

# Fluorescence hyper-spectral imaging to detecting faecal contamination on fresh tomatoes

Roberto Romaniello, Giorgio Peri, Alessandro Leone

Department of the Science of Agriculture, Food and Environment, University of Foggia, Italy

## Abstract

Faecal contamination of fresh fruits represents a severe danger for human health. Thus some techniques based on microbiological testing were developed to individuate faecal contaminants but those tests do not results efficient because their non-applicability on overall vegetable unity. In this work a methodology based on hyper-spectral fluorescence imaging was developed and tested to detecting faecal contamination on fresh tomatoes. Two image-processing methods were performed to maximise the contrast between the faecal contaminant and tomatoes skin: principal component analysis and band image ratio (BRI). The BRI method allows classifying correctly 70% of contaminated area, with no false-positives in all examined cases. Thus, the developed methodology can be employed for a fast and effective detection of faecal contamination on fresh tomatoes.

## Introduction

The contamination of vegetables with pathogens can be the result of their exposure to faecal material during the phases of cultivation and/or harvesting. In particular, these contaminated vegetables become carriers of bacteria and protozoa that are harmful to human health. At this regard, two protozoa (*Giardia* and *Cryptosporidium* spp.) have been responsible for zoonotic infections in humans worldwide (Giangaspero *et al.*, 2007). Since contamination with faecal material cannot be completely removed from fruit and vegetables intended for fresh consumption using current methods of washing/sanitizing, the most effective way to minimise the

risks affecting food safety is to identify and remove from the production flow contaminated products before they are placed on the market (Liu *et al.*, 2007). Nowadays, in order to fulfil the above purpose, molecular analysis techniques are available, but these techniques are expensive and not rapidly applicable. Furthermore, a fundamental limitation is given by the fact that these techniques, being destructive, are not applicable on-line and, consequently, do not allow the control of an entire production lot. The approaches generally used for non-destructive inspection of fruit and vegetables are based on punctual spectroscopic techniques, applied off-line or on conventional imaging techniques (multispectral and hyperspectral imaging), applied on-line. In particular, the hyperspectral imaging is an emerging technique that integrates the conventional imaging and punctual spectroscopy, providing detailed spatial and spectral information of the object under investigation in the form of hyper-spectral image. From the analysis of such hyperspectral image, consisting of hundreds of images acquired at contiguous wavelength in the field of the visible and near infrared spectrum, complex models can be developed in order to identify, quantify and classify small details of the object of investigation, as well as physical and chemical internal properties of the same object (Gowen *et al.*, 2007; Moscetti *et al.*, 2014; Moscetti *et al.*, 2015). It is clear that the added value of the hyper-spectral imaging is represented by the possibility to be used on-line, allowing the accurate control of an entire production lot. Moreover, as well as all the spectroscopic techniques, the hyperspectral imaging can be applied in the food industry to images acquired not only in reflectance and transmittance, but also in fluorescence (Park *et al.*, 2007; Huang *et al.*, 2014; Lee *et al.*, 2014; Wiederoder *et al.*, 2013). Biological matrices, including plants and faecal material of animal origin, emit fluorescence in the visible and near infrared wavelengths when excited by ultraviolet radiation (Kim *et al.*, 2002). Studies conducted by Vargas *et al.* (2004) shown that fluorescence-based techniques are more sensitive than those based on reflectance/transmittance. Thus, hyperspectral fluorescence imaging could be an effective technique to detect the faecal contamination of vegetable products (Vargas *et al.*, 2005). In the specific case of fruit and vegetables intended for fresh consumption, the application of this technique would be very useful to grant the food security, as it would be implemented in the production flow of critical control points aimed at identifying and removing products contaminated with traces of faecal material, invisible to human eye, before placing them on the market. The objective of present work is to develop a fluorescence hyper-spectral imaging system in order to investigate on whose wavelength are useful to detect faecal contamination of fresh tomatoes and to develop an algorithm able to automatically detect the contaminated area, on the basis of the acquired fluorescence spectrum data.

