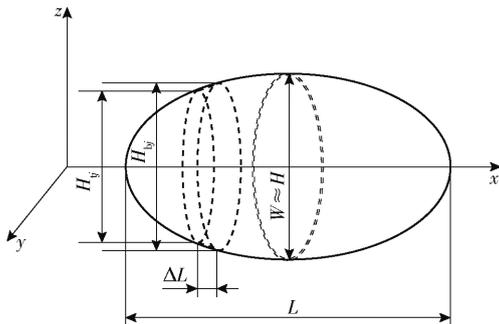


图2 椭球体的划分和体积估算

Fig.2 Partition and volume calculate of ellipsoid

$$V_j = \frac{1}{3} \pi \Delta L \left( \left( \frac{H_{ij}}{2} \right)^2 + \frac{H_{ij} H_{ij}}{4} + \left( \frac{H_{ij}}{2} \right)^2 \right) ; \quad (2)$$

$$V = \sum_{j=1}^n V_j = \sum_{j=1}^n \frac{\pi \Delta L}{12} (H_{ij}^2 + H_{ij} H_{ij} + H_{ij}^2) \quad (3)$$

### 1.3.4 样本理化指标测定

分别采用游标卡尺测定样本的横径和纵径(mm),电子天平称量样本的质量(g),排水法测定样本的体积( $\text{cm}^3$ )。提取光谱采集区域内果肉少许,通过挤压和纱布过滤提取果汁,数字折光仪测量可溶性固形物含量,酸碱中和滴定法测定样本的可滴定酸度(titratable acidity, TA),每个样本重复3次后记录平均值。采用赵晓玲<sup>[18]</sup>方法测定样本的汁胞含水率(moisture content, MC),计算见式(4)<sup>[19]</sup>。

$$\text{汁胞含水率} = \frac{m_1 - m_2}{m_1} \times 100\% \quad (4)$$

式(4)中, $m_1$ ,鲜汁胞的质量,g; $m_2$ ,干汁胞的质量,g。

### 1.4 粒化蜜柚评价分级

邀请10位感官评价人员对不同生长期样本的不同粒化程度进行打分。分级标准参考陈昆松等<sup>[19]</sup>的方法进行,将蜜柚沿着果蒂纵切八瓣,进行粒化评级。根据蜜柚汁胞粒化程度,将蜜柚分为5级,1级为未发生汁胞粒化,2级为汁胞粒化程度小于25%,3级为汁胞粒化程度在26%~50%,4级为汁胞粒化程度在51%~75%,5级为汁胞粒化程度大于75%<sup>[20]</sup>。

### 1.5 模型建立与评价指标确定

#### 1.5.1 模糊C均值聚类

模糊C均值(fuzzy c-means, FCM)聚类是模糊聚类中常用的划分聚类方法,属于软化分算法。FCM算法由MacQueen提出,其原理是首先定义一个准则函数,并随机选取 $c$ 个初始聚类中心,然后根据样本与聚类中心的隶属度来确定聚类关系,再重新计算每类的聚类中心,此过程不断重复,直到准则函数最小。准则函数为样本和聚类中心的平方误差的总和<sup>[21]</sup>。

模糊C均值聚类的准则,设 $x_i (i=1, 2, \dots, n)$ 是 $n$ 个样本组成的样本集合, $c$ 为预定的类别数目, $\mu_j(x_i)$ 是第 $i$ 个样本对于第 $j$ 类的隶属度函数。用隶属度函数定义的聚类损失函数计算见式(5)。

$$J_f = \sum_{j=1}^c \sum_{i=1}^n [\mu_j(x_i)]^b \|x_i - m_j\|^2 \quad (5)$$

其中, $b > 1$ ,是一个可以控制聚类结果模糊程度的常数。在不同的隶属度定义方法下最小化聚类损失函数,就得到不同的模糊聚类方法。其中最具有代表性的是模糊C均值聚类,它要求每个样本对于各个聚类的隶属度之和为1,如式(6)。

$$\sum_{j=1}^c \mu_j(x_i) = 1 \quad (i=1, 2, \dots, n) \quad (6)$$

### 1.5.2 K-近邻法聚类

K-近邻法(K-nearest neighbor, KNN)聚类是一种高效且具备较高分类精度的方法<sup>[22]</sup>。KNN聚类算法的核心思想是在多维空间 $R^n$ 中找到与未知样本最近邻的 $k$ 个点,并根据这 $k$ 个点的类别来判断预测样本的类属。该算法假设所有样本对应于 $n$ 维空间中的点,每个样本的最近邻是根据欧式距离定义,设 $x$ 的特征向量为 $[a_1(x), a_2(x), \dots, a_n(x)]$ 。其中, $a_r(x)$ 表示样本 $x$ 的第 $r$ 个属性值。2个样本 $x_i$ 和 $x_j$ 间的距离定义为 $d(x_i, x_j)$ ,计算见式(7)。

