INTUITIVE, SEMI-SUPERVISED TRAINING OF THE SEGMENTATION OF HYPERSPECTRAL IMAGES

J. van Roy¹*, N. Wouters¹, B. De Ketelaere¹, W. Saeyts¹

¹ KU Leuven, Department Biosystems, Division MeBioS, Kasteelpark Arenberg 30 bus 2456, 3000 Leuven, Belgium

Corresponding author: jeroen.vanroy@kuleuven.be

Hyperspectral imaging has a large potential for automated quality inspection in the food industry. However, the adoption of this technology is hampered by the time needed for building calibration algorithms by experts each time a new product (species, cultivar,..) has to be handled. This hinders a flexible change between products which most food companies desire. As a result, many of them prefer to rely on a subjective human inspection. Accordingly, this research focuses on the development of a semi-supervised training method allowing an operator without knowledge of hyperspectral imaging to train the system in an easy and intuitive way to recognize objects in a hypercube.

The developed algorithm works in three steps: The first step concerns unsupervised segmentation of a training image. The second step is the building of a supervised segmentation model based on the spectra selected from the groups in the first step. To conclude, a training phase is applied to improve the segmentation results. This method has been tested on hyperspectral data of two cultivars of vine tomatoes with a variable ripeness. The achieved results are very promising. Only 2 hypercubes in the training phase of the algorithm were needed to correctly segment 97% of the pixels in the hypercubes of 5 new vine tomatoes. The average used time to segment the images was reduced with a factor 17.5 compared to applying the unsupervised classification methods on each individual hypercube, while the training phase was intuitive and easy as no manual labelling of the images was needed.

KEYWORDS: Hyperspectral imaging, segmentation, machine learning, vine tomatoes

INTRODUCTION

In food companies, quality monitoring is of very high importance. Hyperspectral imaging can be an important tool in automating this process. As it combines spectral and spatial information, it is very useful to look for quality properties that are not visible with normal (RGB) machine vision¹. A good example of this are defects like scars or pinholes in vine tomatoes. However, to be able to correctly use the hyperspectral data, the hypercubes should be segmented into objects before quality analysis can be performed on individual objects. In case of tomatoes, tomato flesh, stalk and background should be classified correctly before the quality can be accurately determined.

Two approaches have been proposed for this: in unsupervised segmentation, the algorithm clusters the pixels in groups with similar spectra². As this method is not supervised, it might find groups which are the result of spectral variation which is not related to the objects of interest. Moreover, these methods are very computationally intensive as they have to evaluate the similarity (distance) for all possible combinations of pixels. This makes these methods unsuitable for in-line use. To avoid these issues, supervised methods have been proposed². However, these methods require a representative training set with pixel spectra for which the corresponding object class has been defined. Such a training set is typically created by letting an expensive human expert manually indicate which part of the hypercube belongs to which object class. This makes this method very labor intensive and prone to human error as the contrast between the different objects in the images on which the human expert manually labels the pixels may be rather low. To tackle these disadvantages a new semi-supervised algorithm is proposed which combines the ease of use of the unsupervised segmentation algorithms with the focus and computational efficiency of the supervised segmentation algorithms. The goal of the algorithm is to build segmentation models capable of dividing a hypercube in its different object classes without the need for expensive human experts and to be fast enough to be applicable in-line on a sorting line.
MATERIALS AND METHODS

1. Classification algorithm

The developed classification algorithm is schematically illustrated in Figure 1. As a case study the segmentation of vine tomatoes in background, stalk and tomato flesh is chosen. The first step is the measurement of vine tomatoes using a Vis-NIR hyperspectral imager (a). Next, the user decides how many different object classes are present in the image (b). This number is used for examining the performance of three different unsupervised segmentation algorithms, namely k-means clustering, hierarchical clustering and Gaussian Mixture Models (c). The best result of this unsupervised segmentation is then selected by the user (d). When the unsupervised segmentation algorithms perform poorly, according to the user, the unsupervised segmentation is searching for an extra object class. This is repeated until a satisfactory segmentation of the objects has been reached, where it is accepted that one object may be split in two different groups. The final classification model is then used to select the most interesting spectra for each object. In the case that an object had been split in two groups, these are both assigned to the same class. These spectra form the training set on which a supervised segmentation algorithm is trained (e). This supervised segmentation model is then applied to 5 independent measurements (f) of new samples and for every segmentation, the user decides if the segmentation is acceptable. If the segmentation was sufficiently accurate for all 5 test samples, the segmentation algorithm is implemented for in-line application (g). However, if the supervised segmentation was, according to the user, not accurate enough for one of the 5 test samples, the corresponding hypercube is reprocessed starting from the unsupervised segmentation algorithms and the correctly segmented hypercube for this test sample is used to augment the training set and build a new supervised classifier.

