

Peer Reviewed Paper **openaccess** [Special Issue on Chemometrics in Hyperspectral Imaging](#)

Semi-supervised learning of hyperspectral image segmentation applied to vine tomatoes and table grapes

Jeroen van Roy,^a Niels Wouters,^b Bart De Ketelaere^c and Wouter Saeys^{d,*}

^aKU Leuven, Department of Biosystems, MeBioS, Kasteelpark Arenberg 30, B-3001 Leuven, Belgium.
ORCID: <https://orcid.org/0000-0003-4123-9395>

^bKU Leuven, Department of Biosystems, MeBioS, Kasteelpark Arenberg 30, B-3001 Leuven, Belgium.
ORCID: <https://orcid.org/0000-0003-1563-3583>

^cKU Leuven, Department of Biosystems, MeBioS, Kasteelpark Arenberg 30, B-3001 Leuven, Belgium.
ORCID: <https://orcid.org/0000-0002-5140-1643>

^dKU Leuven, Department of Biosystems, MeBioS, Kasteelpark Arenberg 30, B-3001 Leuven, Belgium. E-mail: wouter.saeys@kuleuven.be.
ORCID: <https://orcid.org/0000-0002-5849-4301>

Nowadays, quality inspection of fruit and vegetables is typically accomplished through visual inspection. Automation of this inspection is desirable to make it more objective. For this, hyperspectral imaging has been identified as a promising technique. When the field of view includes multiple objects, hypercubes should be segmented to assign individual pixels to different objects. Unsupervised and supervised methods have been proposed. While the latter are labour intensive as they require masking of the training images, the former are too computationally intensive for in-line use and may provide different results for different hypercubes. Therefore, a semi-supervised method is proposed to train a computationally efficient segmentation algorithm with minimal human interaction. As a first step, an unsupervised classification model is used to cluster spectra in similar groups. In the second step, a pixel selection algorithm applied to the output of the unsupervised classification is used to build a supervised model which is fast enough for in-line use. To evaluate this approach, it is applied to hypercubes of vine tomatoes and table grapes. After first derivative spectral preprocessing to remove intensity variation due to curvature and gloss effects, the unsupervised models segmented 86.11% of the vine tomato images correctly. Considering overall accuracy, sensitivity, specificity and time needed to segment one hypercube, partial least squares discriminant analysis (PLS-DA) was found to be the best choice for in-line use, when using one training image. By adding a second image, the segmentation results improved considerably, yielding an overall accuracy of 96.95% for segmentation of vine tomatoes and 98.52% for segmentation of table grapes, demonstrating the added value of the learning phase in the algorithm.

Keywords: hyperspectral imaging, segmentation, clustering, classification, machine learning, vine tomatoes, table grapes

Introduction

In spectral hypercubes of agrofood products multiple classes are often present. To acquire quality properties

from these classes, a calibration method to extract the quality properties of interest has to be built.¹ When

Correspondence

Wouter Saeys (wouter.saeys@kuleuven.be)

Received: 14 September 2017

Revised: 19 December 2017

Accepted: 31 January 2018

Publication: 6 March 2018

doi: 10.1255/jsi.2018.a7

ISSN: 2040-4565

Citation

J. van Roy et al., "Semi-supervised learning of hyperspectral image segmentation applied to vine tomatoes and table grapes", *J. Spectral Imaging* 7, a7 (2018). <https://doi.org/10.1255/jsi.2018.a7>

© 2018 The Authors

This licence permits you to use, share, copy and redistribute the paper in any medium or any format provided that a full citation to the original paper in this journal is given, the use is not for commercial purposes and the paper is not changed in any way.



multiple objects are present in an image, the pixels which have similar properties should be allocated to the corresponding objects,² to enable the building of prediction models.

The segmentation of hyperspectral images can either be done manually or automatically.³ As manual segmentation is time-consuming, tedious and subjective, it is not suitable for industrial applications. Therefore, automatic segmentation is preferred. Many techniques have been proposed for this, which can be divided into three groups:⁴ segmentation in the spatial domain, segmentation in the spectral domain and segmentation combining spatial and spectral information. Segmentation in the spatial domain involves the clustering of spatially connected pixels of similar intensity by algorithms like region growing⁵ and watershed transformation.⁴ These algorithms are either applied to images at a single wavelength with maximal contrast between the different objects to be segmented, or virtual images which are calculated by combining images at different wavelengths to obtain maximal contrast.

Segmentation in the spectral domain ignores the spatial position of the pixel spectra in the image and only looks at the similarity of the pixel spectra. For this purpose, the information from multiple wavelengths is combined to obtain a good classification. For example, ElMasry *et al.*⁶ subtracted images at different wavelengths to classify meat, fat and background. Baranowski *et al.*⁷ used different supervised classification methods, such as support vector machines (SVM), functional trees, nearest-neighbour classifiers and regression methods, to distinguish bruised apple tissue from sound tissue and found that logistic regression gave the best classification results with a correct classification rate of 98.8%. Lü *et al.*⁸ applied SVM on hyperspectral data from kiwi-fruit to detect bruises with a misclassification rate of 12.5%. Partial least squares discriminant analysis (PLS-DA), a supervised classification method based on partial least squares regression (PLSR),⁹ was used to classify undamaged, mechanically damaged and microbiological diseased mushrooms with a correct classification rate of more than 95%¹⁰ and to classify hazelnuts into four quality classes with more than 90% accuracy.¹¹ Nguyen Do Trong *et al.*¹² applied PLS-DA to discriminate between cooked and raw pixels in potatoes with a classification accuracy of 95.82% for raw pixels and 97.91% for fully cooked pixels. Keresztes *et al.*¹³ detected bruises on apples using PLS-DA with an accuracy of 96.25%.

An example of a segmentation algorithm combining both spectral and spatial information was described by Lee *et al.*¹⁴ They first applied PLS-DA to distinguish the spectra from cracked tomato tissue, sound tomato tissue and zones with specular reflection. Afterwards, the images were processed by conducting image filling, followed by a labelling of connected classes. Only the largest classes were preserved. Then, they applied linear discriminant analysis (LDA) and SVM on two morphological parameters quantifying the shape of the stem/scar region, namely roundness and minimum-maximum distance from the centre of the class, which were derived from the binary image. This resulted in a correct classification rate of 94.6% for LDA and 96.4% for SVMs, respectively.¹⁴ Zhu *et al.*¹⁵ used spectral characteristics in combination with textural variables, extracted by using a grey-level co-occurrence matrix, of fresh fish fillets and fish fillets that were frozen-thawed as input for a least-squares SVM, resulting in a correct classification rate of 97.22%. Portalés and Ribes-Gómez¹⁶ developed a system based on camera vision, which was sensitive to light in the near infrared (1000–1100 nm). It was able to distinguish grapes from stalks and leaves in harvest batches. In a first step, they distinguished grapes from bright spots, stems and leaves using thresholds. In a second step, they analysed the shape of parts of an image with comparable properties to distinguish stalks from other, unidentified objects. They achieved minimal root mean squared error (RMSE) values of 0.34% and 0.08% for classification of pixels as stalk or unidentified object, respectively.

As the building of multivariate calibration models for segmentation in the spectral domain is tedious and requires specific expertise, the use of a learning method is very appealing. In machine learning, three different learning methods are distinguished: unsupervised learning, supervised learning and semi-supervised learning. Unsupervised learning does not require a training set of labelled spectra as these methods group the spectra based on a similarity metric.¹⁷ Examples of unsupervised learning techniques are *k*-means clustering and hierarchical clustering. The disadvantage of these techniques is that they are computationally very intensive and consequently still require too much computation time to be suitable for inline use. These techniques also search for groupings in the dataset, but the groups found are not necessarily in accordance with the classes of interest. In supervised learning, the unlabelled data are given to an expert who analyses them and who divides

the data in the different classes present (labelling). Then, the labelled data are sent to a learner. This learner then builds a classifier, based on the labelled data, which is capable of classifying new data. Examples of supervised learning methods, such as SVM and PLS-DA, were presented in the previous paragraph. The execution of such methods is much faster than for the unsupervised methods, but a good training set is required to build the segmentation models. To acquire such a training set, a series of training images have to be segmented manually. As misclassifications in the training set would have a negative impact on the segmentation model trained on it, this manual segmentation should be done very carefully by an expert for a sufficiently large number of images. This makes it a time-consuming and expensive approach.

To overcome the above-mentioned limitations of unsupervised and supervised learning methods, semi-supervised learning methods have been proposed as a strategy in-between supervised and unsupervised learning. Their aim is to obtain a good learning performance without the need for a large training set of manually labelled data, by using labelled as well as unlabelled data to train a classifier with a higher accuracy than classical supervised classification.¹⁸ It is a useful technique when there are far more unlabelled data available than labelled data. This strategy is commonly used in hyperspectral remote sensing,^{19–22} but to our knowledge it has not yet been applied in hyperspectral imaging for food quality assessment. Therefore, its potential for segmentation of hypercubes of vine tomatoes and table grapes has been investigated in this study.

