International Conference of Agricultural Engineering
AgEng 2014 Zurich 6–10 July

Ref: C0392

Detection of decay in citrus fruit using absorption and scattering properties

Delia Lorente, Florentino Juste and José Blasco, Centro de Agroingeniería, Instituto Valenciano de Investigaciones Agrarias (IVIA), Cra. Moncada-Náquera km 5, 46113 Moncada (Valencia), Spain

Manuela Zude and Christian Regen, Leibniz-Institute for Agricultural Engineering Potsdam-Bornim (ATB), Max-Eyth-Allee 100, 14469 Potsdam-Bornim, Germany

Juan Gómez-Sanchis, University of Valencia -Intelligent Data Analysis Laboratory-, Av. Universitat S/N, 46100 Burjassot (Valencia), Spain

Abstract

The detection of fungal infection of citrus fruits in packinghouses is currently performed manually by trained workers illuminating each fruit with dangerous ultraviolet lighting to enhance the contrast between sound tissue and early infection symptoms, such as the water-soaking of the apoplast, the automation of this task being still challenging. The objective of this research was to study the potential of detecting decay in citrus fruits infected with *Penicillium digitatum* by analysing the absorption and reduced scattering coefficients of sound and damaged peel of the fruits using the laser-light backscattering imaging technique. Backscattering images of sound and decaying surface parts of oranges cv. 'Valencia late' were obtained using laser diode modules emitting at five wavelengths. Since the images of backscattered light had radial symmetry with respect to the incident point of the laser beam, they were reduced to a one-dimensional profile through radial averaging. A diffusion theory model was used for determining the absorption and the reduced scattering coefficients from the radial profiles at each wavelength, resulting in ten features that characterised each skin sample. A feature selection method based on mutual information was employed to select the most relevant optical coefficients for the detection of decay in the fruits. The selected coefficients were used as input feature vector for discriminating between sound and decaying skin using a supervised classifier based on linear discriminant analysis. The resulting set of selected features included the two optical coefficients at four laser wavelengths, remaining the optical coefficients at 532 nm outside the selected features, and yielded an average classification success rate of 92.4% with an even higher number of well-classified decaying samples (94.3%). These results highlight the potential of the backscattering imaging technique as an alternative for detecting decay lesions in citrus fruits successfully.

Keywords: fruit inspection, citrus fruit, backscattering imaging, optical properties, diffusion theory model

1. Introduction

Early detection of fungal infections in citrus postharvest is regarded as a primary concern in commercial packinghouses, since they cause severe economic losses in the citrus industry. At the present time, this detection is done by trained workers illuminating each fruit with dangerous ultraviolet (UV) illumination. Nevertheless, this method is harmful for operators, since long exposure to UV radiation can cause damaging effects to the human skin. The use of machine vision systems can be considered as a possible alternative to human inspection for detecting decayed fruit automatically. In this sense, vision systems based on colour cameras are currently used in the citrus industry for detecting external defects visible at first glance (Blasco et al., 2007). However, the detection of decay at early stages is a complicated task for these systems, since the appearance of the damage is very similar to sound skin, thus being hardly visible to the human eye. Therefore, other machine vision technologies have been proposed as a solution for automatic decay detection, such as the use of UV-induced fluorescence (Kurita et al., 2009) or hyperspectral and multispectral sensors (Gómez-Sanchis et al., 2012).

Light backscattering imaging (LBI) has recently emerged as an alternative machine vision technique for fruit inspection. In particular, many works have been reported on assessing quality of different fresh fruit by LBI systems (Qing et al., 2007; Romano et al., 2008). When a light beam interacts with a fruit, a small portion is reflected on the surface of the sample (Fresnel scattering) and the rest penetrates into the tissue. Most of the entering light is scattered backward to the exterior tissue surface after interacting with the internal components of the fruit (diffuse scattering or backscattering), whereas the rest of radiation is absorbed by tissue or transmitted out from the fruit. Light absorption is mainly related to chemical components of the fruit. By contrast, light scattering is affected by the structural and physical properties of the tissue. Thus, materials may be characterised by the absorption (μ_a) and the reduced scattering coefficient (μ_s') (Tuchin, 2000).