Correspondence: Roberto Romaniello, Department of the Science of Agriculture, Food and Environment, University of Foggia, via Napoli 25, 71122 Foggia, Italy.  
Tel.: +39.0881.589.120.  
E-mail: roberto.romaniello@unifg.it

Key words: Hyperspectral imaging; faecal contamination; tomatoes.

Received for publication: 6 July 2015.

Accepted for publication: 24 January 2016.

©Copyright R. Romaniello *et al.*, 2016  
Licensee PAGEPress, Italy  
Journal of Agricultural Engineering 2016; XLVII:491  
doi:10.4081/jae.2016.491

This article is distributed under the terms of the Creative Commons Attribution Noncommercial License (by-nc 4.0) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## Materials and methods

### Tomatoes and faecal contamination treatments

One hundred and fifty *Camodium* red tomatoes were purchased from a local supermarket and immediately transferred in laboratory for

the analyses. The tomatoes were washed and dried to avoid the presence of residual soil. Hyper-spectral images of all 150-tomato samples were first acquired before any treatments in order to acquire their fluorescence spectra. Bovine faeces constituted the faecal contaminant. It was collected in a local breeding cattle farm. Samples of diluted faecal contaminant were obtained, diluting by weight to 1:500 bovine faeces with deionised water. Dry faecal matter concentration for 1:500 dilution was 20 µg/mL as determined by drying samples to constant weight in a 90°C oven. A volume of 0.3 mL of the solution was placed on each tomato sample. These tomato samples were air-dried for several hours until faecal spots were completely dried, and then scanned again. It was observed that the contaminated areas became transparent and not visible to the human eyes, after the tomatoes drying. Therefore, a scanning in visible spectrum of all artificially contaminated samples were acquired immediately after the artificial contamination and prior to taking hyper-spectral fluorescence images in order to identify the effective contaminated area on each sample (binary images called A-images).

### Hyper-spectral imaging system

The hyper-spectral imaging system (DV optic s.r.l., Padova, Italy) is a line-by-line scanning system assembled to work in the visible-infrared (VIS) to near-infrared (NIR) region of the spectrum. The main components of the system are a camera, a spectrograph, and a lens along with a sample transport mechanism and lighting source. The camera is a progressive-scan type (AVT F100-B; Allied Vision Technologies GmbH, Stadtroda, Germany), and its sensor is a 16-bit CCD (Kodak KSI 1020, DE) having a spatial resolution of 1000×1000 pixels. The spectrograph (ImSpector V10; Specim Ltd., Oulu, Finland) is based on prism-grating-prism optics, and is coupled to an  $f$  1.4 C-mount lens. The spectrograph (ImSpector V10; Specim Ltd.) disperses incoming radiation along the scan line into spectral information for each spatial location. The spectral range in the VIS/NIR region of the spectrum is 400-1000 nm with a spectral resolution of 5 nm. The system was modified in order to be equipped with two independent illumination sources for fluorescence and reflectance measurements respectively. For fluorescence measurements a UV-A fluorescent lamp assembly was developed and arranged toward the line of the instantaneous field of view at 30° forward to provide near-uniform excitation energy. The UV-A illumination system has been made to generate a beam of ultraviolet radiation centred on 365 nm, having a width of 500 mm and a thickness of about 50 mm. The radiation source, for fluorescence measurements, consisted of 3 leds with a rated power of 230 mW arranged in a line, spaced 23 mm (Figure 1). Each LED diode was applied to a cylindrical collimating lens for focusing linear, while short-pass filters were placed in front of the lamp housing to prevent transmittance of radiations greater than approximately 400 nm, eliminating potential spectral contamination by pseudo-fluorescence. For reflectance measurement, a halogen lamp (EKE 21V, 150W) was used. The sample transport mechanism moves the sample through the line of the instantaneous field of view in a transverse direction. It consisted in a black metal plane moved by a step motor, synchronised with the digital camera, in order to acquire the spectral reflectance of subsequent lines, according to the camera integration time. The movement of the sample transport mechanism was controlled by a custom software (DV s.r.l.). The software also converts the individual raw data file downloaded from the camera into 16 bit hyper-spectral fluorescence image and the processed image is saved as a standard 8-bit bitmap ASCII file for further analysis.