The aim of this study was to develop a learning algorithm capable of building a classification model requiring little human interaction, which is flexible, but not time consuming, thus making it possible to cope with different products. Taking into account the added value of semi-supervised classification over unsupervised and supervised classification, the efficiency of this technique to segment hyperspectral images of food products has been investigated. As vine tomatoes (*Solanum lycopersicum*) consist of multiple fruits attached to a stalk, and the colour of tomatoes changes significantly during ripening, going from green to red, this is an ideal case to evaluate the performance and robustness of the newly developed segmentation method. To show the flexibility of the algorithm, it is also tested for the segmentation of spectral hypercubes of green table grapes (*Vitis vinifera*), where the stalk and flesh, that are comparable in terms of colour,

have to be segmented from the background and from each other.

Materials and methods

Algorithm

The goal of the developed algorithm is to train a supervised classifier in an intuitive way, without the need for time-consuming manual labelling of the hypercubes by an expert. The training of a supervised classifier is preferred over an unsupervised approach, as unsupervised classifiers are too computationally intensive to be applied inline, may provide variable results and provide no information on the nature of the different classes. The developed algorithm aims to segment good quality products and can be split into three important parts, being the unsupervised classification, the combination of a pixel selection method with a supervised classification algorithm based on those selected pixels and the training of the model. In Figure 1, the proposed classification algorithm is schematically summarised. After the measurement of a product, an unsupervised classification algorithm is started that will look for a number of classes in the dataset. In the cases described here, three classes are defined, namely the background, the stalk and the tomatoes or the grapes. As multiple unsupervised techniques are tested, an operator has to decide which classification performed best and could be used to build a supervised classifier. If no unsupervised classification has provided a satisfactory segmentation according to the operator, this step is repeated looking for an extra

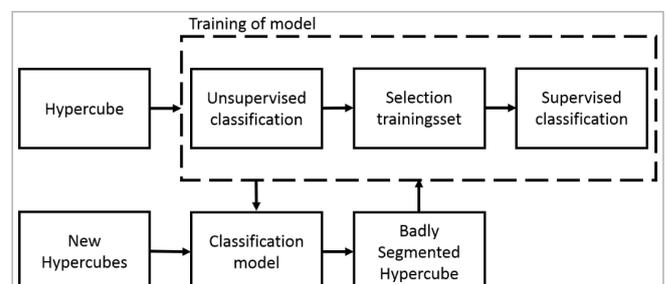


Figure 1. Schematic illustration of the different steps in the classification algorithm. The algorithm consists of three important parts, the unsupervised classification, the building of the supervised classification model and the evaluation of the model on new hypercubes, followed by a model update if necessary.

class in the data. Once a good segmentation has been achieved, a specified number of pixels is selected automatically for each class and used to build the supervised classification model. This model is then tested on a new spectral hypercube acquired for a similar product. If the segmentation for this new spectral hypercube is not satisfactory to the operator, the algorithm is repeated and the selected pixels of this new product are added to the training set to build the supervised classification model. This is repeated 10 or more times to evaluate its performance. If the segmentation is considered adequate for a sufficiently large number of subsequent products (e.g. 10), the algorithm is considered robust enough for in-line use.

Spectral pre-processing

Preprocessing is a very important step in the processing of hyperspectral data, as it aims to remove the effect of physical phenomena like light scattering from the spectra.²³ That is why the effect of four different preprocessing methods on the segmentation results has been tested. As a standard, the use of the reflectance spectra (R) without the application of preprocessing techniques is investigated. The second investigated method is the use of the Lambert–Beer law to linearise the spectra, suggesting a linear relationship between the absorbance (A_a) values and the concentration of the absorbing constituents. A_a is calculated by using the formula: $A_a = \log(1/R)$.²³ The use of scatter correction by taking the first derivative of R is the third investigated technique. Savitzky–Golay filtering²⁴ is applied using a total window size of 15 points and a second order polynomial. The weighting of spectra by means of SNV,²⁵ is the last investigated technique. This technique performs a normalisation of each sample over all variables.

Unsupervised pixel classification

To find the most interesting unsupervised classification method, eight different methods were tested. These methods can be divided into four groups, namely

- a) agglomerative hierarchical clustering in a one-step (1) and a two-step procedure (2),
- b) k -means clustering in a one-step (3) and a two-step (4) procedure,
- c) multivariate Gaussian mixture models in a one-step (5) and a two-step (6) procedure and
- d) a combination of two of the preceding methods, namely combination of (b) and (a) (7), and (b) and (c) (8).

Agglomerative hierarchical clustering (a) is a technique that starts by assigning every individual pixel to a different class. Next, at each step of the algorithm, the closest pair of classes is merged until only a pre-defined number of classes are left.¹⁷ To decide on the similarity between different classes, Ward's method has been used, which aims at minimising the within-class sums of squares.²⁶ This technique has been applied to the collection of all principal components of the pixel spectra that capture at least 1% of the total variance present in the dataset. In the agglomerative hierarchical clustering group, two methods were evaluated: (1) agglomerative hierarchical clustering and (2) agglomerative hierarchical clustering in two steps. In the two-step algorithm, the first step consists of the division of the image into two classes based on agglomerative hierarchical clustering, while the second step splits the foreground, assigned by human intervention, in the pre-defined number of classes based on a second agglomerative hierarchical clustering.

The second group consists of a one-step and a two-step k -means clustering approach (b). k -means clustering is an iterative method searching for k cluster centres, $\mu_1, \mu_2, \dots, \mu_k$, with k_c the number of classes looked for. The starting μ_i is chosen based on the k -means++ algorithm. μ_1 is randomly chosen from the data. Afterwards, each subsequent μ_i is chosen with its probability proportional to the squared distance from the already calculated nearest cluster centres.²⁷ Then, every pixel is classified according to the nearest μ_i , using the squared Euclidean distance as a distance measure. When all the pixels are classified, the mean of each cluster is recalculated. This is repeated until there is no further change in μ_i , or until 100 iterations have been performed.^{28,17} As for agglomerative hierarchical clustering, the one-step and the two-step k -means clustering approaches were applied on the score values of the principal components of the pixel spectra that capture at least 1% of the total variance present in the dataset. The k -means algorithm was repeated four times with new starting μ_i 's to avoid solutions positioned in local minima.

The third group involves the use of multivariate Gaussian mixture models (GMM) in a one-step procedure and a two-step procedure (a). A Gaussian mixture model is a probability density function that consists of a linear combination of different Gaussian distributions. GMMs search for the different Gaussian distributions present in a dataset. The basis of GMM is the expectation–maximisation (EM) algorithm. This is an iterative process consisting

of two steps. In a first step, the algorithm makes an initial guess on the parameters that are used to describe the number of classes using separate Gaussian distributions. Based on these distributions, the posterior probability of each data point to belong to a certain group is calculated (E-step). The second step consists of a recalculation of the parameters describing the Gaussian distributions based on these posterior probabilities (M-step). These steps are repeated until convergence.²⁹ This technique was applied on the preprocessed spectra, because it has been reported that replacement of the original, correlated variables by PC scores can destroy the underlying multimodal structure and make application of GMM ineffective.³⁰

The last group involves combinations of methods from two groups (d), namely the use of *k*-means clustering to distinguish the background from the foreground as step 1 and the use of a multivariate Gaussian mixture model or a hierarchical clustering in step 2. In this fourth group, the same settings have been chosen for the individual methods as in the first three groups.

Pixel selection and supervised classification

As each spectral hypercube contains a spectrum for each pixel in the image, there is a very large number of spectra available for building the classification models. This makes training of the supervised classification models on all available pixel spectra computationally very intensive. Moreover, pixel spectra corresponding to the same object are much more similar than those belonging to different objects. Therefore, careful selection of the training spectra is essential to obtain a robust and accurate classifier. In this study, two-pixel selection methods have been tested for this purpose: random and Kennard–Stone pixel selection.³¹ The effect of the number of selected pixels has been investigated as well. The number of selected pixels per class was chosen based on a Fibonacci sequence with 5 as the starting number and 987 as the maximum number of selected pixels per class. Each selection was repeated five times to minimise the effects of the random sampling.

The subsets of pixels and corresponding spectra that were selected by these methods were then used to train the supervised classification methods. In this study, five methods for supervised classification were tested. The first is the use of a convex hull to constrain the space. A convex hull is the smallest collection of the convex sets containing the points of the particular class.³² To

calculate the convex hull, a PCA model based on all the selected pixels was built first. This model was then applied on each class and the scores of each pixel in the selection were used to calculate the convex hull, defining the boundaries of each class in the two-dimensional space spanned by the first two principal components. For every new image, each pixel will be assigned to a specific class, or to no class, dependent on the position of the pixel in the space. The second supervised technique was PLS-DA, using three latent variables and the third used technique was SVM. The principal components explaining at least 1% of the total variation present in the dataset containing the selected pixels were selected as input variables for the SVM. For each class, a binary SVM classifier using a Gaussian kernel was built to discriminate the pixels belonging to that class from those belonging to the other classes. The fourth investigated technique was soft independent modelling of class analogies (SIMCA), which is a technique that performs a PCA analysis on the selected pixels of each class. For each new spectrum the algorithm calculates a distance measure to determine how likely it is that this sample belongs to that class. This sample is then assigned to the class for which the likeliness is the largest.³³ The first three principal components for each class were used, as these effectively explained the larger part of the variance in the different classes. A combination of a reduced Hotelling's T^2 T_{red}^2 (T^2 divided by its 95% confidence limit $T_{0.95}^2$) and reduced Q-residual Q_{red} (Q divided by its 95% confidence limit $Q_{0.95}$) was used as distance measure (Equation 1). The Q-statistic is a measure for the distance between the sample and its projection on the model hyperplane, while the Hotelling's T^2 describes the distance between this projection on the model hyperplane and the centre of the model hyperplane.

$$D = \sqrt{T_{red}^2 + Q_{red}^2} = \sqrt{\left(\frac{T^2}{T_{0.95}^2}\right)^2 + \left(\frac{Q}{Q_{0.95}}\right)^2} \quad (1)$$

The last investigated technique is multinomial logistic regression, which is a technique that, opposed to discriminant analysis, assumes a multinomial distribution of the dependent variables²⁶ instead of a multivariate normal distribution.³⁴

The performance of these algorithms was compared in terms of the classification accuracy, the time needed to segment one image and the number of selected pixels for each pixel selection method. The classification accuracy was quantified by calculating three different metrics,

using the unsupervised segmentation results as ground-truth: the accuracy (OA)³⁵ (Equation 2), the sensitivity (Equation 3) and the specificity (Equation 4):

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (3)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (4)$$

where TP are the true positives, TN the true negatives, FN the false negatives and FP the false positives. True positives are pixels that are correctly classified in the corresponding class, true negatives are pixels that are legitimately classified as not belonging to the respective class. False negatives are pixels classified outside the class, but that are in fact part of the class and false positives are pixels classified wrongly as belonging to the class.

