



The Society for engineering  
in agricultural, food, and  
biological systems

TECHNICAL LIBRARY

---

# Spectral Image Analysis for Measuring Ripeness of Tomatoes

G. Polder, G. W. A. M. van der Heijden, I. T. Young

---

Published in Transactions of the ASAE Vol. 45(4): 1155-1161 ( ? ? 2002 American Society of Agricultural Engineers ).

Article was submitted for review in September 2000; approved for publication by the Information & Electrical Technologies Division of ASAE in March 2002. Presented at the 2000 ASAE Annual Meeting as Paper No. 003089.

The authors are **Gerrit Polder** , Researcher, and **Gerie W. A. M. Van der Heijden** , Senior Researcher, Plant Research International, Wageningen, The Netherlands; and **I. T. Young** , Professor, Delft University of Technology, Delft, The Netherlands. **Corresponding author** : Gerrit Polder, Plant Research International, Postbus 16, 6700 AA Wageningen, Netherlands; phone: +31-317-476842; fax: +31-317-418094; e-mail: g.polder@plant.wag-ur.nl.

---

**Abstract .** *In this study, spectral images of five ripeness stages of tomatoes have been recorded and analyzed. The electromagnetic spectrum between 396 and 736 nm was recorded in 257 bands (every 1.3 nm). Results show that spectral images offer more discriminating power than standard RGB images for measuring ripeness stages of tomatoes. The classification error of individual pixels was reduced from 51% to 19%. Using a gray reference, the reflectance can be made invariant to the light source and even object geometry, which makes it possible to have comparable classification results over a large range of illumination conditions. Experimental results show that, although the error rate increases from 19% to 35% when using different light sources, it is still considerably below the 51% for RGB under a single light source.*

**Keywords.** Linear discriminant analysis, Color constancy, Imaging spectroscopy, Machine vision, Tomatoes, Ripeness, Multispectral, Hyperspectral.

Tomatoes, with an annual production of 60 million tons, are one of the main horticultural crops in the world, with three million hectares planted every year. In the Netherlands, the total area under cultivation was 1133 hectares in 2000 with a production of 1040 million kilograms and a mean price of \$0.90 per kilogram (van Dijk et al., 2001).

Traditionally, the surface color of tomatoes is a major factor in determining the ripeness of tomato fruits (Arias et al., 2000). A color-chart standard has been specifically developed for the purpose of classifying tomatoes in 12 ripeness classes (The Greenery, Breda, the Netherlands). For automatic sorting of tomatoes, RGB color cameras are used instead of the color chart (Choi et al., 1995). RGB-based classification, however, strongly depends on recording conditions. Next to surface and reflection/absorption characteristics of the tomato itself, the light source (illumination intensity, direction, and spectral power distribution), the characteristics of the filters, the settings of the camera (e.g., aperture), and the viewing position all influence the final RGB image. Van der Heijden et al. (2000) has shown that color information in hyperspectral images can be made invariant to these factors, thus providing a powerful alternative to RGB color cameras. In this way, a hyperspectral imaging system and spectral analysis would permit the sorting of tomatoes under different lighting conditions.

In this study, we want to compare hyperspectral images with standard RGB images for classifying tomatoes in different ripeness classes using individual pixels. Spectral images have been captured under different lighting conditions. By including a gray reference in each image, automatic compensation for different light sources has been obtained.

For comparison and classification, Fisher's linear discriminant analysis (LDA) is used. Van den Broek et al. (1997) showed that this technique is very suitable for classification of spectroscopic images.

The article is organized as follows. First, imaging spectrometry and hyperspectral images are described. Next, we show how hyperspectral images can be transformed into images invariant to the light source and the object geometry. An experiment that compares the two imaging modalities will then be described and the results given.