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^n (a_r(x_i) - a_r(x_j))^2} \quad (7)$$

以 $x_i$ 为预测样本的特征向量, $Y$ 定义样本的类别, $y_i \in Y \{c_1, c_2, \dots, c_k\}$ 为预测样本的类别, $I$ 为指示函数, $N_K(x)$ 为 $n$ 维空间内选取与 $x$ 最邻近的 $k$ 个点的领域,并以 $x$ 预测 $y$ 类别<sup>[23]</sup>,如式(8)。

$$y = \arg \max_{x_i \in N_K(x)} \sum I(y_i = c_j) \quad (8)$$

( $i=1, 2, 3, \dots, N; j=1, 2, 3, \dots, K$ )

### 1.5.3 样本分级的评价指标确定

通过 kennard-stone 法<sup>[24]</sup>将 600 个实验样本按照 2:1 的比例,分为训练集和预测集。运用模糊 C 均值聚类、K-近邻法聚类对样本进行分级,运用混淆矩阵评估二进制分类模型的性能<sup>[25]</sup>。假设将 1 级样本定义为目标样本,2、3、4、5 级样本定义为非目标样本。TP 为目标样本正确分类为 1 级样本的数量;TN 为非目标样本正确分类在 2、3、4、5 级样本中的数量;FP 为非目标样本分类为 1 级样本的数量;FN 为目标样本分类为 2、3、4、5 级样本的数量,即本属于 1 级的样本被分类为 2 级、3 级、4 级或 5 级的样本数量和。当其他级别样本定义为目标样本,皆

可根据以上方法进行假设。再根据式(9)~式(11),分别从聚类模型的准确性、敏感性和特异性进行分级评价。

$$\text{准确性} = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

$$\text{敏感性} = \frac{TP}{TP + FN} \quad (10)$$

$$\text{特异性} = \frac{TN}{TN + FP} \quad (11)$$

## 2 结果与分析

### 2.1 近红外透射光谱分析

600 个样本的近红外透射原始光谱见图 3(a)。在近红外光谱区域内,典型的重叠吸收峰对应水果中水分和碳水化合物的 C—H、O—H 和 N—H 化学键。在短波近红外光谱区域(900~1100 nm),重叠吸收峰出现在大约 948 nm 处,950 nm 附近的吸收峰为水分的吸收带<sup>[26]</sup>。汁胞粒化的过程伴随着样本枯水的过程,其果汁囊泡中损失了更多的水分。在长波近红外光谱区域(1100~1700 nm),最大的吸收峰出现在 1283 nm 处,这是 C—H 键第二官能团振动带<sup>[27]</sup>。而 1410 nm 处吸收峰是木质素相关的吸收带,它是亚甲基 C—H 键与  $R(\text{CH}_2)_nR$  的 C—H 键的组合官能团。柑橘类水果发生汁胞粒化与果汁囊泡的细胞壁结构变化有关,是木质素、纤维素和果胶等细胞壁成分的增加造成的<sup>[3]</sup>。5 种粒化等级的平均光谱见图 3(b)。随着样本的汁胞粒化程度加重,948 nm 处吸收峰呈依次递减趋势;而 1410 nm 处吸收峰呈依次递增趋势。