![Figure 1. Overview of the segmentation algorithm. The blue arrows indicate a bad unsupervised or supervised classification, while the green arrow indicates a good unsupervised segmentation.](image-url)

The first important step in the algorithm is the unsupervised classification of the hypercube. Three, often used, classification algorithms have been investigated. The first algorithm is k-means clustering applied on preprocessed spectra. It is an iterative method searching for the chosen number of cluster centers. Every spectrum is classified according to the nearest cluster center. Each time a new spectrum has been classified, the mean of each cluster is recalculated. This is repeated until there is no change anymore, or until 100 iterations have been performed. The squared Euclidean distance, applied on the original variables, has been used as distance measure. Four repetitions of the algorithm are conducted, using different cluster centers to start with, to reduce the chance of finding suboptimal results. The second classification algorithm investigated is the use of agglomerative hierarchical clustering. This technique assigns every spectrum to a different cluster and combines similar clusters until the wished number of clusters has been found. The method has been applied on the principal component scores which explain more than 1% of the present variance in the dataset. A method to minimize the within-cluster sums of squares, Ward’s method, has been used to decide on the similarity between clusters. Finally, multivariate Gaussian Mixture Models (GMM) based on the Expectation- Maximization algorithm were also used to segment the hypercubes. The Expectation-step (E-step) involves the estimation of a probability density...
function over the missing data, in this case the cluster where a pixel belongs to. The Maximization-step (M-step) consists of a re-estimation of the model parameters by using these results. This is repeated until convergence of the results. A Gaussian mixture distribution has been fitted to the data. The efficiency of these unsupervised techniques has been compared. A supervisor decided if the segmentation of each hypercube was accurate based on a comparison between the segmented image and an RGB image, which was simulated based on the hyperspectral images.

The next step in the algorithm is a classification of new data in a supervised way which requires a training set covering the relevant spectral variation. As hypercubes contain many pixel spectra which are very similar, it is better to select only a subset of the spectra for the training set. In this research, five spectra per object class were selected using the Kennard-Stone selection algorithm with Euclidean distance calculated from the scores values for the first 3 principal components of the data. Prior to applying the Kennard-Stone algorithm, the 1% most extreme spectra (largest Euclidean distance from the median spectrum for each object class) were removed to avoid that erroneous spectra (outliers) would dominate the training set. Three different supervised classification techniques were evaluated: The first is Support Vector Machines (SVM) based on the principal components of the data that explain at least 1% of the total present variation in the training set. A binary SVM classifier using a Gaussian kernel was built for each object class. The second technique investigated is Partial Least Squares-Discriminant Analysis (PLS-DA), which is a technique based on PLS regression to classify groups in a dataset. Three latent variables were used. Lastly, Soft Independent Modeling of Class Analogy (SIMCA) was investigated. This technique conducts a PCA analysis on each object class and keeps the first 3 principal components, describing more than 95% of the present variance in each class. Afterwards, when a new spectrum is shown to the algorithm, the algorithm will search for the smallest distance between the new spectrum and each object class. These techniques were compared based on the acquired accuracy and the time needed to classify every pixel in the hypercube. The achieved accuracy is quantified by comparing the result of the supervised segmentation with the result of a correct unsupervised segmentation of the same hypercube, the total number of correctly segmented pixels is divided by the total number of pixels in the image to calculate the accuracy.

The last step is the evaluation phase. In this phase, a supervisor should decide based on the results of the selected segmentation model on 5 test hypercubes if the segmentation is acceptable. When this segmentation was not acceptable, the corresponding hypercube was processed using the complete algorithm, starting from the unsupervised segmentation. The spectra selected after this unsupervised segmentation are added to the existing training set. To investigate the added value of the evaluation phase in the algorithm, the effect of using two training images in the training set on classification accuracy and time needed to segment a hypercube of vine tomatoes is compared with the use of only one training image.

2. Samples

During this research, vine tomatoes (Solanum lycopersicum) were used to test the efficiency of the algorithm. In June 2014, 36 vines of the cultivar Merice were harvested at ‘Proefstation voor de groenteteelt’ (Sint-Katelijne-Waver, Belgium), and in August 2014, 36 vines of the cultivar Prunus were harvested at ‘Proefcentrum Hoogstraten’ (Hoogstraten, Belgium). Every harvest period, tomatoes in three ripeness stages were selected, namely unripe, commercial harvest and overripe. The tomatoes of every ripeness stage were divided in four groups and every group was measured at a different day, to obtain a dataset covering all relevant 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. The tomatoes were stored in a fridge at a constant temperature of 18 °C and a relative humidity of 80%. Every truss of tomatoes was measured once.
3. Hyperspectral imaging

In this research, a line-scan hyperspectral imaging system (Figure 2) was used. A linear translation stage (Franke GmbH, Aalen, Germany) moves the samples under the line scan hyperspectral camera system consisting of a VisNIR lens with a focal length of 17 mm (CM120 BK15 COMPACT, Schneider Kreuznach, Bad Kreuznach, Germany), a spectrograph (Inspector V10, Spectral Imaging Ltd., Oulu, Finland) and a 12 bit monochrome CCD camera (TXG14NIR, Baumer, Frauenfeld, Switzerland). The camera is sensitive for light with a wavelength between 400 nm and 1000 nm. The illumination consists of 4 DC-halogen lamps (Decostar Alu 35, OSRAM, Munich, Germany) with a diffuser (TSG-LEGB, Knight Optical Ltd., Kent, United Kingdom) mounted in front and placed in an arc-shape around the translation stage to obtain a uniform illumination.