## Imaging Spectrometry

While a gray-value image typically reflects the light intensity over a part of the electromagnetic spectrum in a single band, and a color image reflects the intensity over the red, green, and blue parts of the spectrum in three bands, increasing the number of bands can greatly increase the amount of information in an image. Hyperspectral images commonly contain about 100 to 300 bands with a resolution of 1 to 10 nm. Current techniques offer two basic approaches to spectral imaging. It is implemented by acquiring either a sequence of two-dimensional images at different wavelengths or a sequence of line images in which a complete spectrum is captured for each pixel on the line. The first approach is implemented by employing a rotating filter wheel or a tunable filter in front of a monochrome camera. This approach is preferable if the number of bands needed is limited and the object can be held still in front of the camera during recording. The second approach requires an imaging spectrograph coupled to a monochrome area camera. One dimension of the camera (spatial axis) records the line pixels, and the other dimension (spectral axis) records the spectral information for each pixel. This approach is well suited in a conveyor belt system, using the camera as a line-scan camera.

In this experiment, we used the ImSpector (Spectral Imaging Ltd., Oulu, Finland) straight-axis imaging spectrograph, which uses a prism-grating-prism (PGP) dispersive element and transmission optics

(Herrala and Okkonen, 1996; Hyvonen et al., 1998).

To record a spatial two-dimensional image, the object was moved perpendicular to the optical axis of the camera using a linear translation table.

## Color Invariance

The spectra obtained by the imaging spectrograph depend on the light source and object characteristics. Therefore, these spectra may vary with a change in the intensity and energy distribution of the light source, material characteristics, and viewing mode. The aim of this section is to propose spectra that are invariant to illumination (color constancy) and the object geometry and shadows (normalization). We will follow the method proposed by Stokman and Gevers (1999).

The section is outlined as follows. First, the reflection is given, modeling interaction between light and matter. Then color-constant spectra are presented. Finally, the color-constant spectra are made independent of object geometry and shading using normalization.

## The Reflection Model

Let  $E(\vec{x}, \lambda)$  be the spectral power distribution of the incident (ambient) light at the object surface at  $\vec{x}$  and let  $L(\vec{x}, \lambda)$  be the spectral reflectance function of the object at  $\vec{x}$ . The spectral sensitivity of the  $k$ th sensor is given by  $F_k(\lambda)$ . Then  $\phi_k$ , the sensor response of the  $k$ th channel, is given by:

$$\phi_k(\vec{x}) = \int_{\lambda} E(\vec{x}, \lambda) L(\vec{x}, \lambda) F_k(\lambda) d\lambda \quad (1)$$

where

$\lambda$  = wavelength

$L(\vec{x}, \lambda)$  = a complicated function based on the geometric and spectral properties of the object surface.

The integral is taken over the relevant part of the electromagnetic spectrum (i.e., 380 to 700 nm).

Further, consider an opaque inhomogeneous dielectric object. The geometric and surface reflection component of function  $L(\vec{x}, \lambda)$  can then be decomposed in a body and surface reflection component as described by Shafer (1985):

$$\phi_k(\vec{x}) = G_B(\vec{x}, \vec{n}, \vec{s}) \int_{\lambda} E(\vec{x}, \lambda) B(\vec{x}, \lambda) F_k(\lambda) d\lambda + G_S(\vec{x}, \vec{n}, \vec{s}, \vec{v}) \int_{\lambda} E(\vec{x}, \lambda) S(\vec{x}, \lambda) F_k(\lambda) d\lambda \quad (2)$$

giving the  $k$ th sensor response. Further,  $B(\vec{x}, \lambda)$  and  $S(\vec{x}, \lambda)$  are the surface albedo and Fresnel

reflectance, respectively, at  $\vec{x}$ ,  $\vec{n}$  is the surface patch normal,  $\vec{s}$  is the direction of the illumination source, and  $\vec{v}$  is the direction of the viewer. Geometric terms  $G_B$  and  $G_S$  denote the geometric dependencies on the body and surface reflection component, respectively, independent of the wavelength. Matte surfaces are described by the body reflection alone; shiny surfaces are described by both the body and surface reflection.