The LBI technique can be used for decay detection in citrus fruit, since the decaying process implies structural changes in fruit tissue. There exists a physical model for characterising the backscattering profiles from backscattering images, instead of a purely mathematical model based on finding the parameters of mathematical functions (Lorente et al., 2013). The physical approach consists in extracting some optical properties (the absorption and reduced scattering coefficients) of fruits from the Farrell's diffusion theory model, which describes the shape of the profiles (Qin & Lu, 2008). The main objective of this work was study the potential of detecting decay in citrus fruit infected with *Penicillium digitatum* fungus by analysing the two commented optical coefficients of sound and damaged peel of the fruits using using laser-light backscattering imaging. To this end, diode lasers emitting in the visible and NIR range were used to obtain backscattering images of citrus fruit and the Farrell's model was investigated to determine the optical coefficients from the backscattering profiles at each wavelength. Furthermore, for dimensionality reduction purposes, a feature selection method based on mutual information was employed to select the most relevant optical coefficients effective in the classification of orange skin into sound and decaying.

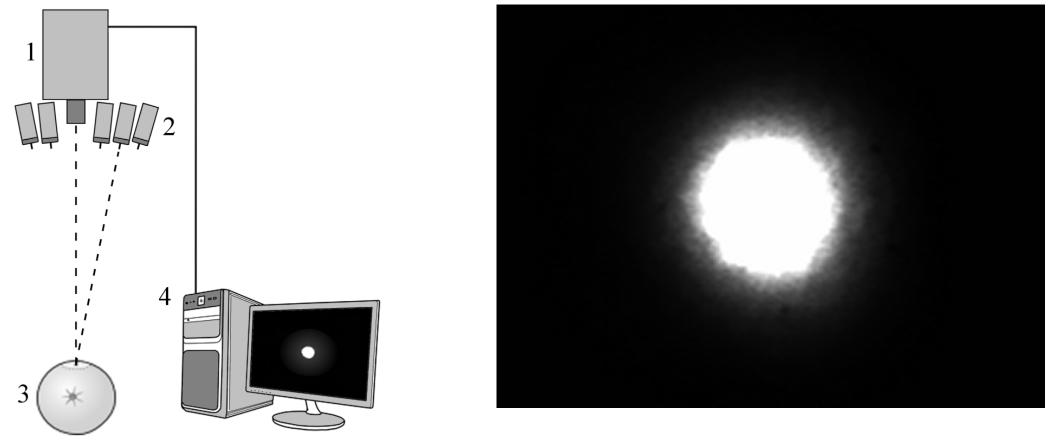
2. Material and methods

2.1 Fruit and fungal inoculation

For the experiments, a total of 40 sweet oranges (*Citrus sinensis* L. Osbeck) cv. Valencia late were superficially wounded on the rind and inoculated with spores of *P. digitatum* fungus. For inoculation, 20 µL of a suspension of spores with a concentration of 1.4×10^6 spores/mL was placed on the equator of each fruit. The fruits were stored in an environment controlled chamber at 22 °C and 55% RH for an enough period of time to get the appearance of decay lesions with a diameter equal or higher than 25 mm on all the fruits. This period was different for each fruit since the development of decay varied between distinct fruits, appearing the first symptoms of decay on some oranges after 4 days‘ storage and the latest after 12 days.

2.2 Imaging system

A laser-light backscattering imaging system was used in this research (Fig. 1a). This system mainly consisted of a monochrome charge-coupled devide (CCD) camera (CV-A50IR, JAI Ltd., Japan) with a zoom lens (12VG1040ASIR-SQ, Tamron Co. Ltd., Japan), five solid-state laser diode modules emitting at different wavelengths (532, 660, 785, 830 and 1060 nm) used alternately as light sources and a personal computer for controlling the camera. Fig. 1b shows an example of a raw backscattering image. After entering into the fruit tissue, the portion of the light backscattered to the fruit surface was recorded by the camera and transferred to the computer by means of a frame grabber board (VRmAVC-1, Vrmagic Holding AG, Germany). Image acquisition took place in a dark room in order to avoid the interference of ambient illumination. The imaging system was configured to acquire 720×576 pixel images with a resolution of 0.133 mm/pixel. The laser diode modules were placed in a fixed position in such a way that all the laser beams were steered to the top of the fruit. Therefore, the incident angle of the light beam was different for each laser, varying from 5º to 15º with respect to the vertical axis. The images were acquired by placing each fruit manually in the imaging system facing the studied parts of the fruit to the camera. Two different surface parts of each orange were analysed: the one presenting decay and the opposite side of the orange, where the skin was sound. Therefore, five images were acquired for each of these two parts of the 40 oranges at the five laser wavelengths, obtaining a total of 400 backscattering images.