样本的近红外透射光谱经预处理后,采用 SPA

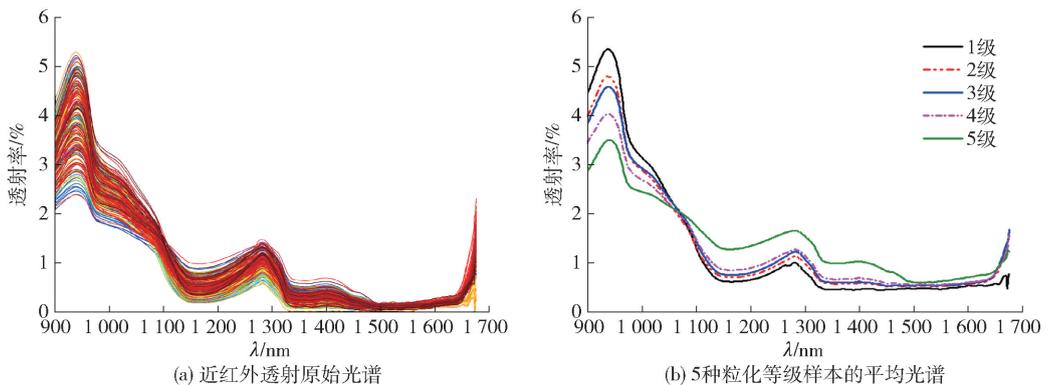


图3 样本的近红外透射光谱

Fig. 3 Near-infrared transmittance spectra of samples

进行光谱特征提取,如图4。SPA共选取17个特征变量,代表样本在900~1700 nm波段的光谱信息,快速高效地完成了数据降维,解决了变量间的共线性问题。且特征变量多集中在950、1283、1410 nm吸收峰附近,与蜜柚汁胞粒化的过程密切相关。

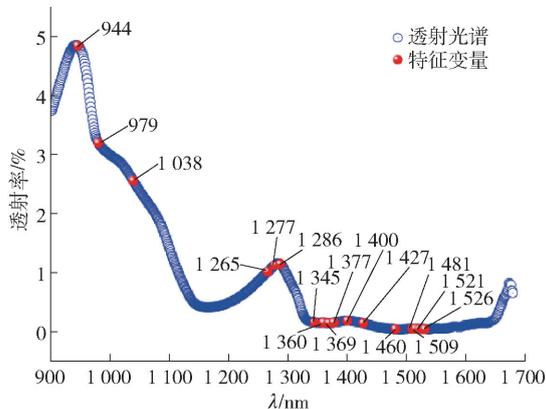


图4 连续投影法的光谱特征提取

Fig. 4 Extraction of spectral feature by successive projections algorithm

## 2.2 样本理化指标与汁胞粒化的关系分析

根据蜜柚的生长期,研究1~5级汁胞粒化样本的特征变化,如图5。实验采集的样本随着汁胞粒化等级的增加,外部形态逐渐减小。根据粒化样本评级方法,样本随着汁胞粒化等级的增加,内部出现

汁液囊变硬、干燥、膨大等特征,颜色呈灰色,几乎没有可提取的汁液(见图6)。随储藏时间的增加,果汁囊泡中损失了更多的水分,存储过程中疾病的严重性增加。

各级粒化样本的理化指标见表1。结果表明:随着汁胞粒化等级由1级升至5级,样本的横径和纵径逐渐减小,伴随着体积也逐渐减小;与此同时,样本的MC、SSC和TA也逐渐降低。样本的外部特征及内部品质均会受到汁胞粒化的影响。

## 2.3 样本特征估算与数据分析

为验证体积估算值的可靠性,将排水法测定的体积测量值( $V_{acc}$ )与图像信息获取系统计算的体积估算值( $V$ )进行比较。对 $V_{acc}$ 与 $V$ 进行线性拟合分析,如图7(a)。由回归方程可知, $V$ 与 $V_{acc}$ 之间存在线性规律,决定系数 $R^2$ 等于0.9879,证明蜜柚体积的估算模型中,估算值相较于真实值,准确率达到98.79%。对 $V_{acc}$ 与 $V$ 进行差异性分析,来确定样品的测量值和估算值之间的差异<sup>[28]</sup>,见图7(b)。 $V_{acc}$ 与 $V$ 的95%一致性极限均位于 $d-1.96$ 至 $d+1.96$ 。 $V_{acc}$ 与 $V$ 平均差异百分比为7.2%,95%一致性极限区间为[4.4, 9.9],证明 $V$ 比 $V_{acc}$ 小约4.4%~9.9%。进一步佐证体积估算值可代替测量值,将其融合光谱数据可用于汁胞粒化的模型分级检测与研究。