### Image processing and performance indexes

The hyper-spectral fluorescence images were processed using algorithms coded in MATLAB® R2013a (Mathworks, Inc., Natick, MA, USA).

Two image-processing methods were tested to improve the contrast between the faecal-contaminated areas and uncontaminated areas in the tomato sample images: principal component analysis (PCA) and band ratio images (BRI).

The Hotelling's transformation (Kim *et al.*, 2002) is a statistical method that permits to obtain, by starting from the 121 component images and performing a linear transformation, a new set of 121 images called principal component images. The principal component images are related to each other and they are arranged in descending order on the basis of the variance of the grey levels of the pixels of each image. Therefore, the first principal component image has the highest variance of the pixels intensity thus the one with the maximum contrast between the image regions corresponding to faecal material and tomatoes surface.

The BRI method is an empirical technique based on the calculation of an image-ratio of two operand images (Vargas *et al.*, 2005). The image-ratio is obtained by performing a punctual division between pixels of the two-operand images. The image obtained can have a greater variance of the pixel intensities than the two operand-images. The above method provides to calculate all possible ratio-images attainable from the 121 component images (14,520 ratio-images) and automatically selects the ratio-image which has the maximum variance between the pixels' grey level, *i.e.* the one having the maximum contrast between the image regions corresponding to the faecal material and the uncontaminated tomato skin.

These methods allow obtaining grey-level images having high contrast between the faecal-contaminated and uncontaminated areas. These images were then converted in logic format using the Otsu's method (Otsu, 1979), obtaining the binary images of the estimated faecal-contaminated areas (called B-images). Two performance indices were selected: pixel fraction of the uncontaminated area correctly classified in the B-images ( $P_s$ ); pixel fraction of the faecal-contaminated area correctly classified in the B-images ( $P_o$ ). The indices  $P_o$  and  $P_s$  were calculated using the following expressions:

$$P_o = 1 - \frac{n_{fn}}{n_o} \quad P_s = 1 - \frac{n_{fp}}{n_s} \quad (1)$$

where  $n_{fn}$  represents the number of false negatives,  $n_o$  the effective number of the pixel belonging to the faecal-contaminated area,  $n_{fp}$  represents the number of false positives and  $n_s$  the effective number of pixel belonging to the uncontaminated area.

The  $n_{fn}$ ,  $n_o$ ,  $n_{fp}$  and  $n_s$  values were determined considering the binary image of the effective contaminated area (A-image), captured immedi-

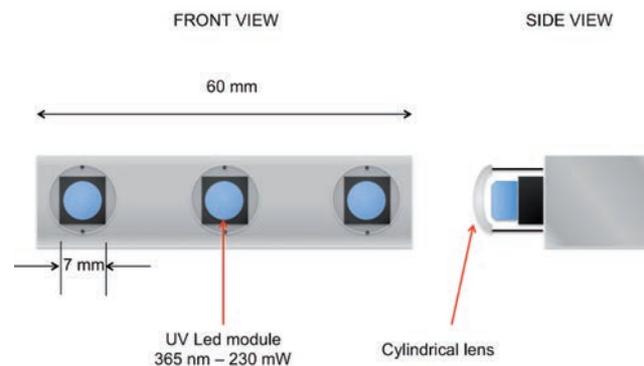


Figure 1. Scheme of UV-A fluorescent lamp assembly.

ately after the contamination when the spot of contaminant was still visible to human eye, and the B-image mentioned above. Two logical operators were applied to these images obtaining the  $n_{fn}$  and  $n_{fp}$  values. In Figure 2 an explicative example is reported: Figure 2A represents the A-image (effective contamination area); Figure 2B represents the segmented image (starting from the enhanced image computed by PCA or BRI methods); Figure 2C represents the binary image C resulted applying the logic operator  $XOR$  to images A and B ( $C = A XOR B$ ); Figure 2D represents the binary image D resulted applying the logic operator  $AND$  to images A and C ( $D = A AND C$ ); finally, Figure 2E represents the binary image E resulted applying the logic operator  $AND$  to images C and D ( $E = C AND D$ ). The sum of the pixels having value 1 (white pixels) in the image A represents the  $n_o$  value, whereas the sum of pixels with value 0 (black pixels) gives the  $n_s$  value. Finally, the sum of the pixels having value 1 in images D and E represents the  $n_{fn}$  and  $n_{fp}$  values respectively.