a) *b)*

Figure 1: a) Schematic view of the laser-light backscattering system. 1: CCD camera; 2: laser sources; 3: fruit sample; 4: computer. b) A typical raw backscattering image.

2.3 Determination of optical properties

Since backscattering images had radial symmetry with respect to the light incident point (Fig. 1b), they were reduced to one-dimensional profiles through radial averaging (Lu, 2004). The radial intensity of the backscattering profiles was then calculated by averaging all pixels within each circular ring with one pixel size (0.133 mm). Some pre-processing was performed on the profiles to fit the backscattering profiles more accurately, leading to better predictions. The data points within and adjacent to the light incident area, with a greyscale level (0–255) higher than 253, were first removed without losing essential scattering information since these points were saturated. Then, each backscattering profile was normalised by its actual maximum value of light intensity occurred at the closest point to the light incident centre (r_{norm}), thus not requiring the measurement of absolute reflectance intensities. Finally, the radial profiles were described using the Farrell's diffusion theory model by extracting some optical properties (the absorption and reduced scattering coefficients) of fruits. In the Farrell's model (Farrell et al., 1992), the diffuse reflectance at the surface of a semi-infinite turbid material with scattering dominance ($\mu_s' >> \mu_a$) is expressed as a function of distance from the light source and the optical properties of the material (Eq. (1)):

$$R_F(r) = \frac{a'}{4\pi}\left[\frac{1}{\mu_t'}\left(\mu_{eff} + \frac{1}{r_1}\right)\frac{\exp(-\mu_{eff} r_1)}{r_1^2} + \left(\frac{1}{\mu_t'} + \frac{4A}{3\mu_t'}\right)\left(\mu_{eff} + \frac{1}{r_2}\right)\frac{\exp(-\mu_{eff} r_2)}{r_2^2}\right] \quad (1)$$

where r is the distance from the light incident centre (mm); a' is the transport albedo, $a' = \mu_s' / (\mu_a + \mu_s')$; μ_{eff} is the effective attenuation coefficient, $\mu_{eff} = [3\mu_a(\mu_a + \mu_s')]^{1/2}$; μ_t' is the total interaction coefficient, $\mu_t' = \mu_a + \mu_s'$; the variables r_1 and r_2 are given by the equations $r_1 = [(1/\mu_t')^2 + r^2]^{1/2}$ and $r_2 = [((1/\mu_t') + (4A/3\mu_t'))^2 + r^2]^{1/2}$, respectively; A is an internal reflection coefficient dependent on the relative refractive index at the interface, and it can be computed by empirical equations. The refractive index for oranges was assumed constant (1.40). After determining A , the Farrell's model was used to fit the backscattering profiles and to estimate μ_a and μ_s' for each sample at the five laser wavelengths. Because of non-uniformity in the scale of experimental backscattering profiles and Farrell's model, both profiles were normalised at a specific distance with respect to the light incident centre. The normalised Farell's model was then fitted to the normalised experimental profiles using a program based on nonlinear least squares regression analysis. Thus, each backscattering profile was uniquely described by the two optical coefficients. Fig. 2 shows how the Farrell's diffusion theory model fits a backscattering profile. All the algorithms used in this work were implemented using Matlab 7.9 (Mathworks, Inc.).