A PC was used to control the linear translation stage and the camera (Labview version 9.0, National Instruments Corporation, Austin, TX, USA). The exposure time was optimized for tomatoes at 35 ms. White and dark references were measured every day. The relative reflectance in a pixel was calculated as described by Wallays et al. (2009). With every step of the translation stage, the camera takes one image, containing spatial and spectral information in the field-of-view of the camera. These images are combined using Matlab forming a hypercube containing as well spatial as spectral information. To reduce the size of the data, the images were binned by a factor 4, resulting in a spatial resolution of 0.8 mm and a spectral resolution of 2.56 nm. All the spectra were preprocessed using a Savitzky-Golay filtering with a window width of 15 points (38.4 nm), a second order polynomial and a first derivative. The developed algorithm was elaborated in Matlab (version 8.4.0, The Mathworks Inc., Natick, MI, USA) using the statistics and machine learning toolbox and combined with the PLS-Toolbox (version 8.0, Eigenvector Research Inc., Manson, WA, USA) on a desktop PC (Intel® Core™ i5, CPU: 2.8GHz, RAM: 6GB).

RESULTS AND DISCUSSION

In Table 1, the results of the unsupervised segmentation are summarized. K-means clustering achieved the best results, providing a correct segmentation of 83.33% of the investigated hypercubes (from the entire data set in background, stalk and tomato flesh). Hierarchical clustering and multivariate Gaussian mixture models achieved considerably lower correct classification rates. However, they showed added value to the k-means clustering as the number of correctly segmented images increases if all three methods are used compared to only k-means clustering.

Table 1. Results of the three unsupervised segmentation techniques on the vine tomato data.

<table>
<thead>
<tr>
<th></th>
<th>k-means clustering</th>
<th>Hierarchical clustering</th>
<th>GMM</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Result (%)</td>
<td>83.33</td>
<td>55.56</td>
<td>38.89</td>
<td>86.11</td>
</tr>
</tbody>
</table>

The results achieved for the unsupervised segmentation methods were used to select the 5 most informative pixel spectra for each object class in the image based on the Kennard-Stone selection algorithm. These spectra were then used to build a supervised model that was validated on 5 different hypercubes. The results are summarized in Table 2. It is clear that, when using only one training image, PLS-DA gives the best accuracy, while
using the least amount of time to process one hypercube. When adding a second training image, PLS-DA still gives the best result, with a correct classification rate of 97 % at pixel level. This indicates that the algorithm is able to cope with tomatoes of very different ripeness and of different cultivars by building only one supervised model, showing the flexibility of the algorithm. As it only needs 0.52 s to process one hypercube in Matlab, it is fast enough for use in in-line quality sorting of vine tomatoes.

**Conclusion**

A semi-supervised segmentation algorithm has been elaborated to classify the pixels in a hyperspectral image into the different object classes. To illustrate the performance of this algorithm, a dataset of vine tomatoes was recorded, and the performance of the algorithm in segmenting the hypercubes in background, stalk and tomato flesh was investigated. When applying the algorithm to this dataset a correct classification rate of 97 % was achieved. The time needed to segment one hypercube is 0.52 s, which makes it very suitable for use in a sorting line where products should be sorted based on quality.

**Acknowledgments**

Jeroen van Roy is a PhD student 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. The research leading to these results has received funding from the European Union’s Seventh Framework Programme for research, technological development and demonstration under grant agreement no 311987 (PicknPack). The authors gratefully acknowledge the support of the Flanders

**References**

5. Do, C.B.; Batzoglou, S.; Nature Biotechnology, 2006, 26(8), 897-899
10. Savitzky, A.; Golay, M.J.E.; Analytical Chemistry, 1964, 36(8), 1627-1639

<table>
<thead>
<tr>
<th>1 Training image</th>
<th>2 Training images</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Accuracy</strong></td>
<td><strong>Time (s)</strong></td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td><strong>Time (s)</strong></td>
</tr>
<tr>
<td>SVM</td>
<td>0.89</td>
</tr>
<tr>
<td>PLS-DA</td>
<td>0.92</td>
</tr>
<tr>
<td>SIMCA</td>
<td>0.82</td>
</tr>
</tbody>
</table>

DOI: 10.17648/nir-2015-34186