The elaboration time was measured both for PCA and BRI method. The time necessary to calculate the enhanced images using PCA and BRI, starting from the influents wavelengths (PCA) and from the best two wavelengths (BRI) respectively was considered.

## Results and discussion

The average fluorescence spectra of faecal-contaminated and uncontaminated areas are reported in Figure 3. The faecal-contaminated area shows two fluorescence peaks centred to 705 and 765 nm, while the uncontaminated areas exhibit a single fluorescence peak centred to 765 nm.

The average fluorescence intensity of the contaminated area results higher than those of the uncontaminated area, for all wavelengths considered. This situation does not allow the differentiation of the two areas by simply selecting single wavelengths, thus a combination of wavelengths is necessary, such as those performed applying PCA and BRI methods. The enhanced images (B-images) resulted by applying the PCA and BRI method are reported in Figures 4 and 5, respectively. The BRI enhanced image was obtained by dividing the image component 705 by image component 815, both selected by the iterative algorithm. Differences in intensity distribution are visible comparing these two images. The enhanced image obtained applying the PCA appears noisy in the areas around the contamination spot on the contrary, the enhanced image obtained applying the BRI method has not noise in the background, but appears with holes in the contamination area. These differences are well visible in the relative segmented images (Figures 4 and 5), obtained using the Otsu's method. In Figure 4 (PCA method) a lot of false positives are visible, represented to the pixel having value 1 (white pixels) outside the contamination area; on the contrary, the Figure 5 shows some false negative, which correspond to the pixels having value 0 (black pixels) in the contamination area. To better visualise the effect of the two methods, the segmented images were processed to obtain the pseudo-colour images relative to PCA (Figure 6) and BRI (Figure 7) method. Table 1 lists the average values of the performance indices  $P_O$  and  $P_S$  for the B-images obtained by the application of PCA and BRI methods.

The average values of the pixel fraction of the faecal-contaminated areas correctly classify were 0.8003 and 0.6998 for PCA and BRI methods respectively. The average values of the pixel fraction of the uncontaminated areas correctly classify were 0.0533 and 1 for PCA and BRI methods respectively.

The false negative were on average 19.97% and 30.02% for the PCA and BRI methods respectively, and the false positive were on average 97.47% and 0% for the PCA and BRI methods respectively.

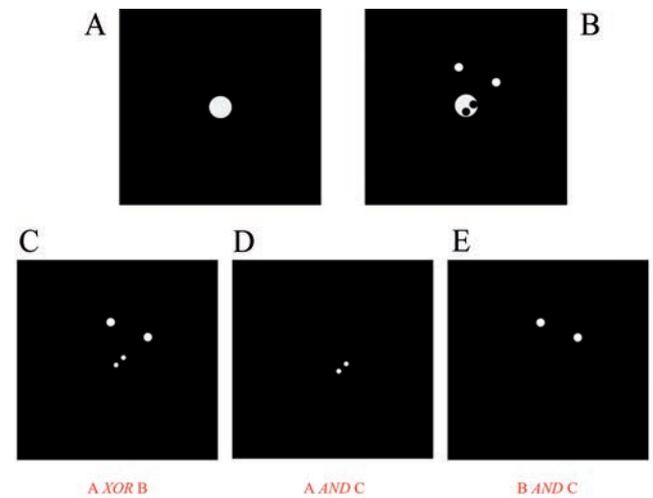


Figure 2. Binary image and logical operation used to determine the  $n_{fn}$ ,  $n_o$ ,  $n_{fp}$  and  $n_s$  values.