The supervised classification models that are built during the training phase are tested on five new spectral hypercubes, forming the validation set. Per number of pixels to select, five different training sets are selected. These are used to train the supervised classification models. By applying these different supervised models to the validation set, the mean OA can be plotted against the number of selected pixels. Next, a straight line is plotted through the mean OA when using five pixels per class and the mean OA using 987 pixels per class. The number of pixels where the distance between this line and the mean OA is the largest is selected based on the technique proposed by Zack *et al.*³⁶ as the best trade-off between training efficiency and classification performance. The number of selected pixels should be low to avoid using many similar pixels in the training phase. Using mainly pixels from, for example, the centre of the tomato, decreases the influence of pixels at the border, introducing mistakes during classification. As the number of selected pixels is not identical for each spectral hypercube in the validation set, the number of pixels that is finally chosen is the number of selected pixels for the five spectral hypercubes that is most common. The computation time required to apply a supervised classification model in the validation phase on a spectral hypercube which has not been used for training the model was also determined, enabling a relative comparison between the times needed for the different classification algorithms.

This is an indication for the relative time lag that these algorithms would introduce in a food processing line.

Augmentation of the training set

The last step of the algorithm is the augmentation of the training set. When the above algorithm has been applied on one spectral hypercube, the developed supervised classification model will be tested on at least 10 new spectral hypercubes to see if the algorithm works well. The segmentation result for each of these spectral hypercubes is visually inspected by the operator to detect unsuccessful segmentations. When a poor segmentation has been detected, the procedure, starting from the unsupervised classification, is repeated for this particular spectral hypercube. After selection of the most successful segmentation from the unsupervised stage, the same number of pixel spectra as in the first used spectral hypercube is selected. This selection is independent from the existing training set, to ensure that information present in this training set is not lost. It is used to augment the training set and to build a new supervised classifier on this set. This is repeated until 10 consecutive spectral hypercubes have been tested and their segmentation is approved. In this way, the training set will cover information from a large variety of samples.

Implementation

All calculations have been carried out using Matlab (version 8.4.0, The Mathworks Inc., Natick, MI, USA). For the spectral preprocessing, unsupervised segmentation and supervised segmentation the PLS Toolbox (version 8.0, Eigenvector Research Inc., Manson, WA, USA) was used, while the convex hull was calculated using a standard Matlab function. For SVM and logistic regression the statistics and machine learning toolbox in Matlab was used.

Hyperspectral data

Samples

Vine tomatoes (*Solanum lycopersicum*) were harvested manually at two different Belgian greenhouses. In June 2014, 36 vines of the cultivar 'Merlice' were harvested at "Proefstation voor de groenteteelt" (Sint-Katelijne-Waver, Belgium), while in August 2014, 36 vines of the cultivar 'Prunus' were harvested at "Proefcentrum Hoogstraten" (Hoogstraten, Belgium). During each harvest period tomatoes were harvested in three different ripeness stages, namely unripe, commercial harvest and overripe.

Each ripeness stage was divided into four groups and each group was measured at a different day, to obtain a dataset with large variability. Group 1 was measured at day 0 (= day of harvest), group 2 at day 3, group 3 at day 7 and group 4 at day 10. After each measurement, the colour of each tomato on the vine was measured using a spectrophotometer calibrated for colour measurements (CM-2600d, Konica Minolta, Osaka, Japan). Between harvest and the time of measurement the tomatoes were stored at a constant temperature of 18°C and a relative humidity of 80%.

To evaluate the general applicability of the algorithm it was also tested on green table grapes. Five vines of the cultivar 'Sundance Seedless' and five vines of the cultivar 'Thompson Seedless' were purchased in a local supermarket and measured on the day of purchase.

Hyperspectral data acquisition

The hyperspectral setup used in this research is a line-scan imaging system that consists of an illumination unit, a spectrograph, a lens and a CCD camera. The setup was developed to be applicable in an industrial environment, requiring a compact and hygienic design. This meant that dome illumination was not acceptable. Based on ray-tracing simulations, an illumination setup was optimised to obtain a uniform illumination.¹³ The illumination unit consists of four DC-halogen lamps (Decostar Alu 35, OSRAM, Munich, Germany) each with a diffuser (TSG-LEGB, Knight Optical Ltd, Kent, United Kingdom) mounted in front of them and positioned in an arc shape around the linear translation stage (Franke GmbH, Aalen, Germany). The DC-halogen lamps were chosen to avoid interference effects in the acquired hypercubes of the

flickering at 100 Hz which is typical for AC lamps. An Inspector V10 spectrograph (Spectral Imaging Ltd, Oulu, Finland) was coupled to a Vis/NIR lens (CM120 BK15 COMPACT, Schneider Kreuznach, Bad Kreuznach, Germany) with a focal length of 17 mm. This combination was mounted on a 12-bit monochrome CCD camera (TXG14NIR, Baumer, Frauenfeld, Switzerland) with a resolution of 1392 × 1040 pixels and a spectral sensitivity between 325 nm and 985 nm and placed above the linear translation stage which has a step resolution of 6 μm (Figure 2). By placing a piece of conveyor belt on top of the translation stage, a conveyor belt was imitated. This setup was placed in a dark room to avoid interference from external light sources.

A PC (Intel® Core™ i5, CPU: 2.8 GHz, RAM: 6 GB), running Labview (version 9.0, National Instruments Corporation, Austin, USA), was used to control the linear translation stage and the camera. The samples were measured by placing the tomatoes on the translation stage. This translation stage was moved with a step size of 0.2 mm under the camera, and each step, a line scan image was captured. The step size was set to 0.2 mm to get a comparable resolution in the x- and y-directions. The exposure time of the camera was optimised for the measurements of the tomatoes, by measuring vines of different ripeness with a range of exposure times. The exposure time was set at a value that was as high as possible without getting saturated regions in the image, resulting in an exposure time of 35 ms. The Labview program saved each line scan image. At the beginning of each session of hyperspectral measurements a white reference and the dark current of the camera were measured. The white reference was measured using a diffuse

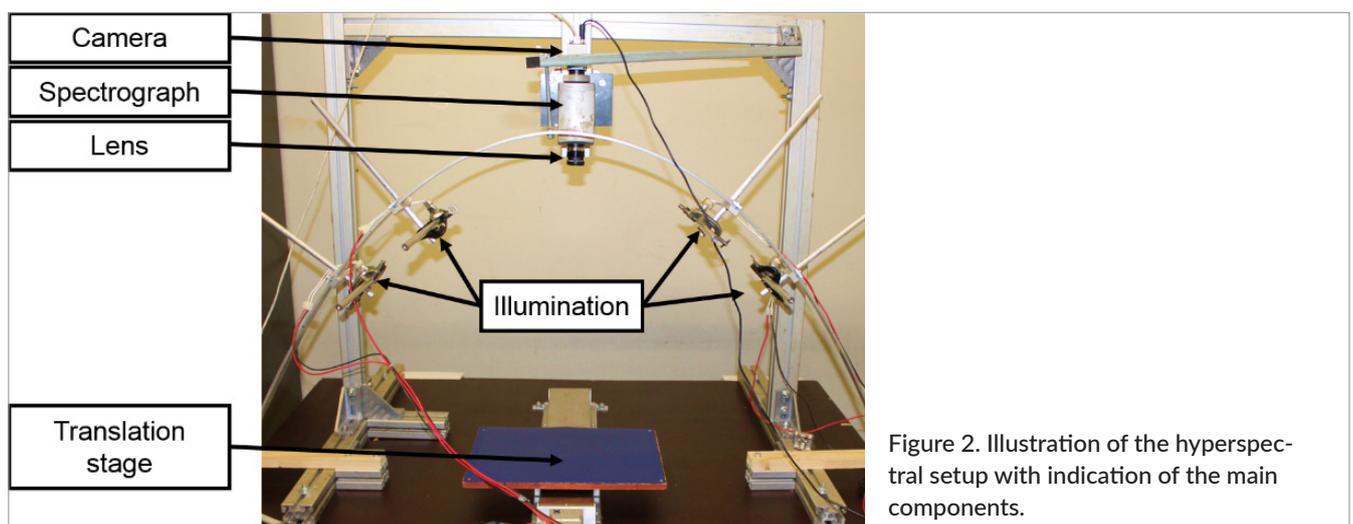


Figure 2. Illustration of the hyperspectral setup with indication of the main components.