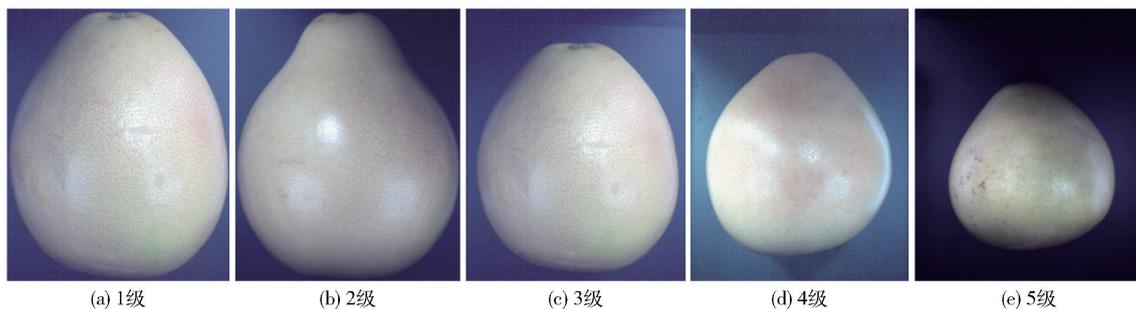


图5 1~5级汁胞粒化样本的外部特征变化

Fig. 5 External characteristic changes of granulation samples from class 1 to class 5

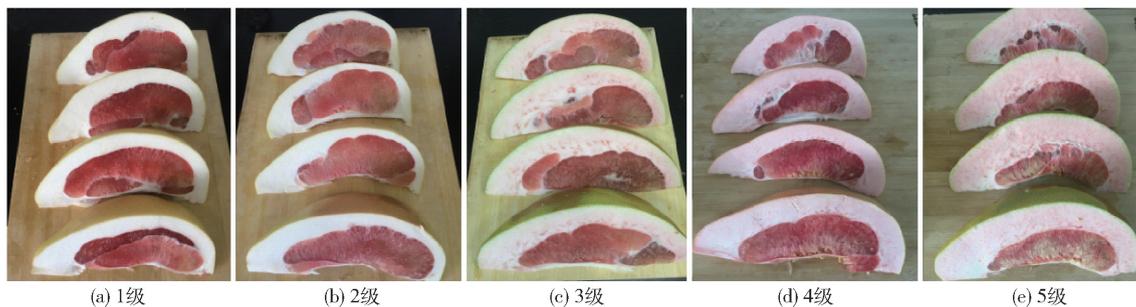


图6 1~5级汁胞粒化样本的内部品质变化

Fig. 6 Inner quality changes of granulation samples from class 1 to class 5

表1 各级粒化样本的理化指标

Tab.1 Physical and chemical indicators of samples in each granulation level

级别	数量/个	H/mm	L/mm	V/cm <sup>3</sup>	MC/%	SSC/%	TA/%
1	146	193.82 ± 3.93 <sup>a</sup>	216.91 ± 1.93 <sup>a</sup>	4 237.03 ± 19.83 <sup>a</sup>	87.9 ± 1.7 <sup>a</sup>	12.4 ± 0.8 <sup>a</sup>	0.86 ± 0.09 <sup>a</sup>
2	166	190.86 ± 8.09 <sup>a</sup>	212.93 ± 8.66 <sup>ab</sup>	3 956.70 ± 17.46 <sup>a</sup>	86.9 ± 1.4 <sup>ab</sup>	10.9 ± 0.6 <sup>b</sup>	0.82 ± 0.13 <sup>a</sup>
3	76	177.72 ± 8.17 <sup>b</sup>	202.90 ± 0.46 <sup>bc</sup>	3 358.68 ± 18.38 <sup>b</sup>	85.2 ± 5.5 <sup>bc</sup>	10.3 ± 0.5 <sup>c</sup>	0.64 ± 0.09 <sup>b</sup>
4	120	177.51 ± 3.68 <sup>b</sup>	197.29 ± 5.88 <sup>c</sup>	3 322.64 ± 12.97 <sup>b</sup>	84.2 ± 3.2 <sup>bc</sup>	10.1 ± 0.8 <sup>c</sup>	0.63 ± 0.07 <sup>b</sup>
5	92	162.03 ± 8.87 <sup>c</sup>	172.58 ± 6.76 <sup>d</sup>	2 468.45 ± 15.79 <sup>c</sup>	83.3 ± 6.5 <sup>d</sup>	10.0 ± 0.6 <sup>c</sup>	0.54 ± 0.05 <sup>c</sup>