## Color Constancy

Consider the reflectance of a perfect reflecting diffuser. A sample is called perfect when the sample reflects all wavelengths of the light source with an equal amount in all directions, producing an isotropic diffuser. This means that there is no surface reflection,  $G_S(\vec{x}, \vec{n}, \vec{s}, \vec{v}) = 0$ . Further, we assume that the diffuser has constant spectral albedo reflectance  $B(\vec{x}, \lambda) = R$ . Assuming diffuse reflectance and that the surface normal  $\vec{n}$  is the same as the viewing direction  $\vec{v}$ , this gives  $G_B(\vec{x}, \vec{n}, \vec{s}) = 1$ . Then the sensor response of the  $k$ th channel of the reference sample (assuming no specular reflections) is given by:

$$\phi_k(\vec{r}) = R \int_{\lambda} E(\vec{r}, \lambda) F_k(\lambda) d\lambda \quad (3)$$

The relative spectral power distribution of the reference sample is measured in this way.

Further, assume that an image of an arbitrary sample is captured under the same illumination conditions. Then the relative reflection of the arbitrary sample with respect to the reference sample is:

$$\frac{\phi_k(\vec{x})}{\phi_k(\vec{r})} = \frac{G_B(\vec{x}, \vec{n}, \vec{s}) \int_{\lambda} E(\vec{x}, \lambda) B(\vec{x}, \lambda) F_k(\lambda) d\lambda}{R \int_{\lambda} E(\vec{r}, \lambda) F_k(\lambda) d\lambda} + \frac{G_S(\vec{x}, \vec{n}, \vec{s}, \vec{v}) \int_{\lambda} E(\vec{x}, \lambda) S(\vec{x}, \lambda) F_k(\lambda) d\lambda}{R \int_{\lambda} E(\vec{r}, \lambda) F_k(\lambda) d\lambda} \quad (4)$$

giving the  $k$ th sensor response of a sample with respect to the reference reflectance. As mentioned before, the spectrograph spectral filters  $F_k(\lambda)$  are narrow-band filters. At channel  $k$ , the wavelength dependency can, therefore, be considered as fixed and we can rewrite equation 4 as:

$$\frac{\phi_k(\vec{x})}{\phi_k(\vec{r})} = \frac{G_B(\vec{x}, \vec{n}, \vec{s}) E(\vec{x}, \lambda) B(\vec{x}, \lambda) F_k(\lambda_k)}{R E(\vec{r}, \lambda) F_k(\lambda_k)} + \frac{G_S(\vec{x}, \vec{n}, \vec{s}, \vec{v}) E(\vec{x}, \lambda) S(\vec{x}, \lambda) F_k(\lambda_k)}{R E(\vec{r}, \lambda) F_k(\lambda_k)} \quad (5)$$

If the spectral distribution of the light at  $\vec{x}$ ,  $E(\vec{x}, \lambda)$ , can be assumed identical to that at the reference sample,  $E(\vec{r}, \lambda)$ , then this can be simplified to:

$$\phi_k(\vec{x}) = \frac{1}{R} [G_B(\vec{x}, \vec{n}, \vec{s})B(\vec{x}, \lambda_k) + G_S(\vec{x}, \vec{n}, \vec{s}, \vec{v})S(\vec{x}, \lambda_k)] \quad (6)$$

Equation 6 implies that, under the assumption that at channel  $k$  the wavelength dependency can be considered as fixed (a unit impulse band filter), the spectral images obtained by the spectrograph can be made independent of the illuminant by dividing the original spectral image by the spectral radiation of the illuminant. The spectral radiation of the illuminant is determined by measuring the reflectance of a perfect isotropic diffuser. Note that the spectral power distribution of the light source is unimportant as long as it contains all relevant wavelengths at a sufficient intensity and is constant over the whole image.