Spectralon plate with a reflectivity of 99% (Labsphere, North Sutton, USA), while the dark current of the camera was measured by covering the lens of the camera with its lens cap.

The table grapes were measured with the same setup as the vine tomatoes, with the same settings except for the exposure time, which was set to 15 ms. The white reference and the dark current of the camera were measured in the same way as in the case of the vine tomatoes.

Data processing and analysis

The acquired line scan images were loaded into Matlab (version 8.4.0, The Mathworks Inc., Natick, MI, USA) and a spectral hypercube was built per sample. A binning-factor of four was used as well in the spatial (x- and y-direction) as the spectral direction, resulting in square pixels with a spatial resolution of 0.8 mm and a spectral resolution of 2.56 nm for both the vine tomatoes and the table grapes. The relative reflectance (R) for each pixel was calculated as described by Wallays *et al.*³⁷ (Equation 5). The wavelengths with a low signal-to-noise ratio were removed, thus reducing the spectra to the 425–985 nm range.

$$R = \frac{I_{\text{sample}} - I_{\text{dark current}}}{I_{\text{reference}} - I_{\text{dark current}}} \quad (5)$$

with I_{sample} , $I_{\text{reference}}$ and $I_{\text{dark current}}$ the measured intensity for the sample at a given pixel, the mean of the measured intensity of the white reference and the mean of the measured intensity of the dark current of the camera.

Results and discussion

Vine tomatoes

Variation in ripeness stages

In Table 1, an overview is given of the hue values which were measured for the tomatoes with the spectropho-

tometer calibrated for colour measurements. It can be seen that the dataset included a wide range of tomato colours with hue values between 42.39° and 83.01°. When the tomatoes were measured 10 days after harvest, the variation between the different ripeness stages at harvest had reduced considerably, as most fruits had become fully ripe. This large variation in fruit colours poses a challenge for the segmentation algorithm as there will be large variability within the class of tomato objects to be segmented.

Effect of preprocessing on pixel spectra

In Figure 3, the mean spectra \pm one standard deviation of a manually selected part of the background, the stalk and unripe, commercially harvested and overripe tomatoes are shown. The mean spectrum of the background is very different from the mean spectra of the biological materials. As the background has a dark blue colour, the background has a very low R in the complete measured wavelength range, while there is a clear signature present in the tomatoes and the stalk. When inspecting the raw spectra in sub-figure a, the mean spectra of the unripe tomato and the stalk show a very clear dip around 680 nm which can be attributed to the absorption by chlorophyll.^{38,39} During ripening, the reflectance R at this wavelength rises, due to the degradation of the chlorophyll. The tomatoes also strongly absorb the blue light, because of the strong presence of carotenoids, especially lycopene.⁴⁰ The rise in lycopene, with three main absorption peaks at 446 nm, 471 nm and 505 nm, during ripening can be seen in Figure 3b, where the absorption of blue light is higher in commercially harvested tomatoes and overripe tomatoes compared to unripe tomatoes. The difference between the commercially harvested tomatoes and the overripe tomatoes is small in this wavelength region, suggesting that the production of lycopene occurs until the last

Table 1. Summary of the average hue (°) values as measured with the calibrated spectrophotometer on different days after harvest for the tomatoes harvested at different ripeness stages (1: Unripe harvest, 2: Commercial harvest, 3: Harvest after the commercial harvest period).

	'Merlice'			'Prunus'		
	1	2	3	1	2	3
Day 0	81.20	59.35	48.95	83.01	60.35	46.09
Day 3	60.01	51.81	46.98	61.84	50.78	44.42
Day 7	48.59	47.50	44.10	57.22	44.87	45.81
Day 10	45.92	45.06	43.17	55.13	42.59	42.39

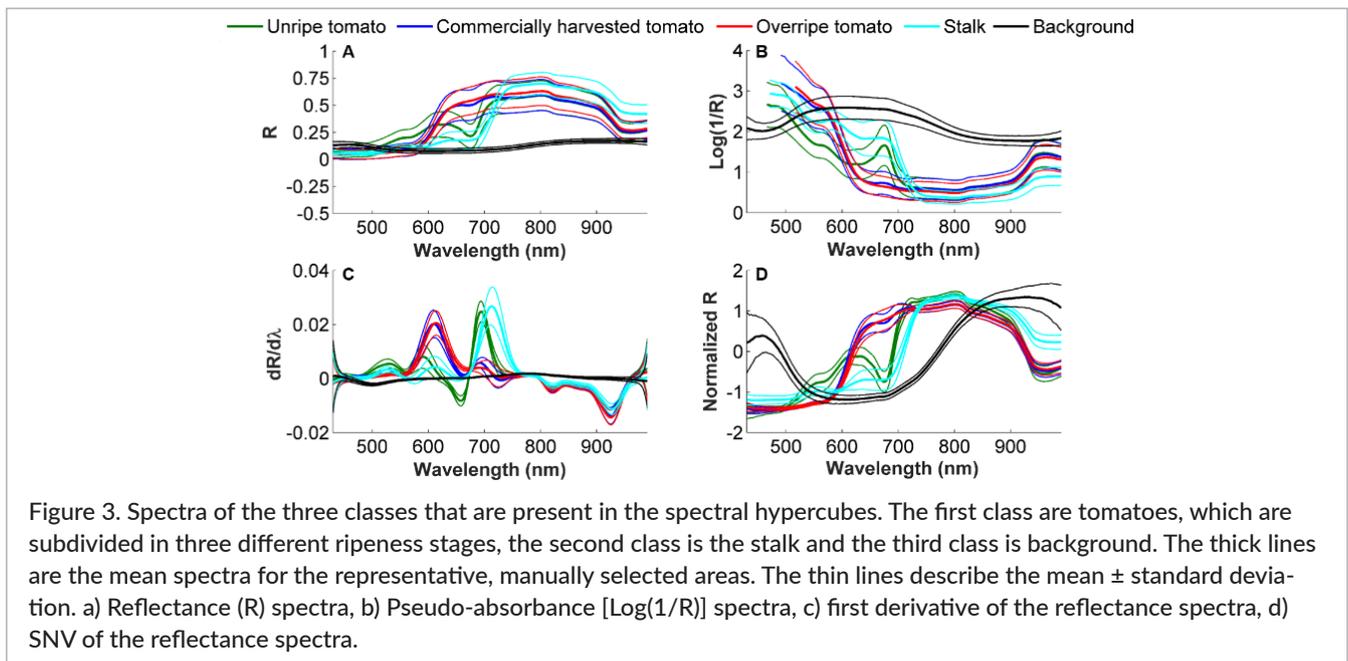


Figure 3. Spectra of the three classes that are present in the spectral hypercubes. The first class are tomatoes, which are subdivided in three different ripeness stages, the second class is the stalk and the third class is background. The thick lines are the mean spectra for the representative, manually selected areas. The thin lines describe the mean \pm standard deviation. a) Reflectance (R) spectra, b) Pseudo-absorbance [$\text{Log}(1/R)$] spectra, c) first derivative of the reflectance spectra, d) SNV of the reflectance spectra.

ripening stages of the tomato, which is in agreement with the observation of Arias *et al.*⁴¹

Unsupervised segmentation

As step 1 of the developed algorithm consists of the use of an unsupervised method to segment the spectral hypercubes, different combinations of unsupervised algorithms and typical spectral preprocessing techniques have been tested. To accurately segment the spectral hypercubes in tomato, stalk and background, a good distinction between the different classes is required. As can be seen from Figure 3a, the difference between the stalk and the tomato spectra is not as clear as that between tomato or stalk and background. This makes it challenging to accurately classify the pixels in an image as either stalk or tomato. To increase the contrast between the different classes, different preprocessing techniques were investigated. The use of a linearisation technique (Figure 3b), by using $\text{log}(1/R)$, does not provide a better distinction. However, first derivative and SNV preprocessing result in an enhancement of the differences between the classes (Figure 3c and Figure 3d), which can be explained by the fact that both techniques are able to correct for variation in the spectra caused by differences in the scatter properties.²³ In Table 2, the segmentation performances are summarised for the different combinations of pre-processing techniques and unsupervised classification algorithms.

The results in Table 2 show that the preprocessing technique used has a large impact on the segmentation performance. As expected from Figure 3, the segmentation results for the reflectance and pseudo-absorbance [$\text{Log}(1/R)$] spectra are poor with only 8.33% and 0% of the total number of images correctly segmented. When using the pseudo-absorbance as a preprocessing step to classify the pixels in the spectral hypercubes, the background was often segmented into shadow and non-shadow parts. The results obtained with SNV preprocessing are better, but still rather poor with a cumulated segmentation accuracy of 26.39% over all unsupervised classification methods. First derivative preprocessing over an interval of 15 wavelength variables, which corresponds to a range of 38.4 nm, is much more successful with a correct segmentation rate of 86.11% over all methods. This better performance can probably be explained by the fact that the first derivative enhances the curvature in the spectra caused by absorption by chemical components, while reducing the influence of the baseline, resulting from differences in light scattering. In this way, the spectral differences between the different classes are increased with respect to the variation within each class, as can be seen from Figure 3c. In Figure 4, the segmented images obtained with the different unsupervised classification algorithms on the first derivative preprocessed spectra are illustrated for one truss of tomatoes of the cultivar 'Prunus' at commercial harvest. Unsupervised algorithms c, e and i did not segment the images correctly,

Table 2. Percentage of the vine tomato images segmented correctly by using different combinations of unsupervised classification algorithms and preprocessing techniques. All the measured vine tomato images were taken into account.

