Non-destructive Detection of Blackheart Potatoes Based on Energy Spectrum of VIS/NIR Transmittance

Ya-fen Han, Cheng-xu Lv, Yan-wei Yuan*, Bing-nan Yang, Qing-liang Zhao, and You-fu Cao

Chinese Academy of Agricultural Mechanization Sciences, Beijing 100083, China

* Corresponding author: yyw215@163.com

Abstract. Blackheart disease can cause internal defects in potatoes, which seriously damages the quality and yield of processed potato products and causes considerable economic losses. To achieve online grading of potato, it is important to develop an accurate and reliable method to identify and separate blackheart potatoes non-destructively. A new method of transmitted energy spectrum determination was analysed and adopted innovatively for blackheart potato detection. Three methods, peak area linear discrimination analysis (PA-LDA), peak value linear discrimination analysis (PV-LDA), and peak difference value linear discrimination analysis (PDV-LDA), were used to identify blackheart potato, respectively. All of the three determination methods (PA between 657nm to 750nm normalized by $T_{580}$, $T_{698}$/T$_{657}$, $(T_{698}-T_{657})/T_{624}$) could identify blackheart potatoes accurately, and the classification accuracy were 91.69%, 92.43% and 93.69%, respectively. PDV-LDA ($(T_{698}-T_{657})/T_{624}$) showed the best performance, whose sensitivity, specificity, and AUC value were 94.86%, 95.19%, and 0.98, respectively. Compared with the general absorbance spectrum determination method, this method has simple operation, stable model, and can realize online detection fastly and accurately.

1. Introduction

Potato (Solanum tuberosum L.) plays an important role in both vegetable and processing uses. However, potatoes are easy to suffer from internal defects such as blackheart disease, when they stay in high carbon dioxide conditions or the environment temperature changes rapidly from high to low during storage [1]. It can seriously damage the quality and yield of processed potato products and cause considerable economic losses. General techniques to identify blackheart potatoes always involve destructive tests on a random sample set, which is both unpractical and inapplicable. Therefore, it is urgent to develop an accurate and reliable method to detect and separate blackheart potatoes non-destructively.

Near-infrared spectroscopy has been widely used in fruit and vegetable quality evaluation. Absorbance spectroscopy combined with chemometric algorithms is always used for quantitative analysis of potato nutrients content [2, 3] and qualitative analysis of potato internal defects [4, 5]. Energy spectrum also has been proved to be a powerful method detecting internal defects in fruits [6-10]. However, no related researches about energy spectrum detection has been applied for identifying blackheart potatoes yet.

In this paper, a new method of blackheart potato detection through energy spectrum has been adopted. Three methods, peak area linear discrimination analysis (PA-LDA), peak value linear discrimination analysis (PV-LDA), and peak difference value linear discrimination analysis (PDV-LDA), were used to identify blackheart potato respectively. Compared with the general determination methods using absorbance spectroscopy, this method does not require complex data conversion and pre-processing, can detect quickly online, and also eliminate the interference information introduced by inappropriate selection of reference.

2. Materials and methods
2.1 Potato materials
Potatoes were provided by China Southern Potato Research Center (Enshi, Hubei province, China). Blackheart potatoes were prepared in the laboratory. After being cleaned, air-dried and disinfected on the surface, potatoes were vacuum-packed and placed in an incubator at a temperature of 38 °C for 48 hours. Then they were immediately transferred to a refrigerator at 4 °C for 3d [11].

2.2 Spectroscopy measurement
A custom-designed laboratory system (Fig. 1) was used to acquire spectra in the transmittance mode. The system composed of a spectrometer, a light source, a potato holder for positioning the sample, an optical fibre combined with a collimating lens for efficient collection of the transmitted ray, and a personal computer for controlling the spectrometer and acquiring spectra.

Due to the high density, potato has a poor light transmission property. To obtain a good transmission spectrum, the spectrometer (PG2000, ideaoptics, China) was equipped with a Si detector covering the visible-near-infrared (vis-NIR) band (368 nm to 1039 nm) which had strong penetration and a slit width of 100μm which increased the amount of transmitted light. The light source was two halogen lamps with high power (50 W, 12 V, Philips Inc., China) to provide enough light intensity. The light source and the optical fibre probe were set on the left and right sides of the potato holder of the grading line. The long axis of the potato was vertical to the conveying direction of the grading line, which was the most stable placement method during potato transportation.

2.3 Assessment of flesh browning
Following Vis/NIR measurements, all potatoes were cut in half through the long axis to determine whether internal flesh browning had occurred. A total of 470 potatoes were used for spectral collection (234 sound potatoes and 236 blackheart potatoes). For model analysis, the random method was used to divide the sample set into a calibration set (176 sound potatoes and 177 blackheart potatoes) and a validation set (58 sound potatoes and 59 blackheart potatoes) according to the ratio of 3:1.

