Pattern recognition-based optical technique for non-destructive detection of *Ectomyelois ceratoniae* infestation in pomegranates during hidden activity of the larvae

Bahareh Jamshidi a,⁎, Ezeddin Mohajerani b, Hossein Farazmand c, Asghar Mahmoudi d, Abolfazl Hemmati d

a Agricultural Engineering Research Institute, Agricultural Research Education and Extension Organization (AREEEO), Karaj, Iran
b Laser and Plasma Research Institute, Shahid Beheshti University, Tehran, Iran
c Iranian Research Institute of Plant Protection, Agricultural Research Education and Extension Organization (AREEEO), Tehran, Iran
d Department of Biosystems Engineering, Faculty of Agriculture, University of Tabriz, Tabriz, Iran.

ARTICLE INFO

Article history:
Received 26 May 2018
Received in revised form 31 July 2018
Accepted 28 August 2018
Available online 29 August 2018

Keywords:
Carob moth
Discriminant analysis
Insect infestation
Near-infrared
Non-destructive
Pomegranate
Spectroscopy

ABSTRACT

In this research, the feasibility of utilizing visible/near-infrared (Vis/NIR) spectroscopy as an optical non-destructive technique combined with both supervised and unsupervised pattern recognition methods was assessed for detection of *Ectomyelois ceratoniae*, carob moth, infestation in pomegranates during hidden activity of the larvae. To this end, some fruits were artificially contaminated to the carob moth larvae. Vis/NIR spectra of the blank samples and the contaminated pomegranates without and with external visual symptoms of larvae infestation were analyzed one and two weeks after contaminating the samples as three groups of “Healthy”, “Unhealthy-A” and “Unhealthy-B”, respectively. Principal component analysis (PCA) as an unsupervised pattern recognition method was used to verify the possibility of clustering of the pomegranate samples into the three groups. Discriminant analysis (DA) based on PCA was also used as a powerful supervised pattern recognition method to classify the samples. The calibration models of linear, quadratic and Mahalanobis discriminant analyses were developed based on different spectral pre-processing techniques. The best PCA-DA model was obtained using Mahalanobis distance method and first derivative (D1) pre-processing. The total percentage of correctly classified samples with the best calibration model was 97.9%. The developed model could also predict unknown samples with total percentage of correctly classified samples of 90.6%. It was concluded that Vis/NIR spectroscopy combined with pattern recognition method of PCA-DA can be an appropriate and rapid technology for non-destructively screening the pomegranates for detection of carob moth infestation during hidden activity of the larvae.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

The carob moth, *Ectomyelois ceratoniae* Zell. (Lepidoptera: Pyralidae), is a destructive worldwide polyphagous insect and the most important pest of pomegranate (*Punica granatum*) in the Middle East, especially in Iran, attacking the fruits before and after harvest and causing 30–80% yield losses. Chemical insecticides are not applicable for controlling this pest because of the hidden activity of the larvae. Non-chemical control methods which are currently used, have no sufficient efficiency. This pest normally lays eggs inside the crown, calyx, of pomegranates. Larvae penetrate into the fruit after hatching. The damage caused by larvae, especially from second or third age, on the fruit is due to their feeding from internal parts of pomegranate without external symptoms. This causes penetration of pathogenic fungi such as *Aspergillus* and *Penicillium* which makes the fruits unmarketable and unfit not only for human consumption but also for the food processing industries [1–4]. Sometimes, the appearance of black spots on the pomegranate is the first symptom of carob moth infection and the beginning of fruit decaying process [4]. However, there is mostly no external visual symptom during hidden activity of the larvae inside the fruit to detect infested pomegranates. Thus, the pomegranates with hidden infestation may pass undetected in packing houses and processing lines. They may damage the surrounding healthy fruits during storage. Moreover, the existence of hidden contamination in pomegranates is a vital challenge for the exportation. Therefore, development of a fast and non-destructive detection technique of infested pomegranates is imperative.

Few publications have addressed the use of some non-destructive techniques for the detection of internal insects and insect infestation on pomegranates. The feasibility of using X-ray computed tomography (CT) coupled with image analysis has been investigated by Magwaza and Opara [5] for non-destructive detection internal structure of pomegranates. Arendse et al. [6] investigated the application of micro-focus X-ray CT (μCT) using a density calibration function for non-destructive
detection of internal defects caused by false codling moth and black-heart in pomegranates. Although X-ray CT scanning technique provides promising results to detect the presence of internal insect larva, disease infestation and disorders, it is not economic and practical for in/on-line applications. This technique is considerably complex and expensive compared with other non-destructive methods. Health and safety issues may arise due to harmful effects of the radiation. In addition, safety of operators, considerable large equipment size and time required for tests, data acquisition, processing and analysis, are of concern [5–9].

