ISSN: 2664-9640

# Research on classification of maize seed varieties based on spectral analysis

Peiye Li

Senior High School, Guangdong Experimental High School, Guangzhou 510055, China

## Abstract

**Because of the shortcomings of traditional maize variety classification and testing methods, which are professional, time-consuming, and destructive to seeds, it is of great practical importance to study rapid, accurate, and nondestructive methods for classifying maize seeds varieties. In this paper, three types of maize seeds, Yue xian nuo 6, Zheng tian 68, and Yue tian 29, were used to extract the region of interest and obtain the average spectral information of this region on each maize sample. The classification of maize varieties was investigated by hyperspectral imaging technology and a machine learning algorithm. The experimental results show that the spectral information of different kinds of maize seeds has some variability, and the algorithm achieves 87.5% classification accuracy on the test set. Therefore, it is feasible to use hyperspectral image technology to extract their spectral information for nondestructive identification and classification of maize seed varieties.**

## Keywords

**Hyperspectral imaging technology; maize species; variety classification; machine learning.**

## 1. Introduction

With the widespread use of technologies such as corn hybridization, corn species have increased rapidly, but many factors can lead to impure corn seeds. A decrease in the purity of corn seed will significantly affect its production yield. Therefore, identifying maize seed purity is an urgent problem to be solved[1]. Traditionally, maize seed varieties are determined mainly by seedling morphology, field planting identification, electrophoresis analysis, and other methods. However, these methods have the disadvantages of long identification time, strong dependence on personnel, destructive to the seeds, and difficulty to use universally[2]. With the development of image processing technology, the method of maize seed identification based on visible light machine vision technology has been widely emphasized by using the differences of different maize seeds in color, texture, shape, and other morphological characteristics. However, the traditional machine vision technology can only obtain the morphological information features of sources in visible light. As the number of seed varieties increases, the overlapping phenomenon of seed features becomes more serious, resulting in poor distinguishability of morphological features, which affects the recognition effect; at the same time, the extraction of morphological features has disadvantages such as more complex algorithms and poor real-time. In recent years, the use of near-infrared spectroscopy to study the identification methods of maize varieties has also received greater attention[3-4]. The basic idea is to use NIR spectroscopy to obtain information on the characteristics of all organic molecules containing hydrogen groups in maize samples and use this characteristic information to classify maize samples. However, NIR spectroscopy can only obtain local chemical information from maize samples. When the chemical information of the samples is not

uniformly distributed, it will certainly cause large testing errors. In contrast, the testing position of the samples, air humidity, and other factors greatly influence the experimental results[5].

Hyperspectral image technology is a new method of nondestructive testing that combines image technology with spectroscopic technology that has emerged in recent years. At present, the Application of hyperspectral image technology in the field of agricultural inspection has become a hot research topic at home and abroad, and the technology has been widely used in the fields of internal and external quality prediction, damage identification and safety detection of agricultural products, and production information acquisition of crops[6]. This paper investigates the hyperspectral image recognition and classification method for 90 corn seeds of three types of varieties. Good results were achieved by extracting the average reflection spectrum of the hyperspectral image region of interest and combining partial least squares discriminant analysis to discriminate and classify the maize seeds for identification.

## 2. Materials and methods

### 2.1. Experimental materials

The experimental seeds were purchased from the Guangdong Academy of Agricultural Sciences on Taobao and contained three types of maize varieties, including Yue xian nuo 6, Zheng tian 68, and Yue tian 29. Thirty kernels of uniform size and similar shape were selected for each variety, with a total of 90 samples, and the purity of each type of maize variety was ≥96%, and the clarity was ≥99%.

### 2.2. Experimental equipment

The experiment uses GaiaField-V10E hyperspectral device, which contains three parts: image acquisition unit, light source, and sample delivery platform, among which the image acquisition unit consists of an imaging spectrometer, CCD camera, lens, etc. The spectral range of the device's light source is 400-1000 nm, and the spectral resolution is 2.8 nm, containing a total of 360 bands.

### 2.3. Sample collection

Based on the size of the images that the system can acquire, it was decided to obtain 30 samples per image, and three images were developed. The initial values of the system were set: exposure time was set to 200 ms, the object distance was set to 30 cm without distortion, and the samples were placed on a 25 cm × 25 cm black background according to the principle that different colors absorb light differently and to reduce the effect of the background. In order to minimize the interference of the light source caused by the temperature during the experiment, one all-black calibration image and one all-white calibration image were acquired for every image in the experiment.

## 3. Data analysis and processing

### 3.1. Spectral image correction

In order to reduce the measurement errors caused by fluctuations in the light source and the variability in quantum efficiency of the CCD camera at different wavelengths, the hyperspectral image must be calibrated. The image obtained by scanning the standard whiteboard is R. Then, the image is captured with a lens covering the camera, and the all-black image D is scanned without light transmission, the original image is calibrated using the whiteboard image and the all-black image, and the calibration equation is

$$I = \frac{I_0 - D}{R - D} \tag{1}$$

ISSN: 2664-9640

where $I$ is the corrected image, $I_0$ is the original image, $R$ is the whiteboard image, and $D$ is the full black image.

### 3.2. Spectral Feature Extraction

Due to the huge amount of data contained in hyperspectral images, feature extraction is needed for image data. Considering that the extraction of morphological feature information in hyperspectral images is tedious. In contrast, the spectral information can reflect the physical structure and chemical composition of the measured object, this experiment only extracts its spectral data. Since manual spectra extraction is tedious, the maximum interclass variance algorithm is used in this paper to separate the foreground and background from obtaining a binary image. Then the average spectrum of the relevant region is extracted using a mask. Figure 1 shows the average spectral curves of 3 types of maize varieties. Figure 1 shows that the difference between the three types of curves is relatively obvious and possesses differentiability.