Besides, samples were divided into four grades according to the proportion of the blackheart area to the entire longitudinal cross-sectional area (1) sound; (2) slight: 0-20%; (3) moderate: 20-50%; (4) severe: more than 50%), which was used for spectral morphology analysis (Fig. 2).

2.4 Data analysis

2.4.1 Classification method
Spectrum data is affected by various factors such as size, hardness, maturity, chemical components, noise. In order to detect the blackheart defect more effectively, the normalized peak area (PA), peak value (PV) and peak difference value (PDV) were computed (Equations (1), (2), and (3)) [12, 13].

\[
PA = \frac{\int\left[F(\lambda) - f(\lambda)\right] T_\lambda}{0.35 \sum_{\lambda_i} \left[F(\lambda) - f(\lambda)\right] T_\lambda} \tag{1}
\]

where \(a\) and \(b\) are the wavelength range of peak area. \(F(\lambda)\) and \(f(\lambda)\) are the transmitted energy value curve equation and the baseline equation. \(F(\lambda_j)\) and \(f(\lambda_j)\) are the transmitted energy value at wavelength \(\lambda_j\) nm and the baseline value at wavelength \(\lambda_j\) nm. \(T_\lambda\) is transmitted energy value at wavelength \(\lambda\) nm used for normalization. 0.35 is the spectral data interval.

\[
PV = \frac{T_{\lambda_p}}{T_{\lambda_q}} \tag{2}
\]

\[
PDV = \frac{(T_{\lambda_p} - T_{\lambda_q})}{T_{\lambda_s}} \tag{3}
\]

where \(T_{\lambda_p}, T_{\lambda_q}, T_{\lambda_m}, T_{\lambda_n}\) and \(T_{\lambda_l}\) are peak values or trough values near the wavelength \(\lambda_p\) nm, \(\lambda_q\) nm, \(\lambda_m\) nm, \(\lambda_n\) nm and \(\lambda_l\) nm, respectively.

\(PA\) values in Equation (1) were obtained by Equation (4). The peak area between 657nm to 750nm (where the peak area of sound potato was distinctly larger than that of blackheart potato) was normalized by the peak values near 580 nm, 657 nm and the trough value near 624 nm.

\[
PA_i = \int_{580}^{750} \left[F(\lambda) - f(\lambda)\right] T_{\lambda_i} \; d\lambda, \quad PA_q = \int_{580}^{750} \left[F(\lambda) - f(\lambda)\right] T_{\lambda_q} \; d\lambda, \quad PA_m = \int_{580}^{750} \left[F(\lambda) - f(\lambda)\right] T_{\lambda_m} \; d\lambda
\]

\[
PV\) values in Equation (2) were acquired by Equation (5). The ratio of two peak values or peak value with trough value were used for classifying the potato samples into the sound and the blackheart.

\[
PV_1 = \frac{T_{\lambda_p}}{T_{\lambda_s}}, \quad PV_2 = \frac{T_{\lambda_p}}{T_{\lambda_l}}, \quad PV_3 = \frac{T_{\lambda_q}}{T_{\lambda_s}}, \quad PV_4 = \frac{T_{\lambda_q}}{T_{\lambda_l}}, \quad PV_5 = \frac{T_{\lambda_m}}{T_{\lambda_s}}, \quad PV_6 = \frac{T_{\lambda_m}}{T_{\lambda_l}} \tag{5}
\]

\(PDV\) values in Equation (3) were gained by Equation (6). The difference of two peak values or peak value with trough value were normalized by the spectral values near 580 nm, 624 nm, and 657 nm.

\[
PDV_1 = \frac{(T_{\lambda_p} - T_{\lambda_s})}{T_{\lambda_l}}, \quad PDV_2 = \frac{(T_{\lambda_p} - T_{\lambda_l})}{T_{\lambda_s}}, \quad PDV_3 = \frac{(T_{\lambda_q} - T_{\lambda_s})}{T_{\lambda_l}}, \quad PDV_4 = \frac{(T_{\lambda_q} - T_{\lambda_l})}{T_{\lambda_s}}, \quad PDV_5 = \frac{(T_{\lambda_m} - T_{\lambda_s})}{T_{\lambda_l}}, \quad PDV_6 = \frac{(T_{\lambda_m} - T_{\lambda_l})}{T_{\lambda_s}} \tag{6}
\]

The classic linear discrimination analysis (LDA) was applied for the model discrimination. Sound potatoes were assigned a value of -1 and blackheart potatoes a value of 1 for the \(Y\) matrix. The discrimination threshold was set as 0.