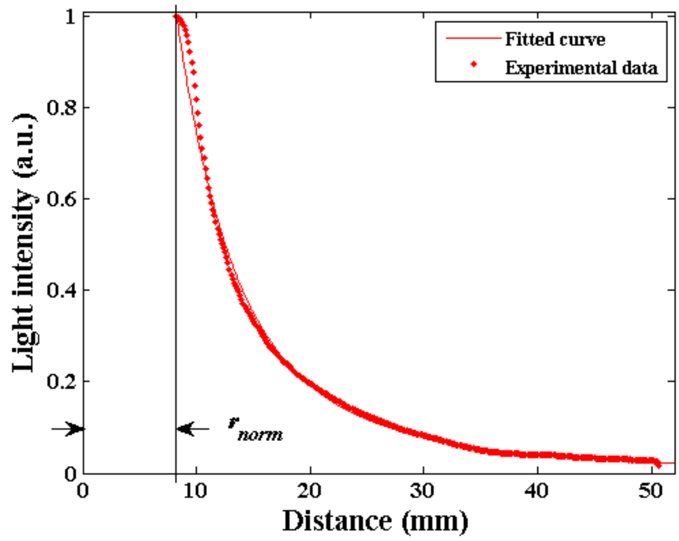


Figure 2: Farrell's diffusion theory model for fitting backscattering profiles.

2.4 Labelled set

Due to the supervised nature of the statistical techniques used in this research, it was necessary to build labelled data set, consisting of 10 features associated to each orange skin sample. Specifically, the absorption and reduced scattering coefficients for each sample extracted from the Farrell's model at the five laser wavelengths. The 80 orange skin samples were assigned to one of the two classes considered in this work: sound and decaying skin. Each sample pattern was therefore composed of 10 features and a class label.

2.5 Feature selection method

Feature vectors could be employed directly as inputs of classification algorithms in order to discriminate between sound and decaying orange skin. However, some of these features may be irrelevant or redundant for the classification task, leading to the overfitting problem and consequently decreasing the classifier performance. For alleviating these problems, a feature selection method based on mutual information (MI; Battiti, 1994) was employed in this work. Methods based on a linear dependence, like correlation analysis, are unable to measure arbitrary dependencies between random variables. On the contrary, the MI function can measure a general dependence between two variables, considering also nonlinear relationships between them. This function measures the reduction in uncertainty about one variable due to the knowledge of the other variable. In this work, the MI function between each input feature and the class variable was evaluated, and the features with the highest MI estimate were selected.

2.6 Classifier

The classifier used in this research was based on linear discriminant analysis (LDA; Fisher, 1936). LDA is a supervised technique based on finding a linear projection of high-dimensional data onto a lower dimensional space where the class separation is maximised. This is achieved by maximising the ratio of the variance between the classes and variance within the classes. The class membership of a sample can be predicted by calculating the distance to the centroid of each class in the transformed space and then assigning the sample to the class with the smallest distance to it.

2.7 Development and validation of the classification models

LDA classification method and feature vectors from the Farrell's model labelled set were used to classify orange skin samples into sound and decaying skin. Furthermore, for dimensionality reduction purposes, the MI feature selection method provided a ranking of features, ordered according to the discriminant relevance of the features to distinguish between both class samples. Then, tests were aimed at selecting a minimum set of features that maximised the classifier performance. To this end, the evolution of classification performance as a function of the number of features was studied. The performance of the classifier was evaluated using the first feature in the ranking, and then successive features were added in an iterative process until all features were employed sequentially. Due to the limited number of samples involved, a cross-validation procedure was used to evaluate and compare the performance of the classification models. In particular, the experiments were performed with 5-fold cross-validation (Hastie et al., 2009), repeating the whole cross-validation process 100 times in order to obtain reliable performance estimation. Therefore, a total of 500 iterations of calibration and validation of each classification model were performed. The validation results were then averaged over all the iterations, thus obtaining a mean confusion matrix created by computing the element-wise mean of the individual confusion matrices for each iteration. In order to assess the classification performance, overall accuracy was computed for each classification model from its associated mean confusion matrix (Fleiss, 1981).

3. Results and discussions

The performance of curve fitting was calculated by averaging the coefficients of determination (R^2) and the root mean squared errors (RMSE) corresponding to the 80 samples of orange skin at each laser wavelength (Table 1). Values related to light intensity were expressed in arbitrary units (a.u.) after normalisation. Farrell‘s model described backscattering profiles with high coefficients of determination ($R^2 \geq 0.982$) and low errors (RMSE ≤ 0.025 a.u.).