	No prep.	Log (1/R)	Derivative (window: 15)	Derivative (window: 25)	SNV
<i>k</i> -means	1.39	0.00	80.56	79.17	1.39
Hier. clust.	1.39	0.00	55.56	50.00	1.39
2-step <i>k</i> -means	4.17	0.00	55.56	52.78	1.39
2-step hier. clust.	0.00	0.00	0.00	0.00	0.00
GMM	0.00	0.00	38.89	38.89	2.78
2-step GMM	0.00	0.00	43.06	34.72	22.22
Combi 1	2.78	0.00	50.00	44.44	0.00
Combi 2	0.00	0.00	0.00	0.00	0.00
Total	8.33	0.00	86.11	86.11	26.39

No prep.: No preprocessing, hier. clust.: agglomerative hierarchical clustering, 2-step hier. clust.: 2-step agglomerative hierarchical clustering, GMM: multivariate Gaussian mixture model, 2-step GMM: 2-step multivariate Gaussian mixture model, Combi 1: combination of *k*-means and multivariate Gaussian mixture model, Combi 2: combination of *k*-means and agglomerative hierarchical clustering, total: percentage of the vine tomatoes classified correctly by at least one of the unsupervised classification algorithms, Total: The number of images that is segmented correctly by one of the unsupervised classification algorithms.

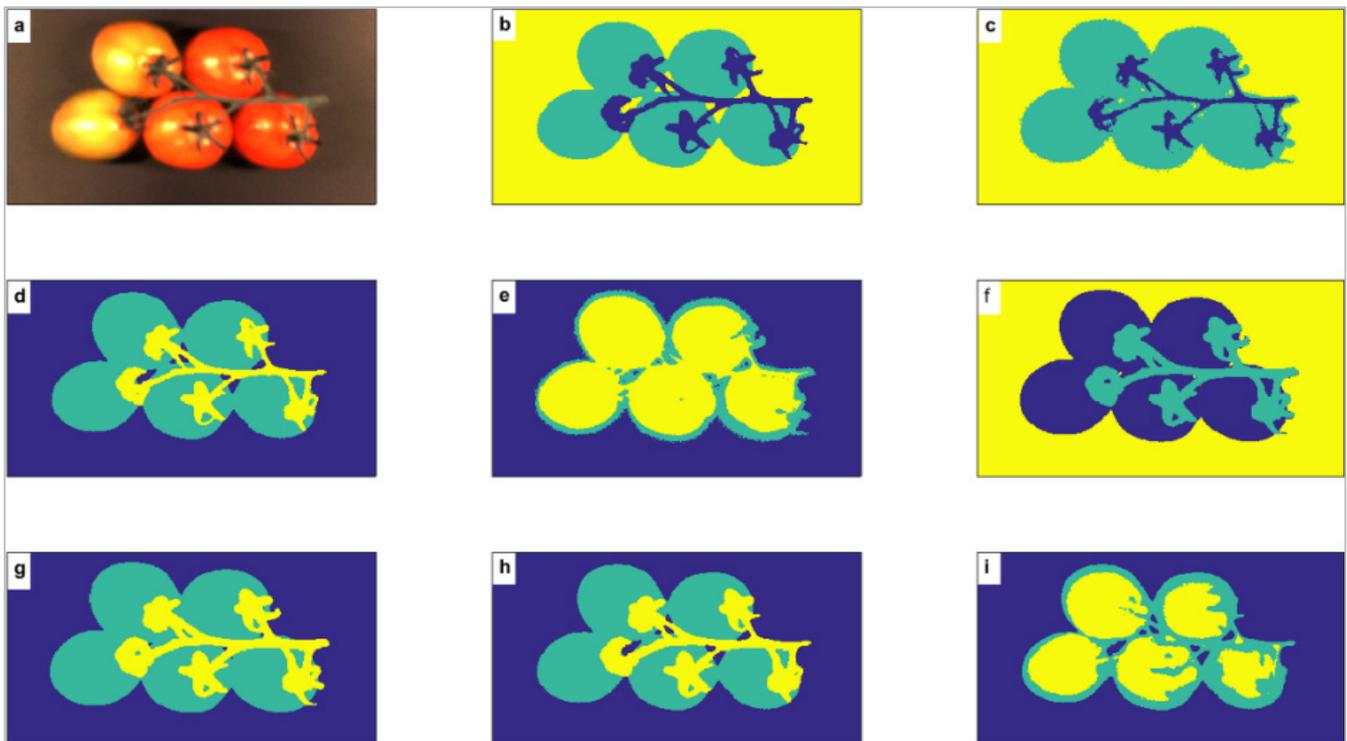


Figure 4. Representative segmented images of a truss of tomatoes of the cultivar 'Prunus' harvested at the commercial harvest time as obtained with the different unsupervised classification algorithms on first derivative pre-processed spectra: a) RGB image; b) *k*-means clustering; c) hierarchical clustering; d) 2-step *k*-means clustering; e) 2-step hierarchical clustering; f) multivariate Gaussian mixture model; g) 2-step multivariate Gaussian mixture model; h) combination of *k*-means and multivariate Gaussian mixture model; i) Combination of *k*-means clustering and multivariate Gaussian mixture model.

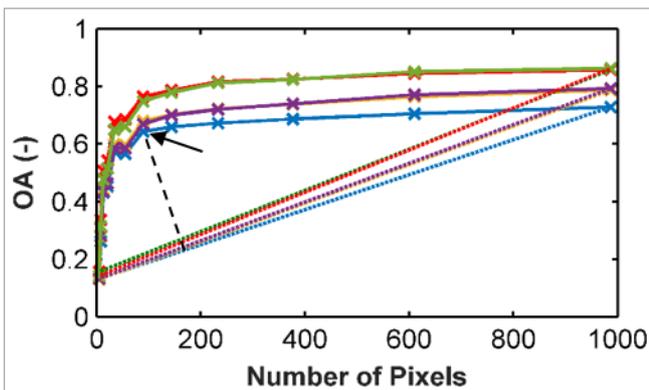


Figure 5. Overall accuracy (OA) as a function of the number of selected pixels for five different tested images. The supervised algorithm used is the convex hull with a random pixel selection. The full lines represent the obtained classification accuracies, the dotted lines represent the straight line between the classification result when using 5 pixels and the classification result when using 987 pixels. The number of pixels where the distance between the full line and the dotted line is the largest was selected. This distance for the blue lines is shown by the black dashed line. The arrow marks the number of selected pixels.

while the other algorithms nicely segment the image in background, tomatoes and stalk. In each investigated situation, algorithm e and i did not segment the images correctly. These algorithms use a two-step architecture with as second step agglomerative hierarchical clustering. This unsupervised hierarchical clustering was not able to accurately discriminate the tomato flesh from the stalk, which can probably be explained by the fact that the number of tomato pixels in the image is quite large compared to the number of stalk pixels.

Pixel selection and supervised classification

Step 2 of the algorithm is a selection of the number of pixels used for training the supervised classification models. As the performance of these supervised models largely depends on the representativeness of the training pixels that are provided for training the model, these pixels should be selected in an efficient way. As the accuracy, specificity and sensitivity depend on the number of training pixels used, it is important to obtain as good results as possible with as few training pixels as possible. In Table 3, the optimal number of training pixels and the corresponding results obtained on a validation set consisting of five spectral hypercubes which were not used for training the supervised classifier, are summa-

rised. The optimal number of training pixels is determined based on Figure 5. The most common number of selected pixels for the five spectral hypercubes of the validation set was decided to be the best suited number of selected pixels per class.

When using only the pixels for one spectral hypercube to train the supervised classifiers, the technique that is most suited for use in industrial applications is PLS-DA. It is the technique with the highest OA (0.969), a high sensitivity (>0.954) and a quite good specificity (>0.846). Also, the calculation time is acceptable for in-line applications. Other techniques with a good overall classification performance are SIMCA and SVM, but their computation time is considerably higher than for PLS-DA. These higher calculation times for SVM and SIMCA can probably be explained by the fact that these methods first have to project the data on one or more PCA models, while PLS-DA only involves multiplication of the spectral vector by the vectors of regression coefficients for the different classes. The technique with the lowest computation time is logistic regression, but the classification performance for this method is also lower. Finally, the convex hull gave results with a very high sensitivity. However, the specificity is rather low, meaning that a considerable number of pixels cannot be classified by this method, resulting in a low OA.

The Kennard–Stone algorithm has been developed to select pixels that cover the dataset uniformly, while the random pixel selection does not take any information regarding the pixels into account. Therefore, it is remarkable that the results obtained with the Kennard–Stone algorithm are not always better or do not always result in a smaller number of selected pixels than the random selection. A possible explanation for this observation may be the large difference between the pixel spectra of the different classes, which makes the selection of the most informative pixels less critical.