2.4.2 Evaluation of model performance

The performance of the classification models was assessed based on the receiver operating characteristic (ROC) analysis, including sensitivity (%), specificity (%), classification accuracy (\(C\), %) and area under the curve (AUC) [13, 14]. Class sensitivity describes the model ability to correctly recognize samples belonging to that class. The class specificity describes the model ability to reject samples of all other classes. \(C_{cal}, C_{val}, C\) referred to the classification accuracy of calibration set, validation set, and total set, respectively. The ROC curve is a plot of ‘sensitivity’ and ‘1-specificity’. The area under the curve (AUC) values is generally used as an indicator of the discriminative ability of a classifier. An AUC value close to 1 indicates a strong discriminative ability, and an AUC value close to 0.5 indicates that the classifier has little discriminative power [13, 15].

All the data analysis was based on Matlab R2016a (Mathworks, Natick, MA, USA).

3. Results and discussion

3.1 Characteristics of the transmitted energy spectrum

Fig. 3 shows the mean transmitted energy spectrum of the sound and blackheart potatoes in different blackheart grades in the wavelength range of 368 nm to 1039 nm. For blackheart potatoes, two peaks were observed near 580nm and 657 nm and an absorption band near 624 nm. However, for sound potatoes, there were one main peak near 698nm and some small peaks near 580nm and 657 nm.
Fig. 3. Mean transmitted energy spectra of potatoes with different blackheart grades.

Fig. 4 shows the mean spectrum of potatoes after normalized by the trough values near 624 nm. After normalization, the spectral characteristics of potatoes in different grades were more obvious. As the blackheart area increased, the transmitted energy increased below 624nm, while that decreased above 624nm. The spectral difference between sound and blackheart potatoes was magnified above 624nm after normalization. Moreover, it was noticed that the value at the peak near 698 nm showed the most apparent difference between sound and blackheart potatoes after normalization. Furthermore, the peak area (between 657nm and 750nm) of sound potatoes was larger than that of blackheart potatoes, which could also be used to separate out the blackheart samples. Therefore, in order to compensate for the additive and multiplicative effects on raw spectra and optimize the signal-to-noise ratio, normalization methods were used to pre-treat the raw energy spectra. PA-LDA, PV-LDA, PDV-LDA were used to classify sound and blackheart potatoes.

3.2 Classification of potato samples by PA-LDA

The peak area between 657nm to 750nm was normalized by the peak values near 580 nm, 657nm and the trough value near 624 nm (Table 1). Among the three pre-treatments, PA1 (peak area between 657nm to 750nm normalized by the peak value near 580 nm) showed the most favourable result. The AUC value was 0.96, indicating that the model had a strong discriminative ability. The classification accuracy reached 91.69%.

Table 1. Classification results of blackheart and sound potatoes by PA-LDA models.

| Pretreatment | Calibration set | Validation set |
|--------------|----------------|----------------|
|              | Sensitivity(%) | Specificity(%) | AUC  | C_cal(%) | C_val(%) | C(%) |

Fig. 4. Mean transmitted energy spectra of potatoes normalized by the trough value near 624 nm.
3.3 Classification of potato samples by PV-LDA

The ratio of two peak values or peak value with trough value were used for identifying blackheart potatoes (Table 2). Among the classification results below, PV3 (T_{698}/T_{657}) presented the best result. The classification accuracy reached 92.43%, which was better than that of PA-LDA.

3.4 Classification of potato samples by PDV-LDA

The difference of two peak values or peak value with trough value were normalized by the spectral values near the 580 nm, 624 nm, and 657nm (Table 3). Among the classification test results, it was noted that PDV2 ((T_{698} - T_{657})/T_{580}), PDV3 ((T_{698} - T_{657})/T_{624}), and PDV5 ((T_{698} - T_{624})/T_{657}) performed strong discriminative ability and high classification accuracy. PDV3 ((T_{698} - T_{657})/T_{624}) performed the best, whose sensitivity, specificity, AUC value and classification accuracy were 94.86%, 95.19%, 0.98, and 93.69%, respectively (Fig. 5), more efficient than PA-LDA and PV-LDA.