同列不同字母表示差异显著( $P < 0.05$ )。

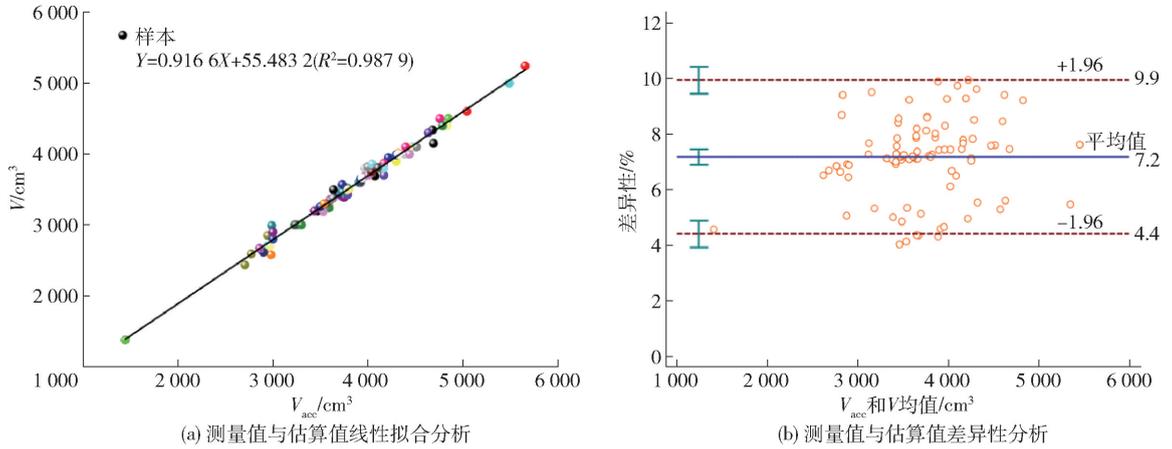


图7 体积测量值和估算值的数据分析

Fig.7 Data analysis of measured and estimated values of volume

## 2.4 预测模型的分级评价

研究采用 SPA-FCM 和 SPA-KNN 2 种分级模型,对各级粒化样本进行分级检测。根据 kennard-stone 法<sup>[24]</sup>,从 600 个实验样本数据中,随机选取 400 个训练集样本数据,进行模型建立和训练;剩余

200 个预测集样本数据进行模型预测。模型输入变量为经预处理和特征提取的近红外透射光谱数据与图像特征提取的样本体积估算值,模型输出变量为汁胞粒化等级。

分级模型的混淆矩阵,如表 2。SPA-FCM 与

表2 分级模型的混淆矩阵

Tab.2 Confusion matrix of classification models

分级模型	数量/个	训练集					数量/个	预测集						
		真实等级	预测等级中样本数量/个					真实等级	预测等级中样本数量/个					
			1	2	3	4			5	1	2	3	4	5
SPA-FCM	131	1	83	48	0	0	0	15	1	11	4	0	0	0
	101	2	12	89	0	0	0	65	2	7	58	0	0	0
	41	3	0	3	38	0	0	35	3	0	0	35	0	0
	70	4	0	0	2	63	5	50	4	0	0	0	45	5
	57	5	0	0	3	17	37	35	5	0	0	0	9	26
SPA-KNN	114	1	104	10	0	0	0	32	1	32	2	0	0	0
	131	2	6	125	0	0	0	35	2	1	34	0	0	0
	21	3	0	0	16	0	5	55	3	0	0	55	0	0
	94	4	0	0	0	88	6	26	4	0	0	0	25	1
	40	5	0	0	3	4	33	52	5	0	0	0	4	48

SPA-KNN 的混淆矩阵中, 预测集的误判样本均少于训练集。模型训练集与模型预测集中, 1 级和 2 级互为误判样本的机率较高, 证明从无汁胞粒化至小于 25% 汁胞粒化, 样本的内部品质变化较小; 因此分级模型完成 1 级和 2 级准确分级难度较大。与此同时, 4 级和 5 级互为误判样本的机率较高, 证明从 50% ~ 75% 汁胞粒化至大于 75% 汁胞粒化, 样本内部均已严重粒化, 伴随着汁胞含水率的降低, 出现严重的枯水现象, 无可提取果汁, SSC 和 TA 的急剧降低, 内部果肉基本不可食, 且口感极差; 因此, 分级模型完成 4 级和 5 级准确分级难度较大。然而, SPA-FCM 与 SPA-KNN 均能完成对 3 级样本的准确分类。