As can be seen from Table 3, the classification results improve considerably when a second spectral hypercube is added in the training phase of the model. Especially the specificity improves significantly, while the sensitivity only improves marginally. The combined effect is an increase in the OA of at least 4%. The effect on the segmentation performance of including a second spectral hypercube is illustrated in Figure 6. The improvement in the segmentation is clearly visible with correct classification of all the background pixels and a correct classification of green tomatoes as tomato. Especially

Table 3. Segmentation results for the validation set consisting of five spectral hypercubes of vine tomatoes with indication of the selected number of training pixels. The used ground-truth was the result of an unsupervised classification which was evaluated as correct through visual inspection.

		No. of pixels	OA	Sens.	Spec.	Time (s)
Random pixel selection	Convex hull	89	0.7005	0.9951	0.5316	0.4037
	SVM	89	0.8888	0.9538	0.8458	0.8250
	PLS-DA	5	0.9211	0.9654	0.8791	0.4864
	SIMCA	34	0.8204	0.8819	0.8070	1.2727
	Log. regression	89	0.8918	0.9321	0.8145	0.0242
Kennard–Stone pixel selection	Convex hull	55	0.7411	0.9974	0.5955	0.3996
	SVM	377	0.9155	0.9627	0.8614	1.2136
	PLS-DA	21	0.9202	0.9657	0.8464	0.4822
	SIMCA	13	0.9105	0.9449	0.8828	1.2546
	Log. regression	89	0.8674	0.9097	0.7790	0.0258
Random pixel selection using 2 images	Convex hull	89	0.8340	0.9868	0.6966	0.486
	SVM	55	0.9557	0.9735	0.8985	1.018
	PLS-DA	5	0.9689	0.9808	0.9332	0.5167
	SIMCA	55	0.9013	0.9624	0.9065	1.3487
	Log. regression	55	0.9566	0.9695	0.8953	0.0570
Kennard–Stone pixel selection using 2 images	Convex hull	144	0.8685	0.9863	0.7314	0.5004
	SVM	5	0.9648	0.9728	0.9123	1.3171
	PLS-DA	5	0.9695	0.9837	0.9463	0.5139
	SIMCA	13	0.9533	0.981	0.9298	1.3170
	Log. regression	55	0.9462	0.9620	0.8722	0.0415

for the left image, this results in a higher OA, specificity and sensitivity. In the right image, the segmentation of small instances, like some parts of the stalk, is improved. It should also be noted that in this case pixel selection by the Kennard–Stone algorithm in the training phase resulted in slightly better results than random pixel selection. This might indicate that the Kennard–Stone pixel selection algorithm is more efficient in capturing the additional information contained in the second spectral hypercube than the random pixel selection.

Table grapes

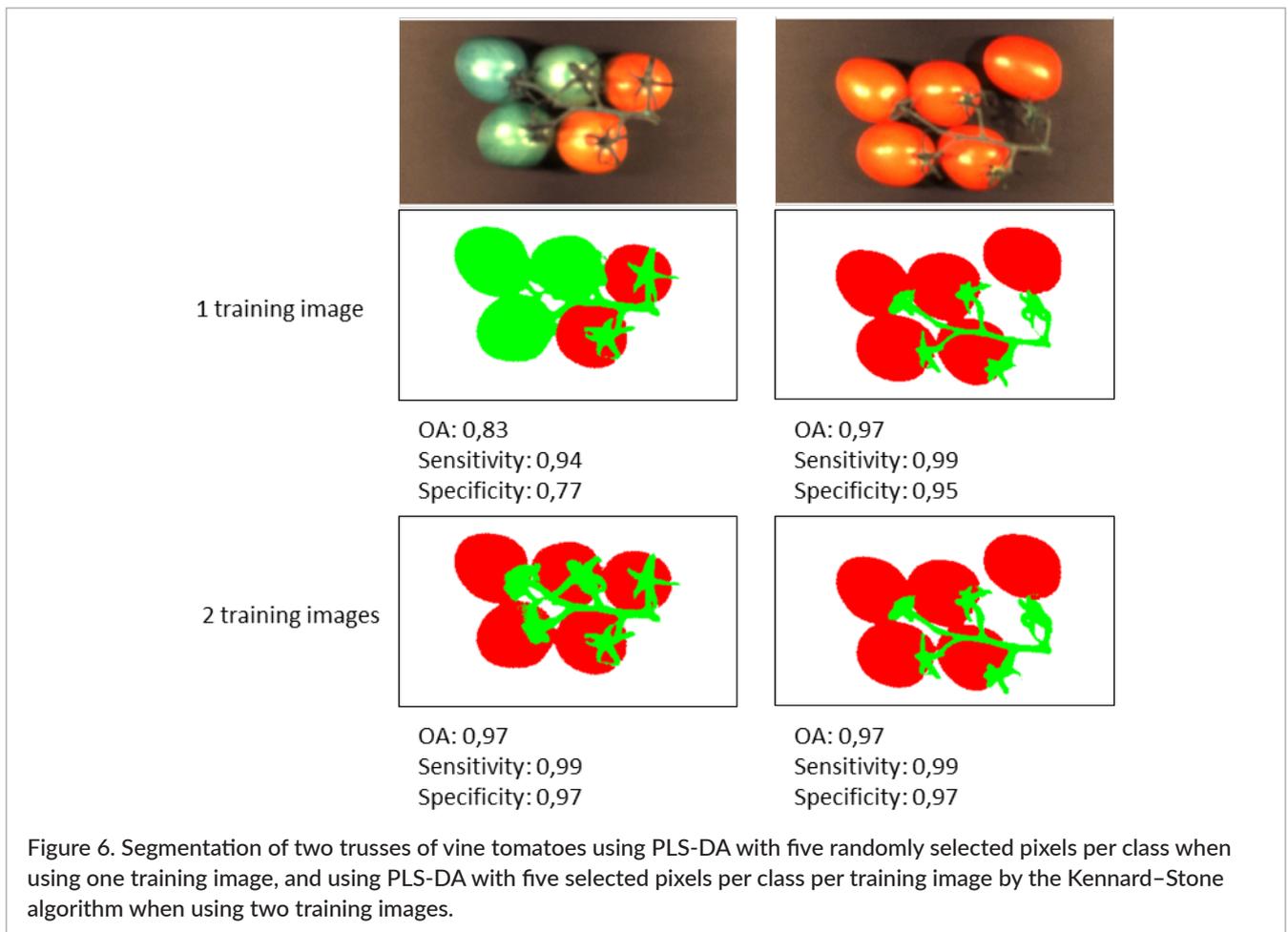
Unsupervised segmentation

In Table 4 the results for the unsupervised segmentation of the table grape hypercubes are summarised. While the segmentation of the table grapes was not successful without pre-processing, 1st derivative pre-processing in combination with the combination of *k*-means and GMM resulted in a correct segmentation of 80% of the images. The algorithms that were able to segment at least one

image correctly were all based on a multivariate Gaussian mixture model. This can be explained by the small number of stalk pixels in the images and the better ability of the GMMs to deal with unequal numbers of pixels between the classes. When comparing these results with the unsupervised results obtained for the vine tomatoes, it becomes clear that different unsupervised techniques may be more appropriate for segmenting different products. This can be explained by the different nature of the unsupervised classification algorithms which have been considered. For example, GMMs are better suited to find classes of different size compared to *k*-means clustering. Similar to the case of vine tomatoes, agglomerative hierarchical clustering in the second step of the algorithm did not result in good segmentation results for the table grapes.

Pixel selection and supervised classification

The results for the supervised pixel classification for five trusses of green table grapes are summarised in Table 5.



SVM, PLS-DA and logistic regression all resulted in a very high OA (>0.95). The difference between these techniques is in the specificity, sensitivity and computation time required to segment one image. When ignoring the time needed to segment the images, SVM gives the best results. Next to a very high OA, the specificity and sensitivity are also very high. However, as the algorithm is slower than the other two techniques, it is less suitable for in-line use. PLS-DA, trained on pixels selected using the Kennard–Stone algorithm, also has a very good OA, specificity and sensitivity and is faster than SVM. Logistic regression is very fast compared to the other techniques, but the specificity is lower (<0.86). When investigating this lower specificity, it was observed that the pixels are mostly wrongly assigned to the stalk, while belonging to the background or the table grapes (specificity stalk = 0.628, specificity table grapes = 0.957, specificity background = 0.978). This result has a limited influence on the OA and the sensitivity as the number of pixels belonging to the stalk is much smaller than the number of

pixels belonging to the background or the grapes. When using pixels from two different spectral hypercubes in the training set of the algorithm, the results of PLS-DA and logistic regression slightly improve, while the results of SVM and SIMCA decrease. It should be noted that the improvement was small, because the initial result, using only one hypercube already gave a high OA.

Discussion

Most studies on hyperspectral imaging of agrofood products have been aimed at detecting defects in fruit and vegetables or classification of a product in a certain quality class.^{10,14,42–44} While the majority reported quite high accuracies, it is difficult to compare these as results were typically not reported on the pixel level. On the pixel level, classification accuracies of more than 99.6% were achieved to discriminate between terrestrial and fish species in animal protein by-prod-

Table 4. Percentage of all the hypercubes acquired for table grapes which were segmented correctly by using different combinations of unsupervised classification algorithms.