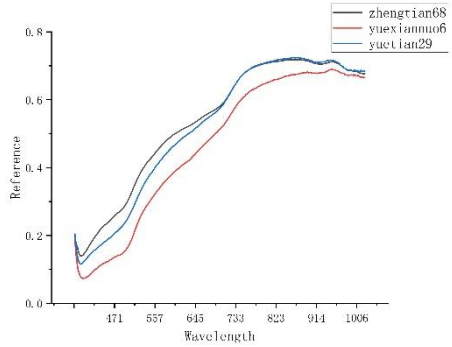



Fig. 1 Average reflected light intensity of different kinds of samples

### 3.3. Partial least squares discriminant analysis

The partial least squares discriminant analysis (PLSDA) is based on the PLS regression method, which linearly regresses the categorical feature parameters with the category information variables. Based on the model established in the training set, the value of the category information variable $(Y_p)$ of the test samples in the validation set is estimated; if $Y_p$ is greater than the threshold value, the sample is judged to belong to the class. Otherwise, it does not. The threshold value is obtained from the Bayesian function.

## 4. Result analysis and discussion

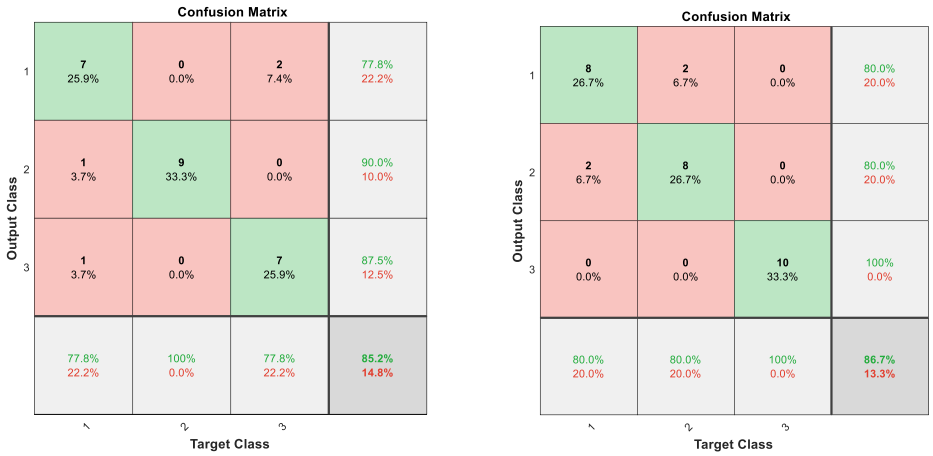
In the experiment, 3/4 of the samples in each class were randomly selected as the training set samples (63 seeds in total), and the rest were used as the validation set samples (27 seeds in total). The average reflected light intensity of 63 corn seeds in 360 bands was used as the feature variable, and the classification recognition model was established by combining PLSDA. In addition, SVM was used in the experiment for experimental comparison, and the experimental results are shown in Table 1.

Table 1 Classification identification results

| | Yue tian 29 | Yue xian nuo 6 | Zheng tian 68 | total |
|---|---|---|---|---|
| SVM | 77.8% | 100% | 77.8% | 85.2% |
| PLSDA | 80% | 80% | 100% | 86.7% |

The experimental results show that both PLSDA and SVM algorithms can effectively enter the identification of corn varieties, and their classification accuracy is 86.7% and 85.2%,

respectively, with PLSDA algorithm classifying slightly better than SVM algorithm. Figure 2 shows the confusion matrix of the algorithm classification, and its results are consistent with the experimental statistics. In addition, it can be seen that Yue tian 29 and Zheng tian 68 are easily misclassified.




(a)SVM (b)PLSDA

Fig. 2 Confusion matrix for both algorithms

## 5. Conclusion

This paper investigates the classification and identification of maize seeds using hyperspectral reflectance image techniques in the wavelength range of 400 - 1 000 nm. The type of maize varieties was achieved by extracting the average relative reflected light intensity of maize in the full wavelength region of interest as the classification feature, combined with the PLSDA classification model, and the classification accuracy of the final prediction set was 86.7%, which was higher than the SVM algorithm. Therefore, the extraction of spectral information of maize seeds by hyperspectral techniques can be well used to classify maize seeds and provide a reference for the Application of hyperspectral images on maize variety identification.

## References

[1] Wang L, Sun D W, Pu H, et al. Application of hyperspectral imaging to discriminate the variety of maize seeds[J]. Food Analytical Methods, 2016, 9(1): 225-234.

[2] Zhao Y, Zhu S, Zhang C, et al. Application of hyperspectral imaging and chemometrics for variety classification of maize seeds[J]. Rsc Advances, 2018, 8(3): 1337-1345.

[3] Rahman A, Cho B K. Assessment of seed quality using nondestructive measurement techniques: a review[J]. Seed Science Research, 2016, 26(4): 285-305.

[4] Ambrose A, Kandpal L M, Kim M S, et al. High speed measurement of corn seed viability using hyperspectral imaging[J]. Infrared Physics & Technology, 2016, 75: 173-179.

[5] Zhang J, Dai L, Cheng F. Corn seed variety classification based on hyperspectral reflectance imaging and deep convolutional neural network[J]. Journal of Food Measurement and Characterization, 2021, 15(1): 484-494.

[6] Yang J, Sun L, Xing W, et al. Hyperspectral prediction of sugarbeet seed germination based on gauss kernel SVM[J]. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2021, 253: 119585.