### Table 2. Classification results of blackheart and sound potatoes by PV-LDA models.

| Pretreatment | Calibration set | Validation set |
|--------------|-----------------|----------------|
|              | Sensitivity(%)  | Specificity(%) | AUC | C_cal(%) | C_val(%) | C(%) |
| PV1          | 94.86           | 86.54          | 0.97 | 91.76    | 86.96    | 89.36 |
| PV2          | 93.71           | 93.27          | 0.98 | 93.55    | 89.13    | 91.34 |
| PV3          | 92.57           | 95.19          | 0.98 | 93.55    | 91.30    | 92.43 |
| PV4          | 86.29           | 97.12          | 0.98 | 90.32    | 90.22    | 90.27 |
| PV5          | 83.43           | 98.08          | 0.97 | 88.89    | 88.04    | 88.47 |

### Table 3. Classification results of blackheart and sound potatoes by PDV-LDA models.

| Pretreatment | Calibration set | Validation set |
|--------------|-----------------|----------------|
|              | Sensitivity(%)  | Specificity(%) | AUC | C_cal(%) | C_val(%) | C(%) |
| PDV1         | 95.43           | 89.42          | 0.98 | 93.19    | 88.04    | 90.62 |
| PDV2         | 96.57           | 91.35          | 0.98 | 94.62    | 92.39    | 93.51 |
| PDV3         | 94.86           | 95.19          | 0.98 | 94.98    | 92.39    | 93.69 |
| PDV4         | 92.00           | 92.31          | 0.97 | 92.11    | 89.13    | 90.62 |
| PDV5         | 92.00           | 96.15          | 0.98 | 93.55    | 93.48    | 93.52 |
| PDV6         | 90.29           | 96.15          | 0.97 | 92.47    | 89.13    | 90.80 |
Fig. 5. Classification result of blackheart and sound potatoes by using the normalized peak differences \( (T_{698}-T_{657})/T_{624} \)

4. Conclusion
The blackheart disease of potato developing during storage and transportation is a serious industry problem. Transmitted energy spectrum of potatoes was analysed and adopted innovatively for determining blackheart potato. Three methods, PA-LDA, PV-LDA, and PDV-LDA, were used to identify blackheart potato respectively. All of the three determination methods could detect blackheart potatoes accurately. PDV-LDA \((T_{698} - T_{657})/T_{624}\) showed the best performance, whose sensitivity, specificity, AUC value and classification accuracy were 94.86%, 95.19%, 0.98, and 93.69%, respectively. Compared with the general absorbance spectrum determination method, this method has simple operation, stable model, and can realize online detection fastly and accurately.

Acknowledgments
This study was financially supported by Modern Agricultural Industry Technology System (CARS-10).

References
[1] H. F. Bergman, Bot. Rev. 25(3) 417-485 (1959)
[2] F. Wang, Y. Li, Y. Peng, B. Yang, L. Li, Y. Liu, Spectrosc. Spectr. Anal. 38(12) 3736-3742 (2018) (in Chinese)
[3] X. Dong, X. Sun, Spectrosc. Spectr. Anal. 33(12) 3216-3220 (2013) (in Chinese)
[4] Z. Zhou, S. Zeng, X. Li, J. Zheng, J. Spectrosc. 2015 Paper No. 786709 (2015).
[5] X. Zhang, B. Yang, Y. Cao, S. Li, Q. Zhao, L. Xing, Spectrosc. Spectr. Anal. 38(08) 2379-2385 (2018) (in Chinese)
[6] X. Sun, Y. Liu, Y. Li, M. Wu, D. Zhu, Postharvest Biol. Technol. 116(2016) 80-87 (2016)
[7] J. Wang, X. Sun, L. Pan, Q. Sun, D. Han, Spectrosc. Spectr. Anal. 28(9) 2098-2102 (2008) (in Chinese)
[8] D. S. Ryu, S. Y. Lee, S. H. Noh, Proc. Acad. (PLAZA, FOOMA Japan) 12 243-246 (2005)
[9] D. S. Ryu, S. H. Noh, H. Hwang, Int. J. Agric. Biosyst. 4(1) 1-7 (2003)
[10] Y. Liu, Y. Li, Z. Gong, X. Sun, Spectrosc. Spectr. Anal. 37(12) 3714-3718 (2017) (in Chinese)
[11] F. Tian, Y. Peng, W. Wei, X. Zheng, W. Wang, Trans. CSAE, 33(5) 287-294 (2017) (in Chinese)
[12] Y. W. Seo, S. H. Noh, K. J. Lee, ASABE, Paper No. 066121 (2006)
[13] M. Li, J. Wang, W. Liu, L. Wu, C. Wan, S. Yang, D. Han, Int. Agric. Eng. J. 27(2) 273-281 (2018)
[14] Z. Li, H. Lu, J. Yang, X. Zeng, L. Zhao, H. Li, Q. Liao, S. Peng, M. Zhou, M. Wu, J. Xiang, Y. Wang, G. Li, Clin. Chim. Acta, 425 221-226 (2013)
[15] D. Ballabio, V. Consonni, Anal. Methods, 5 3790-3798 (2013)