Table 1: Average determination coefficients (R^2) and average root mean squared errors (RMSE) from fitting backscattering profiles by Farrell's model for all samples at the five laser wavelengths.

Wavelength (nm)	R^2 (unitless)	RMSE (a.u.)
532	0.9864	0.0144
660	0.9816	0.0250
785	0.9869	0.0205
830	0.9855	0.0218
1060	0.9893	0.0221

Fig. 3 shows the evolution of the overall accuracy of the LDA classifier as a function of the number of ranked features for the MI feature selection method using Farrell's model labelled set. It can be observed that the overall accuracy reached a maximum value of 92.42% using the first eight features ranked with the MI feature selection method.

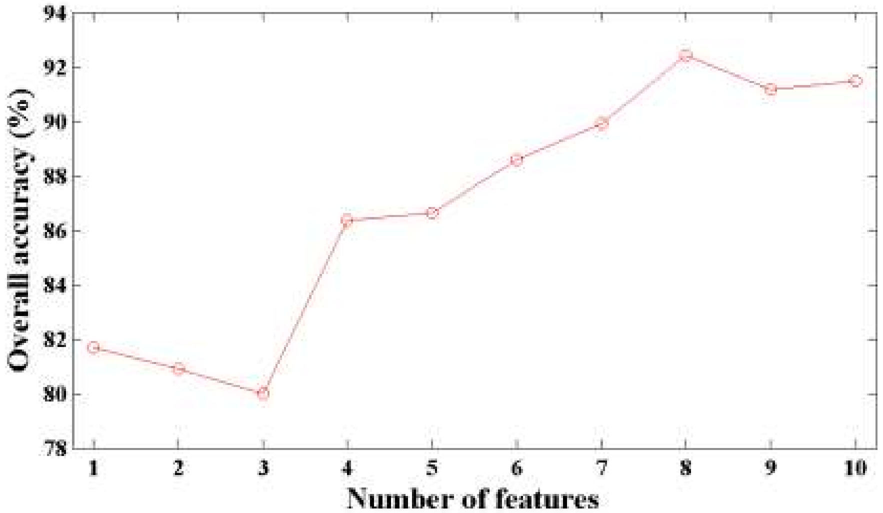


Figure 3: Evolution of the classifier overall accuracy with the number of ranked features for MI selection method using the Farrell's model labelled set.

For a further discussion, Table 2 shows the selected features leading to the best classification performance, as well as the classification results, including overall accuracies, and the associated mean confusion matrices. The resulting set of selected features included the absorption and reduced scattering coefficients at four laser wavelengths, remaining the optical coefficients at 532 nm outside the feature selection. Therefore, it could be assumed that these two features may have little impact on the detection of decay in oranges. From observing the mean confusion matrix, it can be noticed that the number of well-classified decaying samples (94.25%) was higher than the number of well-classified sound samples (90.60%). In practice, the detection of decaying fruit is even more important than the detection of sound fruit for a potential inspection system, since only a few infected fruits can spread the infection to a whole batch, thus causing great economic losses.

To conclude the discussion of the classification results, it can be said that the reduced set of selected features for the Farell's model resulted in good classification results for the decay detection problem, with a percentage of well-classified samples exceeding 90% for both classes despite the similarity between sound and decaying orange skin. Therefore, this work lays the foundation for future implementation of an automatic system based on the laser-light backscattering imaging capable of detecting decay in early stages, which is very important from the agricultural point of view.

Table 2: Selected features and the corresponding classification results for the Farrell's model labelled set.

<table>
<tr><th>Selected features</th><th>Overall accuracy (%)</th><th colspan="3">Mean confusion matrix</th></tr>
<tr><td rowspan="4">μ_a at 660 nm; μ_s' at 660 nm;
μ_a at 785 nm; μ_s' at 785 nm;
μ_a at 830 nm; μ_s' at 830 nm;
μ_a at 1060 nm; μ_s' at 1060 nm</td><td rowspan="4">92.43</td><td></td><td>Sound (%)</td><td>Decay (%)</td></tr>
<tr><td>Sound</td><td>90.60</td><td>5.75</td></tr>
<tr><td>Decay</td><td>9.40</td><td>94.25</td></tr>
<tr><td></td><td></td><td></td></tr>
</table>

4. Conclusions

This work proved that laser-light backscattering imaging technique was useful for detecting early decay symptoms in citrus fruit caused by *P. digitatum*. Backscattering images of oranges at five laser wavelengths in the visible and NIR range were used for this detection. The Farrell's diffusion theory model described backscattering profiles accurately at the five laser wavelengths using the absorption and reduced scattering coefficients, with average R^2 values higher than or equal to 0.982.