	No preprocessing	1st derivative (window: 15)
<i>k</i> -means	0	0
Hier. clust.	0	0
2-step <i>k</i> -means	0	0
2-step hier. clust.	0	0
GMM	0	30
2-step GMM	0	40
Combi 1	0	80
Combi 2	0	0
Total	0	90

hier. clust.: agglomerative hierarchical clustering, 2-step hier. clust.: 2-step agglomerative hierarchical clustering, GMM: multivariate Gaussian mixture model, 2-step GMM: 2-step multivariate Gaussian mixture model, Combi 1: combination of *k*-means and multivariate Gaussian mixture model, Combi 2: combination of *k*-means and agglomerative hierarchical clustering, total: percentage of the table grapes segmented correctly by at least one of the unsupervised classification algorithms; Total: The number of images that is segmented correctly by at least one of the unsupervised classification algorithms.

ucts by using four different algorithms (Mahalanobis distance, Kennard–Stone, spatial interpolation and binning) to select an appropriate subset for building a PLS-DA model. Spatial interpolation was decided to be the optimal selection criterion.⁴⁵ The classification results they reported are slightly better than the results achieved in our research, but the investigated application is too different to allow a useful comparison. Portalés and Ribes-Gómez¹² were able to detect stalks and unknown objects in batches of vines with a very high accuracy. The RMSE they achieved was as low as 0.34% in the case of the detection of grape stalks and 0.08% when detecting unknown objects. However, these results were obtained through spectral filtering followed by spatial filtering. It should be noted that the described researches used an extensive training phase to build segmentation models. This makes the algorithms very application dependent and limits their flexibility. The algorithm proposed in this study can be transferred more easily to other applications, as was demonstrated by its successful application to both vine tomatoes and table grapes.

It should be noted that the algorithm was developed to train segmentation models to segment good quality products. However, when small defects are present in the dataset, it is expected that this approach will still provide good quality segmentation when the spectral differences between defects and good quality products are small compared to the differences between the different classes. If necessary, this method could be made robust against the presence of defect pixels by replacing PCA and PLS-DA by their robust counterparts.⁴⁶

Conclusions

A semi-supervised algorithm for training the segmentation of hyperspectral images of agrofood products has been proposed and tested on vine tomatoes and table grapes. With only little human interaction, it was possible to build highly accurate supervised models which provided an accurate segmentation of the classes in a spectral hypercube. The segmentation of spectral hypercubes of vine tomatoes in tomato flesh, stalk and background was used as a first case study. First derivative preprocessing of the spectra resulted in a large improvement in the performance of the unsupervised classification methods by removing the intensity variation due to product curvature and light scattering effects. Second, PLS-DA was found to be the most suitable supervised classification method for segmentation of hyperspectral images of vine tomatoes, as it achieved very good classification results within a short calculation time. By augmenting the training set with spectra of a second hyperspectral image which was poorly segmented by the initial supervised segmentation algorithm, the results improved considerably. This augmentation resulted in an overall accuracy for vine tomatoes of 96.95%, a specificity of 98.37% and a sensitivity of 94.63% for the test set of five additional hypercubes. Next to segmenting vine tomatoes, the algorithm was successfully applied for the segmentation of table grapes, showing its general applicability. The best performing unsupervised segmentation algorithm for table grapes was different from the best algorithm for vine tomatoes, proving the need to investigate several different unsupervised segmentation algorithms for different products. The final PLS-DA model to segment table grapes resulted in an overall accuracy of 98.52%, a

Table 5. Segmentation result based on maximum of pixels needed for OA. These are the results for the validation set consisting of five spectral hypercubes of table grapes. The used ground-truth was the result of a correct unsupervised classification.

		No. of pixels	OA	Sens.	Spec.	Time (s)
Random pixel selection	Convex hull	34	0.4650	0.9860	0.3114	0.2068
	SVM	144	0.9793	0.9915	0.9574	0.4893
	PLS-DA	55	0.9530	0.9795	0.8674	0.3347
	SIMCA	21	0.7678	0.9204	0.8194	0.9088
	Log. regression	89	0.9662	0.9808	0.8448	0.0078
Kennard–Stone pixel selection	Convex hull	144	0.5550	0.9910	0.3765	0.2222
	SVM	144	0.9865	0.9921	0.9302	0.7021
	PLS-DA	21	0.9826	0.9894	0.9232	0.3345
	SIMCA	13	0.7956	0.9306	0.8327	0.9058
	Log. regression	89	0.9642	0.9777	0.8542	0.0077
Random pixel selection using 2 images	Convex hull	144	0.5000	0.9807	0.3555	0.2231
	SVM	89	0.9750	0.9903	0.9723	0.5633
	PLS-DA	13	0.9656	0.9836	0.8909	0.3352
	SIMCA	13	0.7402	0.9114	0.8041	0.9109
	Log. regression	55	0.9772	0.9863	0.8547	0.0078
Kennard–Stone pixel selection using 2 images	Convex hull	144	0.5462	0.9921	0.3592	0.2280
	SVM	21	0.9844	0.9922	0.9606	0.4233
	PLS-DA	144	0.9852	0.9915	0.9452	0.3350
	SIMCA	233	0.7915	0.9292	0.8298	0.9344
	Log. regression	55	0.9706	0.9820	0.8433	0.0077

Sens.: Sensitivity; Spec.: Specificity; Time is the mean time needed for segmenting one spectral hypercube using the supervised classifier.

specificity of 94.52% and a sensitivity of 99.15% for the test set of five additional hypercubes.

Acknowledgements

The research leading to these results was partially funded by the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 311987 (PicknPack). Jeroen van Roy has been funded by the Institute for the Promotion of Innovation through Science and Technology (IWT-Vlaanderen, Grant Number 121620). Bart De Ketelaere is research manager financed through the Industrial Research Fund at KU Leuven. Special thanks go to Dr ir. Ans de Roeck and Dr ir. Bert Verlinden for their advice with respect to tomatoes and to the "Proefstation voor de groenteteelt" and "Proefcentrum Hoogstraten" for providing the tomatoes.

References

1. G. ElMasry and D.-W. Sun, "Principles of hyperspectral imaging technology", in *Hyperspectral Imaging for Food Quality Analysis and Control*, Ed by D.-W. Sun. Academic Press, San Diego, pp. 3–43 (2010). doi: <https://doi.org/10.1016/B978-0-12-374753-2.10001-2>
2. M.O. Ngadi and L. Liu, "Hyperspectral image processing techniques", in *Hyperspectral Imaging for Food Quality Analysis and Control*, Ed by D.-W. Sun. Academic Press, San Diego, pp. 99–127 (2010). doi: <https://doi.org/10.1016/B978-0-12-374753-2.10004-8>
3. D. Wu and D.-W. Sun, "Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: a review-part II: Applications", *Innov. Food Sci. Emerg.* **19**, 15–28 (2013). doi: <https://doi.org/10.1016/j.ifset.2013.04.016>