根据分级评价指标, 结合模型的混淆矩阵, 对预测集样本进行分级预测, 见表 3。SPA-FCM 预测结果的准确性、敏感性和特异性分别达到 0.925 9、0.733 3 和 0.935 3 以上; SPA-KNN 预测结果的准确性、敏感性和特异性分别达到 0.970 0、0.923 1 和 0.987 4 以上。而且, SPA-FCM 与 SPA-KNN 对 3 级样本完全准确预测, 其准确性、敏感性和精确性均达到 1。结果表明: SPA-KNN 相较于 SPA-FCM, 对各级粒化样本的分级预测能力更好, 即能更好地完成对样本汁胞粒化的分级检测。

表 3 分级模型对样本预测集的预测能力

Tab. 3 Predictive performances of classification models for predicted set of samples

等级	SPA-FCM			SPA-KNN		
	准确性	敏感性	特异性	准确性	敏感性	特异性
1	0.940 9	0.733 3	0.959 1	0.984 5	0.941 2	0.993 8
2	0.940 9	0.892 3	0.966 9	0.984 5	0.971 4	0.987 4
3	1	1	1	1	1	1
4	0.925 9	0.900 0	0.935 3	0.984 5	0.961 5	0.988 1
5	0.925 9	0.742 9	0.967 5	0.970 0	0.923 1	0.993 2

### 3 结 论

采用近红外透射光谱和机器视觉技术的分级预测模型, 可用于蜜柚汁胞粒化的分级检测。近红外透射光谱可以捕获由汁胞粒化引起的内部品质的化学变化。机器视觉技术可快速估算样本的外观特征, 探究由汁胞粒化引起的外部特征的物理变化。此方法对粒化样本的预测能力好, 分级准确率高, 分

级检测过程中, 对样本无破坏, 适用于实时在线检测。因此, 基于近红外透射光谱和机器视觉技术对蜜柚汁胞粒化的分级检测具有较高的可行性, 以期对厚皮水果或柑橘类水果的内部品质检测提供了一定的研究基础和参考。

#### 参考文献:

- [1] ZHOU Y, HE W Z, ZHENG W L, et al. Fruit sugar and organic acid were significantly related to fruit Mg of six citrus cultivars[J]. Food Chemistry, 2018, 259: 278 - 285.
- [2] BARTHOLOMEW E, SINCLAIR W, TURRELL F. Research on granulation of valencia oranges shows only limited control measures exist[J]. California Agriculture, 1947, 1:106 - 108.
- [3] ILAN S, EDO C, ROSA V, et al. Scierification of juice sacs in pummelo (*Citrus grandis*) fruit[J]. Canadian Journal of Botany, 1989, 67(3):625 - 632.
- [4] BURNS J K, ACHOR D S. Cell wall changes in juice vesicles associated with section drying in stored late-harvested grapefruit[J]. Journal of the American Society for Horticultural Science, 1989, 114:283 - 287.
- [5] PAILLY O, TISON G, AMOUROUX A. Harvest time and storage conditions of 'Star Ruby' grapefruit (*Citrus paradisi* Macf.) for short distance summer consumption [J]. Postharvest Biology and Technology, 2004, 34(1): 65 - 73.
- [6] 刘燕德, 吴明明, 李轶凡, 等. 苹果可溶性固形物和糖酸比可见/近红外漫反射与透射在线检测对比研究[J]. 光谱学与光谱分析, 2017, 37(8):2424 - 2429. LIU Y D, WU M M, LI Y F, et al. Comparison of reflection and diffuse transmission for detecting solid soluble contents and ratio of sugar and acid in apples by online Vis/NIR spectroscopy[J]. Spectroscopy and Spectral Analysis, 2017, 37(8):2424 - 2429.
- [7] 刘燕德, 徐海, 孙旭东, 等. 西红柿成熟度的近红外漫透射光谱无损检测[J]. 激光技术, 2019, 43(1): 25 - 29. LIU Y D, XU H, SUN X D, et al. Non-destructive measurement of tomato maturity by near-infrared diffuse transmission spectroscopy[J]. Laser Technology, 2019, 43(1):25 - 29.
- [8] 孟幼青, 翁海勇, 岑海燕, 等. 潜伏期柑橘黄龙病宿主糖代谢及近红外光谱特征[J]. 浙江农业学报, 2019, 31(3):428 - 435. MENG Y Q, WENG H Y, CEN H Y, et al. Carbohy-