The diffuser used in this article is patch 21 from the GretagMacBeth standard color checker (New Windsor, N.Y.) with a spectral reflectance of 0.36 from 300 to 900 nm. Before dividing the spectra by the spectral radiation of the illuminant, the dark current has been subtracted. The dark current was estimated by taking an image with a closed lens cap and all illumination set to zero.

## Normalization

In this section, the color-constant spectra are normalized, yielding spectra that are independent of the object geometry, under the assumption of matte surfaces (negligible Fresnel reflection).

Consider the body reflection term of equation 6:

$$D_k(\vec{x}) = G_B(\vec{x}, \vec{n}, \vec{s})B(\vec{x}, \lambda_k) \quad (7)$$

giving the  $k$  th sensor response of the spectral reflectance curve of a matte surface.

According to equation 7, the color depends only on surface albedo, and the brightness depends on factor  $G_B(\vec{x}, \vec{n}, \vec{s})$ . As a consequence, a uniformly painted surface may give rise to a broad variation in sensor values due to the varying circumstances induced by the image-forming process, such as a change in object geometry.

In contrast, a normalized sensor space can be insensitive to surface orientation, illumination direction, and intensity by dividing each channel by the sum of channels:

$$\begin{aligned} C_k(\vec{x}) &= \frac{D_k(\vec{x})}{D_1(\vec{x}) + D_2(\vec{x}) + \dots + D_N(\vec{x})} \\ &= \frac{G_B(\vec{x}, \vec{n}, \vec{s})B(\vec{x}, \lambda_k)}{G_B(\vec{x}, \vec{n}, \vec{s})B(\vec{x}, \lambda_1) + B(\vec{x}, \lambda_2) + \dots + B(\vec{x}, \lambda_N)} \\ &= \frac{B(\vec{x}, \lambda_k)}{B(\vec{x}, \lambda_1) + B(\vec{x}, \lambda_2) + \dots + B(\vec{x}, \lambda_N)} \end{aligned} \quad (8)$$

The result is taken over  $N$  wavelengths and only dependent on the surface albedo. It is independent of the illumination spectral power distribution, illumination direction  $\vec{s}$ , and object geometry  $\vec{n}$ .

# Linear Discriminant Analysis

Linear discriminant analysis (LDA) is a supervised classification technique. For each class in a training set, it calculates the means and covariance matrix of the features (in our case the spectral bands) and then assigns new objects (pixels) to the category with the smallest Mahalanobis distance to that object (Fukunaga, 1990; Ripley, 1996). The Mahalanobis distance between pixel  $x$  and class  $c$  is defined as:

$$D_M(x, c) = (\mu_c - x)^T G_c^{-1} (\mu_c - x) \quad (9)$$

where

$\mu_c$  = mean vector of class  $c$

$G_c$  = covariance matrix of class  $c$

$x$  = spectrum of a pixel belonging to class  $c$ .

We assume  $G_i = G_j$  for all  $i, j \in \{1 \dots c\}$ . To obtain the classes, the LDA model is calculated by applying the transformation  $y = W^T x$ . The weight matrix  $W$  is used to reduce the features  $x$  to the discriminant space  $y$ , where the ratio of the between-class scatter matrix of the projected samples to the within-class scatter matrix of the projected samples is maximized.

From these transformed training pixels, the class means  $\mu$  and corresponding covariance matrices  $G_c$  of each class are calculated. The class membership of the pixels in the test set can be predicted by calculating the Mahalanobis distances to each class. A vector in which integer-valued elements represent the membership of each pixel to a particular class expresses the classification result.

## Experiment: Image Recording of Tomatoes

Five tomatoes (Capita F1 from De Ruiter Seeds, Bergschenhoek, The Netherlands) in ripeness stage 7 (orange) were harvested. The ripeness stage was defined using a tomato color chart standard (The Greenery, Breda, The Netherlands), which is commonly used by breeders. Each day over a time period of 5 days, a color RGB image and spectral images were taken of the five fruits on a black velvet background.