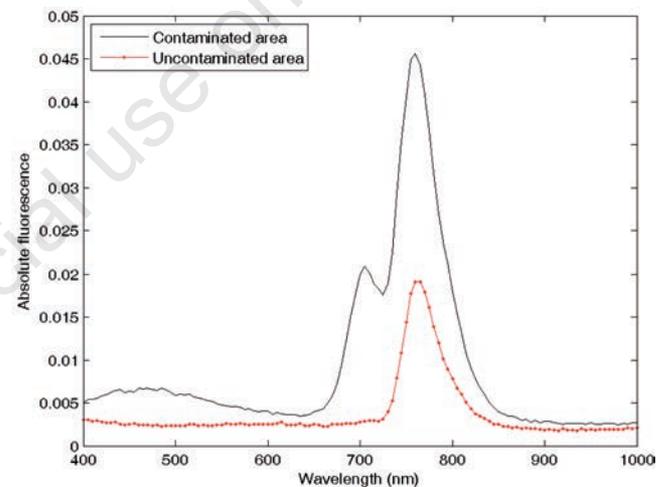


Figure 3. Average fluorescence spectra of faecal-contaminated and uncontaminated areas of the 150 tomato samples considered.

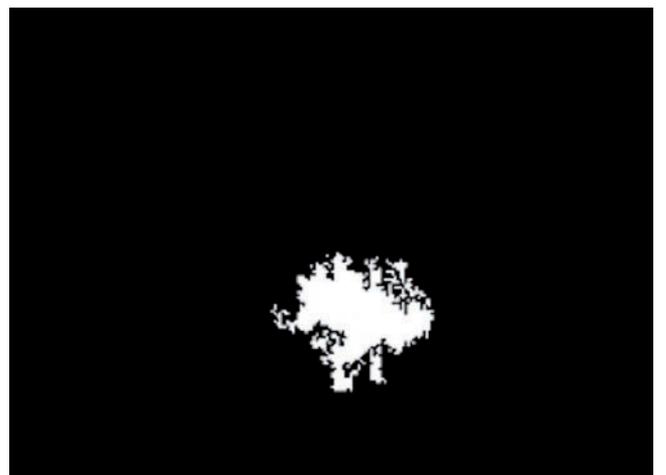


Figure 4. Binary image obtained by segmentation of B-image relative to principal component analysis method.

When the contrast between the faecal-contaminated areas and uncontaminated areas was improved using the PCA method, about 80% of the faecal-contaminated areas was correctly classified. When the contrast between the faecal-contaminated areas and the uncontaminated areas was improved using BRI method about 70% of the faecal-contaminated areas was correctly classify. However, the BRI method exhibits more useful results due to absence on average of false positives. The large amount of false positive in PCA methods derives, probably, from the high over enhancement obtained by combining linearly the 121 wavelengths. At this stage, the high (contaminated area) and the low pixel intensities (background area) are amplified at the same time. On the other hand, the BRI method is based on the ratio of component images. The ratio between pixels of two different images tends to emphasise the pixel of the dividend image having higher intensity than the correspondent pixel in the dividend image, and to cut off the pixels having lower intensities in the dividend image than the correspondent pixels in the divisor image. This produces images well contrasted, without noise in the background. Finally, the elaboration times for the two methods were determined. PCA and BRI methods required 12 seconds and 6.8 milliseconds, respectively, to generate the enhanced image starting from the principal component images selected by PCA and from the two best images selected for BRI method. It is clear that the PCA method is too slow for an on-line application. On the contrary, the BRI method allows investigating a sample in less than 7 milliseconds and it can be implemented in an on-line application for a fast detection of faecal contamination on *Camodium* tomatoes.