Nuclear magnetic resonance spectroscopy (NMR) and nuclear magnetic resonance imaging (MRI) have been shown to be effective techniques for non-destructive internal structure and quality assessment in pomegranates [10–12]. Although NMR-based techniques promise good results to discriminate healthy and defective fruits, they are less sensitive than other analytical methods. NMR systems are not economic and practical for in/on-line use in fruit packing houses because of the high-cost constraints, complexity and potential health hazard due to maintenance of magnetic [7].

Near-infrared (NIR) spectroscopy is one of the most promising non-destructive methods which is flexible for food qualitative and quantitative analyses especially for fruits. While NMR and X-ray CT techniques are able to show only the internal structure of the fruit, not the compositional or nutritional details, NIR spectroscopy is very successfully being used to measure the compositional quality of a fruit [13]. It is rapid, safe, non-contaminant and can be used in processing lines [14,15]. Compare to NMR and X-ray CT systems, NIR spectroscopy technique is very low-cost. Visible/near-infrared (Vis/NIR) and NIR spectroscopy have been widely used for rapid internal quality and safety assessment and chemical compounds prediction of both thin-rind and thick-rind fruits [16–23]. Few works have recently addressed the use of these techniques for detection of internal insect larvae or insect infestation in fruits such as tart cherries [9], wild blueberries [24], Mangoes [25], jujubes [26,27], Chestnuts [28,29] and olives [30,31]. All these researches confirm the possibility and reasonability of detection of hidden insect infestation in such fruits using NIR-based spectroscopy. Relating to pomegranate fruit, it is more difficult to detect hidden infestation especially in primary stages because of the complexity and variety of internal different parts. In spite of that, good results have been reported by Khodabakhshian et al. [32] on carob moth damage detection in pomegranates (Ashraf variety) during four maturity stages using reflectance Vis/NIR spectroscopy. In their research, the pomegranates containing different infection levels were selected based on visual observation of the symptoms of defect on fruit at each maturity stage. However, there is mostly no external visual symptom during hidden activity of the larvae inside the fruit. Moreover, selecting the infested pomegranates based on visual observation is not a precise method for confidence about the carob moth infestation and the symptoms may be due to other factors.

This research aims to investigate the capability of Vis/NIR spectroscopy based on both supervised and unsupervised pattern recognition techniques for non-destructive detection of internal carob moth infestation in the harvested pomegranates (Malas Saveh, an export variety) during hidden activity of the larvae inside the fruits without any symptom until the appearance of the symptoms on them for the first time. To be sure about the contamination of the pomegranates to the carob moth no other insects’ larvae, fruits were artificially contaminated to the carob moth larvae. To reduce the rind color effect of the fruit on detection results, the spectroscopy measurements were conducted in the interactance mode.

2. Materials and Methods

2.1. Pomegranate Samples

Pomegranate samples (cv. Malas Saveh with red aril, thin-rind and sour-sweet delicious taste) which were free from external defects and symptoms of insect infestation were prepared during October 2017 from the pomegranate orchards in Saveh region, central part of Iran.

![Fig. 1. Contaminating the samples to the carob moth larvae. The pomegranates after cutting the above edges of the calyx and removing the stamen clusters of it (a). Putting the carob moth larva of second age into the calyx after creating a small hole (b). The closed calyx with a small mesh piece after contaminating the sample to the carob moth larva (c).](image)

![Fig. 2. Spectra acquisition of a blank sample at the position around and near the calyx (a). Vis/NIR absorbance spectra of all healthy samples and unhealthy pomegranates during decaying process with the carob moth larva (b).](image)
The fruits were stored in 18 °C and appropriate humidity for 6 days to appear any symptom if they had any hidden insect infestation. After that, a total of 70 safe and healthy pomegranate fruits with similar color, size and shape were selected for the experiments. Ten of the pomegranates were labeled as blank samples and the remained fruits were contaminated to the carob moth larvae artificially.