4. M. Fauvel, Y. Tarabalka, J.A. Benediktsson, J. Chanussot and J.C. Tilton, "Advances in spectral-spatial classification of hyperspectral images", *Proc. IEEE*, **101(3)**, 652 (2013). doi: <https://doi.org/10.1109/JPROC.2012.2197589>
5. Y. Gao, J.F. Mas, N. Kerle and J.A.N. Pacheco, "Optimal region growing segmentation and its effect on classification accuracy", *Int. J. Remote Sens.* **32(13)**, 3747–3763 (2011). doi: <https://doi.org/10.1080/01431161003777189>
6. G. ElMasry, N. Wang, C. Vigneault, J. Qiao and A. ElSayed, "Early detection of apple bruises on different background colors using hyperspectral imaging", *LWT-Food Sci. Techn.* **41(2)**, 337–345 (2008). doi: <https://doi.org/10.1016/j.lwt.2007.02.022>
7. P. Baranowski, W. Mazurek and J. Pastuszka-Woźniak, "Supervised classification of bruised apples on the basis of hyperspectral imaging data", *Postharvest Biol. Tech.* **86**, 249–258 (2012). doi: <https://doi.org/10.1016/j.postharvbio.2013.07.005>
8. Q. Lü, M.J. Tang, J.R. Cai, J.W. Zhao and S. Vittayapadung, "Vis/NIR hyperspectral imaging for detection of hidden bruises on kiwifruits", *Czech J. Food Sci.* **29(6)**, 595 (2011).
9. S. Chevallier, D. Bertrand, A. Kohler and P. Courcoux, "Application of PLS-DA in multivariate image analysis", *J. Chemometr.* **20(5)**, 221–229 (2006). doi: <https://doi.org/10.1002/cem.994>
10. E. Gaston, J.M. Frias, P.J. Cullen, C. O'Donnell and A. Gowen, "Visible near-infrared hyperspectral imaging for the identification and discrimination of brown blotch disease on mushroom (*Agaricus bisporus*) caps", *J. Near Infrared Spectrosc.* **18(5)**, 341–353 (2010). doi: <https://doi.org/10.1255/jnirs.894>
11. R. Moschetti, W. Saeys, J.C. Kereztes, M. Goodarzi, M. Cecchini, M. Danilo and R. Massantini, "Hazelnut quality sorting using high dynamic range short-wave infrared hyperspectral imaging", *Food Bioproc. Technol.* **8(7)**, 1593–1604 (2015). doi: <https://doi.org/10.1007/s11947-015-1503-2>
12. N. Nguyen do Trong, M. Tsuta, B.M. Nicolai, J. De Baerdemaeker and W. Saeys, "Prediction of optimal cooking time for boiled potatoes by hyperspectral imaging", *J. Food Eng.* **105(4)**, 617–624 (2011). doi: <https://doi.org/10.1016/j.jfoodeng.2011.03.031>
13. J.C. Kereztes, M. Goodarzi and W. Saeys, "Real-time pixel based early apple bruise detection using short wave infrared hyperspectral imaging in combination with calibration and glare correction techniques", *Food Control* **66**, 215–226 (2016). doi: <https://doi.org/10.1016/j.foodcont.2016.02.007>
14. H. Lee, M.S. Kim, D. Jeong, S.R. Delwiche, K. Chao and B.K. Cho, "Detection of cracks on tomatoes using a hyperspectral near-infrared reflectance imaging system", *Sensors* **14(10)**, 18837–18850 (2014). doi: <https://doi.org/10.3390/s141018837>
15. F. Zhu, D. Zhang, Y. He, F. Liu and D.-W. Sun, "Application of visible and near infrared hyperspectral imaging to differentiate between fresh and frozen-thawed fish fillets", *Food Bioproc. Technol.* **6(10)**, 2931–2937 (2013). doi: <https://doi.org/10.1007/s11947-012-0825-6>
16. C. Portalés and E. Ribes-Gómez, "An image-based system to preliminary assess the quality of grape harvest batches on arrival at the winery", *Comput. Ind.* **68**, 105–115 (2015). doi: <https://doi.org/10.1016/j.compind.2014.12.010>
17. R.O. Duda, P.E. Hart and D.G. Stork, *Pattern Classification, 2nd Edition*. John Wiley & Sons, Inc., New York, USA (2001).
18. X. Zhu and A.B. Goldberg, *Introduction to Semi-Supervised Learning*. Morgan & Claypool Publishers, San Rafael, USA (2009). doi: <https://doi.org/10.2200/S00196ED1V01Y200906AIM006>
19. M. Marconcini, G. Camps-Valls and L. Bruzzone, "A composite semisupervised SVM for classification of hyperspectral images", *IEEE Geosci. Remote Sens. Lett.* **6(2)**, 234 (2009). doi: <https://doi.org/10.1109/LGRS.2008.2009324>
20. M. Chi and L. Bruzzone, "Semi-supervised classification of hyperspectral images by SVMs optimized in the primal", *IEEE Trans. Geosci. Remote Sens.* **45(6)**, 1870 (2007). doi: <https://doi.org/10.1109/TGRS.2007.894550>
21. J. Li, J.M. Bioucas-Dias and A. Plaza, "Semi-supervised hyperspectral image classification using soft sparse multinomial logistic regression", *IEEE Geosci. Remote Sens. Lett.* **10(2)**, 318 (2013). doi: <https://doi.org/10.1109/LGRS.2012.2205216>
22. G. Jun and J. Ghosh, "Semi-supervised learning of hyperspectral data with unknown land-cover classes", *IEEE Trans. Geosci. Remote Sens.* **51(1)**, 273 (2013). doi: <https://doi.org/10.1109/TGRS.2012.2198654>

23. Å. Rinnan, F. van den Berg and S.B. Engelsen, "Review of the most common pre-processing techniques for near-infrared spectra", *Trends Anal. Chem.* **28(10)**, 1201–1222 (2009). doi: <https://doi.org/10.1016/j.trac.2009.07.007>
24. A. Savitzky and M.J.E. Golay, "Smoothing and differentiation of data by simplified least squares procedures", *Anal. Chem.* **36(8)**, 1627–1639 (1964). doi: <https://doi.org/10.1021/ac60214a047>
25. R.J. Barnes, M.S. Dhanoa and S.J. Lister, "Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra", *Appl. Spectrosc.* **43(5)**, 772–777 (1989). doi: <https://doi.org/10.1366/0003702894202201>
26. S. Sharma, *Applied Multivariate Techniques*. John Wiley & Sons, Inc., New York, USA (1996).
27. D. Arthur and S. Vassilvitskii, "K-means++: the advantages of carefully seeding", *SODA '07: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, 1027 (2007).
28. S.P. Lloyd, "Least squares quantization in PCM", *IEEE Trans. Inform. Theory* **28**, 129 (1982). doi: <https://doi.org/10.1109/TIT.1982.1056489>
29. C.B. Do and S. Batzoglou, "What is the expectation maximization algorithm?" *Nature Biotechnol.* **26(8)**, 897–899 (2006). doi: <https://doi.org/10.1038/nbt1406>
30. W. Li, S. Prasad, J.E. Fowler and L.M. Bruce, "Locality-preserving dimensionality reduction and classification for hyperspectral image analysis", *IEEE Trans. Geosci. Remote Sens.* **50(4)**, 1185 (2012). doi: <https://doi.org/10.1109/TGRS.2011.2165957>
31. R.W. Kennard and L.A. Stone, "Computer aided design of experiments", *Technometrics* **11(1)**, 137–148 (1969). doi: <https://doi.org/10.2307/1266770>
32. E.W. Weisstein, "Convex hull", *MathWorld—A Wolfram Web Resource*. <http://mathworld.wolfram.com/ConvexHull.html> [Accessed 19 September 2016].
33. C. Albano, W.J. Dunn, U. Edlund, E. Johansson, B. Norden, M. Sjoström and S. Wold, "Four levels of pattern recognition", *Anal. Chim. Acta* **103(4)**, 429–443 (1978). doi: [https://doi.org/10.1016/S0003-2670\(01\)83107-X](https://doi.org/10.1016/S0003-2670(01)83107-X)
34. P.J. Brown, T. Fearn and M.S. Haque, "Discrimination with many variables", *J. Amer. Stat. Assoc.* **94(448)**, 1320–1329 (1999). doi: <https://doi.org/10.1080/01621459.1999.10473884>
35. R.G. Congalton, "A review of assessing the accuracy of classifications of remotely sensed data", *Remote Sens. Environ.* **37(1)**, 35–46 (1991). doi: [https://doi.org/10.1016/0034-4257\(91\)90048-B](https://doi.org/10.1016/0034-4257(91)90048-B)
36. G.W. Zack, W.E. Rogers and S.A. Latt, "Automatic measurement of sister chromatid exchange frequency", *J. Histochem. Cytochem.* **25(7)**, 741–753 (1977). doi: <https://doi.org/10.1177/25.7.70454>
37. C. Wallays, B. Missotten, J. De Baerdemaeker and W. Saeys, "Hyperspectral waveband selection for on-line measurement of grain cleanness", *Biosyst. Eng.* **104**, 1–7 (2009). doi: <https://doi.org/10.1016/j.biosystemseng.2009.05.011>
38. A.E. Watada, K.H. Norris, J.T. Worthington and D.R. Massie, "Estimation of chlorophyll and carotenoid contents of whole tomato by light absorbance technique", *Appl. Sci. Eng.* **41(2)**, 329–332 (1976). doi: <https://doi.org/10.1111/j.1365-2621.1976.tb00611.x>
39. G. Polder, G.W.A.M. van der Heijden, H. van der Voet and I.T. Young, "Measuring surface distribution of carotenes and chlorophyll in ripening tomatoes using imaging spectroscopy", *Postharvest Biol. Tech.* **34(2)**, 117–129 (2004). doi: <https://doi.org/10.1016/j.postharvbio.2004.05.002>
40. B. Farneti, R.E. Schouten and E.J. Woltering, "Low temperature-induced lycopene degradation in red ripe tomato evaluated by remittance spectroscopy", *Postharvest Biol. Tech.* **73**, 22–27 (2012). doi: <https://doi.org/10.1016/j.postharvbio.2012.05.008>
41. R. Arias, T.-C. Lee, L. Logendra and H. Janes, "Correlation of lycopene measured by HPLC with the L*, a*, b* color readings of a hydroponic tomato and the relationship of maturity with color and lycopene content", *J. Agr. Food Chem.* **48(5)**, 1697–1702 (2000). doi: <https://doi.org/10.1021/jf990974e>
42. J. Xing and J. De Baerdemaeker, "Bruise detection on 'Jonagold' apples using hyperspectral imaging", *Postharvest Biol. Tech.* **37(2)**, 152–162 (2005). doi: <https://doi.org/10.1016/j.postharvbio.2005.02.015>
43. A. Siedliska, P. Baranowski and W. Mazurek, "Classification models of bruise and cultivar detection on the basis of hyperspectral imaging data", *Comput. Electron. Agr.* **106**, 66–74 (2014). doi: <https://doi.org/10.1016/j.compag.2014.05.012>
44. G. ElMasry, A. Iqbal and D.-W. Sun, "Quality classification of cooked, sliced turkey hams using NIR hyperspectral imaging system", *J. Food Eng.* **103(3)**,

- 333–344 (2011). doi: <https://doi.org/10.1016/j.jfoodeng.2010.10.031>
45. C. Riccioli, D. Pérez-Marín, J.E. Guerrero-Ginel, W. Saeyns and A. Garrido-Varo, “Pixel selection for near-infrared chemical imaging (NIR-CI) discrimination between fish and terrestrial animal species in animal protein by-product meals”, *Appl. Spectrosc.* **65(7)**, 771–781 (2011). doi: <https://doi.org/10.1366/10-06177>
46. V. Todorov and P. Filzmoser, “Software tools for robust analysis of high-dimensional data”, *Austrian J. Stat.* **43(3–4)**, 255–266 (2014). doi: <https://doi.org/10.17713/ajs.v43i4.44>