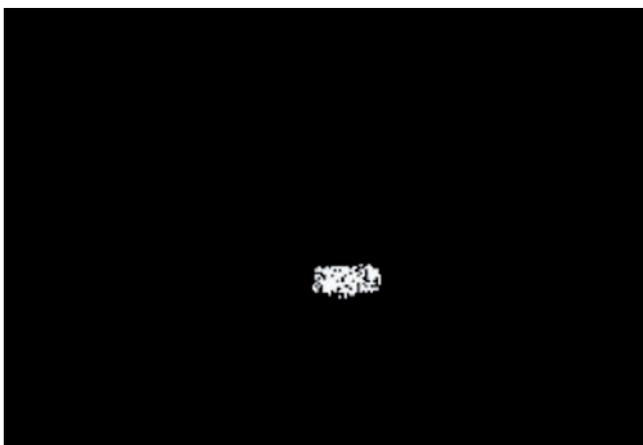


Figure 5. Binary image obtained by segmentation of B-image obtained by band image ratio method.

## Conclusions

In this study a hyper-spectral fluorescence imaging system has been developed to evaluate the potential for detection of faecal contamination on red tomatoes. The results shown in this investigation are further manifestation that fluorescence imaging techniques are very sensitive tools and indicate that hyper-spectral fluorescence imaging systems can be used to detect faecal contamination on red tomatoes that

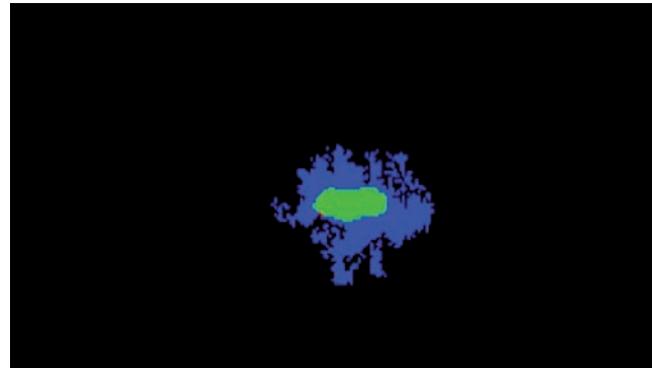


Figure 6. Principal component analysis method - pseudo-colour image: in green colour the pixels correctly classified; in red colour the false positives; in blue colour the false-positives pixels.

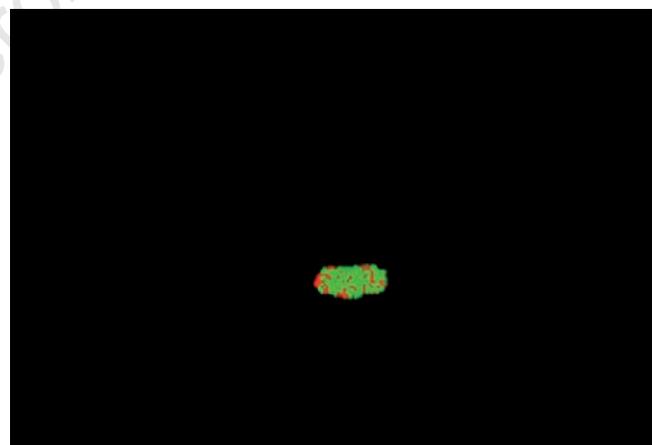


Figure 7. Band image ratio method - pseudo-colour image: in green colour the pixels correctly classified; in red colour the false-negative pixels.

Table 1. Average values of  $P_0$  and  $P_s$  obtained from the principal component analysis method and the band image ratio method, and statistical indices.

Method Performance index	PCA		BRI	
	$P_0$	$P_s$	$P_0$	$P_s$
Average	0.8003	0.0553	0.6998	1.0000
Standard deviation	0.2193	0.0453	0.0823	0.0000
Minimum	0.0000	0.0091	0.6662	1.0000
Maximum	0.9917	0.2665	0.9535	1.0000
Variation coefficient	0.2740	0.8197	0.1176	0.0000

PCA, principal component analysis; BRI, band image ratio.

is not visible to the human eyes. The algorithm, developed for automatic detection of contaminated area works better by using BRI method than PCA. This result indicates that by the identification of two wavelengths, to be used to acquire the fluorescence images, it is possible to get an image-ratio, which expresses the greatest possible variance between the pixels of the contaminated area and the pixels of the uncontaminated area. The application of a segmentation algorithm to such image-ratio allows identifying the contamination area without any risk of generating false positives.