2.2. Contaminating the Samples to Carob Moth Larvae

First, the above edges of the calyx in the pomegranate except the blank samples were cut with a cutter (after cleaning it up with alcohol) to make a flat edge. Then, the stamen clusters were removed from the calyx. A thin metal rod with a sharp point on the end (after cleaning it up with alcohol) was used to create a very small hole in the calyx of the pomegranates for fast penetrating the larvae into the fruit. The carob moth larva of second age, the best age for penetration into the fruit and causing more damages, with 2–4 mm length and after 24 h of starvation were put into the calyx of each fruit quietly with a soft painting-brush. While the larvae found the holes and penetrated into the fruits after minutes, the above of the calyx was closed for 48 h with a small mesh piece in each contaminated fruit to be sure that they cannot leave the fruit. Fig. 1 shows some steps of contaminating the samples to the carob moth larvae. All the samples containing healthy and unhealthy (contaminated) pomegranates were kept in the same conditions and temperature about 25 °C before the experiments.

2.3. Vis/NIR Spectra Acquisition

First experiment was conducted one week after contaminating the samples to carob moth larvae when the contaminated pomegranates had no external visual symptoms of carob moth infestation during hidden activity of the larvae inside the fruit. Vis/NIR spectra of both blank and the contaminated pomegranates were collected using the spectroscopy set-up containing a USB2000 spectrometer (Oceanoptics Inc., USA) with charge coupled device (CCD) detector, a tungsten halogen light source (LS-1, Oceanoptics Inc., USA) and a fiber optic of P400-7-Vis-NIR model (Oceanoptics Inc., USA) to acquire the reflectance spectra of the samples in interactance mode at the range of 500–1000 nm with 1.5 nm resolution. Data acquisition for each fruit was done at four positions around and near the calyx (Fig. 2a) with five scans at each position using OOIBase32 software (Oceanoptics Inc., USA). For each sample, the mean Vis/NIR spectrum was calculated from a total of 20 scans and converted to absorbance values (log 1/R). One week after the first experiment when most of the contaminated samples had some external visual symptoms of larvae infestation such as spots, softening the texture or decaying of the pomegranate near the calyx, the second experiment was conducted for all the samples and the absorbance Vis/NIR spectra of them were collected as the same. Therefore, 140 index spectra (20 spectra for the blank samples and 120 spectra for the contaminated pomegranates containing 60 spectra during hidden activity of the larvae inside the fruit and 60 spectra after appearing the external symptoms of larvae infestation) were acquired from both healthy and unhealthy samples (Fig. 2b).

2.4. Qualitative Analysis

After spectra measurements, qualitative analyses based on pattern recognition methods were conducted to assess the feasibility of Vis/NIR spectroscopy technology for non-destructive detection of carob moth infestation in pomegranates during hidden activity of the larvae. To this end and before performing any spectral pre-processing method, outliers which are the samples containing interferences with a negative influence on model development were removed [15]. After running principal component analysis (PCA), they were detected using the Q-residual versus Hotelling’s T² plot (Fig. 3). The Q-residual on the ordinate axis describes the sample distance to model. The Hotelling’s T² explains how well the sample is described by the model [33]. The red lines in Fig. 3 display the associated critical limits with significance levels of 5% and the marked samples outside these limits are the detected outliers. This plot was constructed with the third PC where total residual variance of PCA model goes to zero with as few components as possible.

After removing the outliers (13 samples) and to verify the possibility of clustering of the pomegranate samples in three groups of “Healthy”, unhealthy without any external symptoms “Unhealthy-A”, and unhealthy with the external symptoms “Unhealthy-B”. PCA model was performed on the remaining samples (127 samples) as an unsupervised pattern recognition method. PCA provides a visual representation of the relationships between samples and variables and makes insights into how measured variables cause some samples to be similar to, or how they differ from each other.