In the evaluation of the MI feature selection method, the optimal set of features leading to the best performance in the classification of sound and decaying skin included the two optical coefficients at four laser wavelengths, remaining the optical coefficients at 532 nm outside the selected features, and reached an overall accuracy of around 92.4% with an even higher number of well-classified decaying samples (94.3%). In conclusion, the optimal set of features provided good classification results, with a percentage of well-classified samples above 90% for both classes despite the similarity between sound and decaying skin. The next step will be the development of a prototype for in-line real-world tests.

5. Acknowledgements

This work has been partially funded by the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria de España (INIA) through research project RTA2012-00062-C04-01 with the support of European FEDER funds. Delia Lorente thanks INIA for the support trough grant FPI-INIA number 42.

6. References

Battiti, R. (1994). Using mutual information for selecting features in supervised neural net learning. *IEEE Transactions on Neural networks, 5,* 537–550.

Blasco, J., Aleixos, N., & Moltó, E. (2007). Computer vision detection of peel defects in citrus by means of a region oriented segmentation algorithm. *Journal of Food Engineering, 81,* 535–543.

Farrell, T. J., Patterson, M. S., & Wilson, B. (1992). A diffusion-theory model of spatially resolved steady-state diffuse reflectance for the noninvasive determination of tissue optical-properties in vivo. *Medical Physics, 19,* 879–888.

Fisher, R. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics, 7,* 179-188.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions*. (2nd ed.). Wiley-Interscience.

Gómez-Sanchis, J., Martín-Guerrero, J. D., Soria-Olivas, E., Martínez-Sober, M., Magdalena-Benedito, R., & Blasco, J. (2012). Detecting rottenness caused by *Penicillium* in citrus fruits using machine learning techniques. *Expert Systems with Applications, 39,* 780-785.

Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. (2nd ed.). New York: Springer-Verlag.

Kim, D. G., Burks, T. F., Qin, J., & Bulanon, D. M. (2009). Classification of grapefruit peel diseases using color texture feature analysis. *International Journal of Agricultural and Biological Engineering, 2,* 41–50.

Kurita, M., Kondo, N., Shimizu, H., Ling, P., Falzea, P. D., Shiigi, T., Ninomiya, K., Nishizu, T., & Yamamoto, K. (2009). A double image acquisition system with visible and UV LEDs for citrus fruit. *Journal of Robotics and Mechatronics, 21,* 533–540.

Lorente, D., Zude, M., Regen, C., Palou, L., Gómez-Sanchis, J., & Blasco, J. (2013). Early decay detection in citrus fruit using laser-light backscattering imaging. *Postharvest Biology and Technology, 86,* 424-430.

Lu, R. (2004). Multispectral imaging for predicting firmness and soluble solids content of apple fruit. *Postharvest Biology and Technology, 31,* 147–157.

Qin, J., & Lu, R. (2008). Measurement of the optical properties of fruits and vegetables using spatially resolved hyperspectral diffuse reflectance imaging technique. *Postharvest Biology and Technology, 49,* 355–365.

Qing, Z., Ji, B., & Zude, M. (2007). Predicting soluble solid content and firmness in apple fruit by means of laser light backscattering image analysis. *Journal of Food Engineering, 82,* 58–67.

Romano, G., Baranyai, L., Gottschalk, K., & Zude, M. (2008). An approach for monitoring the moisture content changes of drying banana slices with laser light backscattering imaging. *Food and Bioprocess Technology, 1*, 410–414.

Tuchin, V. (2000). *Tissue Optics: Ligh Scattering Methods and Instruments for Medical Diagnosis*. Bellingham (Washington): SPIE Press.