Using the fluorescence hyper-spectral imaging system developed, other vegetable matrices will be analysed in order to identify the specific wavelengths to be used, for each matrix, to acquire the fluorescence images useful to obtain the optimal image-ratio for the automatic detection of faecal contamination. In this way, it is possible to develop specific multispectral systems implemented as part of a faecal contamination recognition point in processing plants, in order to quickly identify contaminated products before commercialisation. The identification of the two best wavelengths (705 and 815 nm) that allow to obtain an image useful to discriminate the faecal contaminant from tomatoes surface is the starting step to develop an imaging system based on a sensor able to acquire only these two wavelengths; a specific software for the acquisition of hyperspectral images and their further processing can be developed, in order to construct an efficient environment to be used on-line, at reasonable costs.

## References

- Giangaspero A., Berrilli F., Brandonisio O. 2007. Giardia and Cryptosporidium and public health: the epidemiological scenario from the Italian perspective. *Parasitol. Res.* 101:1169-82.
- Gowen A.A., O'Donnell C.P., Cullen P.J., Downey G., Frias J.M. 2007. Hyperspectral imaging - an emerging process analytical tool for food quality and safety control. *Trend Food Sci. Technol.* 18:590-8.
- Huang M., Wang Q., Zhang M., Zhu Q. 2014. Prediction of color and moisture content for vegetable soybean during drying using hyperspectral imaging technology. *J. Food Engine.* 128:24-30.
- Kim M.S., Lefcourt A.M., Chao K., Chen Y.R., Kim I., Chan, D.E. 2002. Multispectral detection of faecal contamination on apples based on hyperspectral imagery: Part I. Application of visible and near-infrared reflectance imaging. *Trans. ASAE.* 45:2027-37.
- Lee W., Kim M., Lee H., Delwiche S., Bae H., Kim D., Cho B. 2014. Hyperspectral near-infrared imaging for the detection of physical damages of pear. *J. Food Engine.* 130:1-7.
- Liu Y., Chen Y.R., Kim M.S., Chan D.E., Lefcourt A.M. 2007. Development of simple algorithms for the detection of faecal contaminants on apples from visible/near infrared hyperspectral reflectance imaging. *J. Food Engine.* 81:412-8.
- Moscetti R., Monarca D., Cecchini M., Massantini R., Haff R.P., Saranwang S. 2014. NIR spectroscopy is suitable to detect insect infested chestnuts. *Acta Hort.* 1042:153-9.
- Moscetti R., Saeys W., Keresztes J.C., Goodarzi, M., Cecchini M., Monarca D., Massantini, R. 2015. Hazelnut quality sorting using high dynamic range short-wave infrared hyperspectral imaging. *Food Bioproc. Technol.* 8:1593-604.
- Otsu N. 1979. A threshold selection method from gray-level histogram. *IEEE Trans Syst Man Cyber* 9:62-6.
- Park B., Windham W.R., Lawrence K.C., Smith D.P. 2007. Contaminant classification of poultry hyperspectral imagery using a spectral angle mapper algorithm. *Biosyst. Engine.* 96:323-33.
- Vargas A.M., Kim M.S., Tao Y., Lefcourt A., Chen Y.R. 2004. Safety inspection of Cantaloupes and strawberries using multispectral fluorescence imaging techniques. Written for presentation at the 2004 ASAE Annual International Meeting Sponsored by ASAE Fairmont Chateau Laurier, July 31-August 4, Ottawa, Ontario, Canada.
- Vargas A.M., Kim M.S., Tao Y., Lefcourt A.M., Chen Y.R., Luo Y., Song Y., Buchanan R. 2005. Detection of fecal contamination on Cantaloupes using hyperspectral fluorescence imagery. *J. Food Sci.* 70:471-6.
- Wiederoder M., Liu N., Lefcourt A., Kim M., Martin Lo Y. 2013. Use of a portable hyperspectral imaging system for monitoring the efficacy of sanitation procedures in produce processing plants. *J. Food Engine.* 117:217-26.