Discriminant analysis (DA), the simplest of all possible classification techniques which are based on Bayes’ formula, as a powerful supervised pattern recognition method was also used to classify the samples. Before developing DA model, some pre-processing methods such as moving average (MA) filter with segment size of five for averaging, standard normal variate correction (SNV), and first derivative (D1) of the spectra based on Savitzky-Golay (SG) algorithm with five smoothing points and polynomial order of two were performed for smoothing and denoising the spectra, normalizing and improving the spectral resolution, respectively. The samples of each group were randomly divided into calibration (approximately 75% of the samples) and validation (the remained 25% of the samples) sets. Therefore, 95 and 32 samples...
containing all three groups were selected for calibration and validation sets, respectively. The calibration set was contained 13, 38 and 44 samples of “Healthy”, “Unhealthy-A”, and “Unhealthy-B” groups, respectively. Moreover, 4, 13 and 15 samples of these three groups were selected for the validation set, respectively. Then, DA based on PCA with 5 components using different methods of linear, quadratic and Mahalanobis distance with different pre-processing techniques of MA, SNV, and D1. Finally, the developed models were used to classify unknown samples. The closer the samples are in the scores plot, the more similar they are with respect to these two PCs concerned. On the contrary, samples far away from each other are different from each other. As it can be seen in Fig. 5, PC1 and PC2 explained 84% and 14% of all data variance, respectively. The three groups of “Healthy”, “Unhealthy-A”, and “Unhealthy-B” samples are also seen in different colors and symbols. The PCA clustering result indicates a clear discrimination between the groups of “Healthy” and “Unhealthy-B” as well as between the classes of “Unhealthy-A”, and “Unhealthy-B”. Therefore, the contaminated samples with the external symptoms of carob moth larvae infestation were well separated from the healthy and contaminated pomegranates with no external symptoms during the hidden activity of the larvae. However, the samples of “Healthy” and “Unhealthy-A” classes had some overlapping and were not well distinguished. The similarities between these two groups could be due to weak or slow activity of the larvae in some contaminated fruits which makes them more similar to healthy samples in terms of spectral information. Therefore, a powerful classifier of PCA-DA as a supervised pattern recognition method was also applied to make separations more clear.

3.3. PCA-DA Classification

Table 1 indicates the results of PCA-DA models developed based on different methods of linear, quadratic and Mahalanobis distance with different pre-processing methods for discriminating the pomegranates in three groups of “Healthy”, “Unhealthy-A” and “Unhealthy-B”.

<table>
<thead>
<tr>
<th>Pre-processing</th>
<th>DA</th>
<th>Calibration set</th>
<th>Validation set</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Correctly classified (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Healthy</td>
<td>Unhealthy-A</td>
<td>Unhealthy-B</td>
</tr>
<tr>
<td></td>
<td>Linear</td>
<td>69.2</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quadratic</td>
<td>84.6</td>
<td>97.4</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MA</td>
<td>84.6</td>
<td>97.4</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Linear</td>
<td>69.2</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quadratic</td>
<td>84.6</td>
<td>97.4</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SNV</td>
<td>84.6</td>
<td>97.4</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Linear</td>
<td>76.9</td>
<td>92.1</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quadratic</td>
<td>84.6</td>
<td>89.5</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mahalanobis</td>
<td>84.6</td>
<td>86.8</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Linear</td>
<td>53.8</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quadratic</td>
<td>84.6</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mahalanobis</td>
<td>84.6</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Fig. 5 shows the scores plot of the first component (PC1) versus the second component (PC2) obtained from PCA for clustering of the pomegranate samples into the three groups. These two PCs summarize more variation in the data than any other pair of components. Therefore, this plot can be used to interpret differences and similarities among the samples. The closer the samples are in the scores plot, the more similar they are. As it can be seen in Fig. 5, PC1 and PC2 explained 84% and 14% of all data variance, respectively. The three groups of “Healthy”, “Unhealthy-A”, and “Unhealthy-B” samples are also seen in different colors and symbols. The PCA clustering result indicates a clear discrimination between the groups of “Healthy” and “Unhealthy-B” as well as between the classes of “Unhealthy-A”, and “Unhealthy-B”. Therefore, the contaminated samples with the external symptoms of carob moth larvae infestation were well separated from the healthy and contaminated pomegranates with no external symptoms during the hidden activity of the larvae. However, the samples of “Healthy” and “Unhealthy-A” classes had some overlapping and were not well distinguished. The similarities between these two groups could be due to weak or slow activity of the larvae in some contaminated fruits which makes them more similar to healthy samples in terms of spectral information. Therefore, a powerful classifier of PCA-DA as a supervised pattern recognition method was also applied to make separations more clear.
According to Table 1, all the developed models of PCA-DA based on all three methods of linear, quadratic and Mahalanobis had excellent ability to discriminate the samples into the three groups. The total percentage of correctly classified samples was above of 92% for all the calibration models developed using the different pre-processing methods. Therefore, it was concluded that Vis/NIR spectral information of the pomegranates around and near the calyx combined with supervised pattern recognition of PCA-DA can be useful for non-destructive detection of internal carob moth infestation in the pomegranates during hidden activity of the larvae inside the fruits without any symptom until the appearance of the symptoms on them. The results of classification using the calibration models developed with pre-processing of MA were similar to those achieved with no pre-processing method. In these calibration models, methods of quadratic and Mahalanobis distance could model the classes better than linear method (accuracy of 96.8% versus 95.8%). It was noted that the variability within the groups is not the same structure. The classification results indicated that SNV pre-processing method decreases the percentage of correctly classified samples of “Unhealthy-A” group and the total percentage of correctly classified pomegranates. Moreover, the calibration models of PCA-DA based on linear and quadratic methods developed after pre-processing of D1 caused decreasing in the percentage of correctly classified samples of “Healthy” group and the total percentage of correctly classified pomegranates in comparison with the results obtained from the models developed with no pre-processing or after MA pre-processing. However, the best results of classification were achieved using the calibration model of PCA-DA developed based on Mahalanobis method and D1 pre-processing. The total percentage of correctly classified samples with the best calibration model was 97.9%. This conclusion was in agreement with the results reported by Moscetti et al. [28] who assessed the feasibility of NIR spectroscopy to detect hidden insect damage in chestnuts using a genetic algorithm in combination with a linear discriminant analysis. They found that the best classification of the infested and non-infested chestnuts can be achieved using D1 pre-processing method based on SG algorithm.