- drate metabolism and near-infrared spectral characteristics in asymptomatic Huanglongbing infected leaves [J]. *Acta Agriculturae Zhejiangensis*, 2019, 31(3):428-435.
- [9] ZHANG H L, ZHAN B S, PAN F, et al. Determination of soluble solids content in oranges using visible and near infrared full transmittance hyperspectral imaging with comparative analysis of models [J]. *Postharvest Biology and Technology*, 2020, 163:1-7.
- [10] AKAR R, AYDIN C. Some physical properties of gumbo fruit varieties [J]. *Journal of Food Engineering*, 2005, 66(3):387-393.
- [11] OMID M, KHOJASTEHN AZHAND M, TABATABAEEFAR A. Estimating volume and mass of citrus fruits by image processing technique [J]. *Journal of Food Engineering*, 2010, 100(2):315-321.
- [12] 孙潇鹏,徐赛,陆华忠. 基于光谱技术与光学仿真的柚果在线检测托盘设计与试验 [J]. *食品与机械*, 2019, 35(12):56-62.
- SUN X P, XU S, LU H Z. Design and experiment of online detection pallets for pomelo fruit based on spectral technique and optical simulation [J]. *Food & Machinery*, 2019, 35(12):56-62.
- [13] 潘嘹,董迪,卢立新. 基于漫反射可见光谱无损检测生鲜牛肉和猪肉肌红蛋白的研究 [J]. *食品科学技术学报*, 2020, 38(6):103-110.
- PAN L, DONG D, LU L X. Non-destructive detection for quantifying myoglobin in fresh beef and pork by using a surface diffuse reflection visible spectrum [J]. *Journal of Food Science and Technology*, 2020, 38(6):103-110.
- [14] 马文强,张漫,李源,等. 核桃仁脂肪含量的近红外光谱无损检测 [J]. *农业机械学报*, 2019, 50(s1):374-379.
- MA W Q, ZHANG M, LI Y, et al. Non-destructive detection for fat content of walnut kernels by near infrared spectroscopy [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2019, 50(s1):374-379.
- [15] 韩建,李雨昭,曹志民,等. SPA-PLS 的高含水原油近红外光谱含水率分析 [J]. *光谱学与光谱分析*, 2019, 39(11):3452-3458.
- HAN J, LI Y Z, CAO Z M, et al. Water content prediction for high water-cut crude oil based on SPA-PLS using near infrared spectroscopy [J]. *Spectroscopy and Spectral Analysis*, 2019, 39(11):3452-3458.
- [16] 方孝荣,章海亮,黄凌霞,等. 近红外光谱技术结合 RCA 和 SPA 方法检测土壤总氮研究 [J]. *光谱学与光谱分析*, 2015, 35(5):1248-1252.
- FANG X R, ZHANG H L, HUANG L X, et al. Measurement of soil total nitrogen using near infrared spectroscopy combined with RCA and SPA [J]. *Spectroscopy and Spectral Analysis*, 2015, 35(5):1248-1252.
- [17] 袁雷明. 基于多视成像及近红外光谱技术的巨峰葡萄品质无损检测研究 [D]. 镇江:江苏大学, 2016:35-38.
- YUAN L M. Study on non-destructive detection of 'Kyoho' grape's quality by multi-perspective imaging and NIR spectroscopy techniques [D]. Zhenjiang: Jiangsu University, 2016:35-38.
- [18] 赵晓玲. 琯溪蜜柚 [ *Citrus grandis* (L.) Osbeck. cv. guanxi miyou] 成熟阶段粒化过程的若干生理生化变化 [D]. 福州:福建农林大学, 2007:20-63.
- ZHAO X L. Physiological and biochemical changes of pomelo fruit [ *Citrus grandis* (L.) Osbeck. cv. guanxi miyou] in the process of granulation during maturity [D]. Fuzhou: Fujian Agriculture and Forestry University, 2007:20-63.
- [19] 陈昆松,张上隆,李方,等. 胡柚果实采后枯水的研究 [J]. *园艺学报*, 1995(1):35-39.
- CHEN K S, ZHANG S L, LI F, et al. A study on the postharvest granulation of 'Huyou' fruit [J]. *Acta Horticulturae Sinica*, 1995(1):35-39.
- [20] 王森,张晶,贺妍,等. 基于低场核磁共振的柑橘汁胞粒化评级 [J]. *农业工程学报*, 2016, 32(7):290-295.
- WANG M, ZHANG J, HE Y, et al. Evaluation of juicy sac granulation in citrus with low field nuclear magnetic resonance [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2016, 32(7):290-295.
- [21] 武斌,傅海军,武小红,等. 可能模糊鉴别 C 均值聚类的茶叶 FTNIR 分类研究 [J]. *光谱学与光谱分析*, 2020, 40(2):512-516.
- WU B, FU H J, WU X H, et al. Classification of FTNIR spectra of via possibilistic fuzzy discriminant C-means clustering [J]. *Spectroscopy and Spectral Analysis*, 2020, 40(2):512-516.
- [22] DENG Z Y, ZHU X S, CHENG D B, et al. Efficient KNN classification algorithm for big data [J]. *Neurocomputing*, 2016, 195(26):143-148.
- [23] 陈真诚,杜莹,邹春林,等. 基于 K-nearest neighbor 和神经网络的糖尿病分类研究 [J]. *中国医学物理学杂志*, 2018, 35(10):1220-1224.
- CHEN Z C, DU Y, ZOU C L, et al. Classification of diabetes based on K-nearest neighbor and neural network [J]. *Chinese Journal of Medical Physics*, 2018, 35(10):1220-1224.

