

张 初,刘 飞,孔汶汶,章海亮,何 勇.利用近红外高光谱图像技术快速鉴别西瓜种子品种[J].农业工程学报,2013,29(20):270-277

利用近红外高光谱图像技术快速鉴别西瓜种子品种

Fast identification of watermelon seed variety using near infrared hyperspectral imaging technology

投稿时间: 2013-05-05 最后修改时间: 2013-08-31

中文关键词: [近红外光谱](#), [遗传算法](#), [图像处理](#), [西瓜种子](#), [连续投影算法](#), [极限学习机](#)

英文关键词: [near infrared spectroscopy](#), [genetic algorithm](#), [imaging processing](#), [watermelon seed](#), [successive projections algorithm](#), [extreme learning machine](#)

基金项目: 国家863课题资助 (2013AA102301)、十二五支撑计划项目资助 (2011BAD21B04)

作者 单位

[张 初](#) [1. 浙江大学生物系统工程与食品科学学院, 杭州 310058](#)

[刘 飞](#) [1. 浙江大学生物系统工程与食品科学学院, 杭州 310058](#)

[孔汶汶](#) [1. 浙江大学生物系统工程与食品科学学院, 杭州 310058](#)

[章海亮](#) [1. 浙江大学生物系统工程与食品科学学院, 杭州 310058](#). [华东交通大学机电工程学院, 南昌 330013](#)

[何 勇](#) [1. 浙江大学生物系统工程与食品科学学院, 杭州 310058](#)

摘要点击次数: 160

全文下载次数: 85

中文摘要:

为了研究采用近红外高光谱图像技术对西瓜种子品种快速无损鉴别的可行性, 该文采用近红外高光谱图像技术, 通过提取西瓜种子的光谱反射率, 结合Savitzky-Gol G)平滑算法, 经验模态分解算法 (empirical mode decomposition, EMD) 和小波分析 (wavelet transform, WT) 对提取出的光谱数据进行去除噪声处理, 采用连续投影 (successive projections algorithm, SPA) 和遗传-偏最小二乘法 (genetic algorithm-partial least squares, GA-PLS) 进行特征波长选择。基于全波段光谱建立了偏最小二判别分析 (partial least squares-discriminant analysis, PLS-DA), 基于特征波长建立了反向传播神经网络 (back-propagation neural network, BP NN) 判别模型和极限学习机 (extreme learning machine, ELM) 判别模型。试验结果表明, 基于特征波长的BPNN模型和ELM模型的结果优于基于全部波长的PLS-DA模型, 基于SG预处理光谱提取的特长建立的ELM模型取得最优的判别效果, 建模集和预测集的判别正确率均为100%。结果表明应用近红外高光谱成像技术对西瓜种子品种鉴别是可行的, 为西瓜种子的快速鉴别提供了一种新方法。

英文摘要:

Abstract: Watermelon seed variety selection plays a vital role in watermelon planting, and the variety of watermelon seeds directly affect the yield and quality of watermelons. In this study, we aimed to identify the cultivars of watermelon seeds by using a novel, rapid, non-invasive, and low cost technique named hyperspectral imaging. 121 samples of four different cultivars of watermelon seeds were investigated, and a near-infrared hyperspectral imaging system (874-1734 nm with 256 bands) was established to acquire the hyperspectral images of the samples. A region of interest (ROI) with 15×15 pixels of the hyperspectral image of each sample was defined, and the average reflectance spectrum of the ROI were extracted. To remove the absolute noises of the spectra, only the spectral range 1 042-1 646 nm was used for analysis, and to reduce the noises existed in spectral range 1 042-1 646 nm the extracted 121 reflectance spectra were preprocessed by Savitzky-Golay smoothing (SG), Empirical Mode Decomposition (EMD), and Wavelet Transform (WT) methods. The preprocessed spectra were then used to select sensitive wavelengths by Successive Projections Algorithm (SPA) and Genetic Algorithm-partial least squares (GA-PLS) methods. Different numbers of sensitive wavelengths were selected by different variable selection methods with different preprocessing methods. 24, 16, and 15 sensitive wavelengths were selected by SPA with spectra preprocessed by SG, EMD, and WT, respectively. Moreover, 38, 33, and 32 sensitive wavelengths were selected by GA-PLS with spectra preprocessed by SG, EMD, and WT, respectively. Partial least squares - discriminant analysis (PLS-DA) was used to build discriminant models with the full spectra, and back-propagation neural network (BPNN) and extreme learning machine (ELM) were applied to build discriminant models with the selected wavelength variables. A PLS-DA model with spectra preprocessed by EMD obtained the best identification rate among all PLS-DA models, with an identification rate of 91.57% in the calibration set and 78.95% in the prediction set. SPA-BPNN model showed relatively worse results than GA-PLS-BPNN models with the same spectral preprocessing methods. The SG-GA-PLS-BPNN model obtained the best performance among a BPNN models, with an identification rate of 92.77% in the calibration set and 86.84% in the prediction set. Compared with the PLS-DA models and the BPNN models, ELM models obtained the best results. All ELM models obtained an identification rate over 90% in the calibration set and the prediction set, and the SG-SPA-ELM model, SG-GA-PLS-ELM model and WT-SPA-ELM model obtained the identification rate of 100% of calibration and prediction. The overall results showed that BPNN and ELM models performed better than PLS-DA models, and the ELM models with the selected wavelengths based on SG preprocessed spectra obtained the best results, with 100% classification accuracy for both the calibration and the prediction set. The SG preprocessing method showed the best performance in all PLS-DA, BPNN, and ELM models. The results indicated that it was feasible to use near-infrared hyperspectral imaging to identify the watermelon seed varieties, and near-infrared hyperspectral imaging provided an alternate way of rapid identification of watermelon seed variety. ELM, as a single hidden layer feed-forward network, was an effective classification method in watermelon seed cultivar identification. Moreover, the results in this paper showed the great potential of hyperspectral imaging in the seed industry for on-line identification of seed cultivars and detection of the seed quality parameters.

[查看全文](#) [下载PDF阅读器](#)

关闭