The confusion matrix of the best classification model is presented in Table 2. All the “Unhealthy-A” and “Unhealthy-B” samples in calibration set were correctly attributed to their groups. Only 2 samples with actual class of “Healthy” in calibration set were misclassified and predicted as “Unhealthy-A” group. Therefore, the percentage of correctly classified samples of “Healthy” group was 84.6%.

The discrimination results of the best PCA-DA developed model are shown in Fig. 6 for the calibration set. Every sample is displayed, color-coded by class, and the axes are for two classes of “Healthy” and “Unhealthy-A” in the model. Samples lying close to zero for a class are associated with the class.

As it can be seen in Table 1, all the calibration models of PCA-DA developed based on different methods of linear, quadratic and Mahalanobis with different pre-processing techniques had excellent ability to discriminate the unknown samples of the validation set into the three groups of “Healthy”, “Unhealthy-A” and “Unhealthy-B”. The total percentage of correctly classified unknown samples was above of 87.5% for all the models developed using the different pre-processing methods. According to the confusion matrix (Table 2), all the “Healthy” and “Unhealthy-B” samples in validation set were correctly classified to their groups. However, 3 samples with actual class of “Unhealthy-A” in validation set were predicted as “Healthy” group. Therefore, the percentage of correctly classified samples of “Unhealthy-A” group was 76.9%.

Consequently, the results indicated the potential of interactance Vis/NIR spectroscopy at the range of 500–1000 nm combined with pattern recognition method of PCA-DA based on Mahalanobis distance method for non-destructively screening the pomegranates to detect carob moth infestation during hidden activity of the larvae. This conclusion was in agreement with the results reported by Wang et al. [26,27] who assessed the feasibility of Vis/NIR spectroscopy to detect internal

---

**Table 2**

Confusion matrix of classification using the best PCA-DA calibration model developed based on Mahalanobis method and D1 pre-processing.

<table>
<thead>
<tr>
<th>Class</th>
<th>Actual (calibration set)</th>
<th>Actual (validation set)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Healthy</td>
<td>Unhealthy-A</td>
</tr>
<tr>
<td>Predicted</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Healthy</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>Unhealthy-A</td>
<td>2</td>
<td>38</td>
</tr>
<tr>
<td>Unhealthy-B</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

---

**Fig. 6.** Discrimination plot of the best PCA-DA developed model based on Mahalanobis method.
insect-infested jujubes using Mahalanobis discriminant analysis and found the best results in intractance mode.

4. Conclusion

The feasibility of using visible/near-infrared (Vis/NIR) spectroscopy combined with pattern recognition methods was investigated for detection of carob moth infestation in pomegranates (cv. Malas Saveh) during hidden activity of the larvae. The achieved results confirmed the capability of this technology for discrimination of healthy pomegranates from contaminated fruits to carob moth larvae with and without external symptoms of larvae infestation. Therefore, Vis/NIR spectroscopy can be useful for fast, low-cost and non-destructive screening and preliminary health control of the pomegranates. Nevertheless, further works should be considered on other varieties of pomegranates at different regions which naturally contaminated to carob moth larvae for adapting the Vis/NIR spectroscopy to detect carob moth infestation during hidden activity of the larve.

Acknowledgement

This research was funded by Iran National Science Foundation (INSF), project No. 94018222 and supported by Laser & Plasma Institute of Shahid Beheshti University. The authors are grateful to Iranian Research Institute of Plant Protection for the cooperation in preparing the contaminated samples.

References