- [24] ZHANG L, LI G, SUN M, et al. Kennard-stone combined with least square support vector machine method for noncontact discriminating human blood species[J]. *Infrared Physics and Technology*, 2017, 86:116–119.
- [25] CHONG I G, JUN C H. Performance of some variable selection methods when multicollinearity is present[J]. *Chemometrics & Intelligent Laboratory Systems*, 2005, 78(1/2):103–112.
- [26] GUO Z M, WANG M M, AKWASI A A, et al. Quantitative detection of apple watercore and soluble solids content by near infrared transmittance spectroscopy[J]. *Journal of Food Engineering*, 2020, 279: 1–5.
- [27] MARTON J, SPARKS H E. Determination of lignin in pulp and paper by infrared multiple internal reflectance [J]. *Tappi Journal*, 1967, 50(7):363–368.
- [28] BLAND J M. Measuring agreement in method comparison studies[J]. *Statistical Methods in Medical Research* 1999, 8(2):135–160.

## Detection of Honey Pomelo in Different Granulation Levels Based on Near-Infrared Transmittance Spectroscopy Combined with Machine Vision

SUN Xiaopeng<sup>1</sup>, LIU Cancan<sup>2</sup>, LU Huazhong<sup>1,3</sup>, XU Sai<sup>4,\*</sup>

(1. College of Engineering, South China Agricultural University, Guangzhou 510642, China;

2. College of Light Industry and Food Engineering, Guangxi University, Nanning 530004, China;

3. Guangdong Academy of Agricultural Sciences, Guangzhou 510642, China;

4. Public Monitoring Center for Agro-Product, Guangdong Academy of Agricultural Sciences, Guangzhou 510642, China)

**Abstract:** Granulation is a physiological disorder of juice sacs in citrus fruit, which made juice sacs become hard and dry and damaged the internal quality of citrus fruit. Honey pomelo is a thick-skinned citrus fruit, and it is hard to identify the granulation levels by observing the outer peel and fruit shape. In this study, a rapid and non-destructive detection method based on near-infrared transmittance spectroscopy combined with machine vision technology was used to classify honey pomelo by the granulation levels. 600 honey pomelos in different growth stages were harvested and divided into five granulation levels according to the granulation changes of samples. Spectral data of samples were recorded in the range of 900 ~ 1700 nm, which were combined with chemometrics to research the chemical changes of inner quality caused by granulation. Machine vision technology can be used to study the physical changes of external characteristics caused by granulation. Therefore, comparison of the traditional method, this method has better predictive performances in classification models. In particular, the predictive performances of accuracy, sensitivity, and specificity were respectively not less than 0.970 0, 0.923 1, and 0.987 4 in the SPA-KNN (successive projections algorithm-K nearest neighbor) predicted model. The results showed that this method could be used for classification and evaluation of granulation, and had a great potential. The method provides a reference and theoretical basis for the online sorting and inner quality detecting of thick-skinned fruits.

**Keywords:** machine vision; near-infrared transmittance spectroscopy; honey pomelo; granulation; classification model