The RGB images were recorded with a Sony DX-950P 3-CCD color camera. The light source consisted of fluorescent daylight tubes (6500 K) with a high-frequency (50 kHz) electronic ballast, which prevents flickering caused by interference of the camera frame rate and the AC mains frequency. The frame grabber used was a Matrox Meteor RGB (Dorval, Quebec).

The imaging spectrograph used in this experiment is the ImSpector (Spectral Imaging Ltd., Oulu, Finland). The ImSpector is available in several different wavelength ranges. Type V7, which was used, has a spectral range of 396 to 736 nm and a slit size of 13  $\mu$ m, resulting in a spectral resolution of 1.3 nm.

The spectral images were recorded using two Dolan-Jenner PL900 illuminators (Andover St. Lawrence, Mass.), with 150 W quartz halogen lamps. These lamps have a relatively smooth emission between 380

and 2000 nm. Glass fiber optic line arrays of 0.02 inch • 6 inch aperture and rod lenses for the line arrays (Vision Light Tech, Uden, The Netherlands) were used for illuminating the scene.

To test the effect of using different light sources for the training set and test set using color-constant and normalized images, an additional experiment was carried out with halogen illumination for the training set and three other illumination sources for the test set:

1. Halogen combined with a Schott KG3 filter in front of the camera lens.
2. Halogen with an additional TLD58W (Philips, The Netherlands) fluorescence tube.
3. Halogen with an additional blue fluorescence tube (Marine Blue Actinic, Arcadia, U.K.).

The camera used was a Qimaging PMI-1400 EC Peltier cooled camera with a NIKON 55 mm lens and the ImSpector V7 between the lens and the camera. The frame grabber used was a Datacell Limited Snapper board (Berkshire, U.K.). The translation table used to move the object with respect to the camera was a Lineartechnik Lt1-Sp5-C8-600 translation table (Andelst, The Netherlands) driven by a SDHWA 120 programmable microstepping motor driver (Ever Elettronica, Italy). The resolution of this translation table was  $\approx 30 \mu\text{m}$ , and the maximum speed 250 mm/s.

Full-size hyperspectral images are large. If the full spatial resolution of the camera (1320 • 1035 pixels) for the x-axis and spectral axis was used, and with 1320 pixels in the y-direction, a single hyperspectral image would be 3.6 GB (using 16 bits/pixel). Due to limitations in lens and ImSpector optics, such a spectral image is oversampled and binning can be used to reduce the size of the image without losing information (Polder and van der Heijden, 2001). The software allows binning of the image separately in both the spatial and the spectral axes during capture of the image. The binning factor in both the spatial x-axis and the spectral axis was 4. The step size of the stepper table was chosen to match the binned spatial resolution in the x-direction. The number of steps was chosen to capture five tomatoes and the gray reference in one image. The resulting images have a spatial dimension of 318 • 512 square pixels and a spectral dimension of 257 bands (about 84 MB).

The software to control the stepper table and frame grabber, to construct the hyperspectral images, and to save and display them was locally developed in a single computer program written in Java.

## Data Analysis

Data analysis was performed in the following steps. First, image preprocessing was performed as described in the section below. The preprocessed images were then analyzed with linear discriminant analysis (LDA). All analyses were done using Matlab (The Mathworks, Inc., Natick, Mass.) and the Matlab PRTools toolbox (Faculty of Applied Physics, Delft University of Technology, The Netherlands).

## Image Preprocessing

To separate the tomatoes from the background and process each tomato image as a separate object, a threshold was performed on the intensity image of the RGB and spectral images. The intensity image was calculated as the sum of the R, G, and B components of the RGB image and as the sum of the different bands for the spectral image. The binary image that was obtained was labeled, giving the locations of the

five fruits and the gray reference patch in each image.

The spectral images contain high-intensity regions caused by specular reflection of the illumination source at the tomato surface. These specular regions act like a mirror and predominantly show the spectral power distribution of the light source, distorting the measurements. These pixels were discarded by thresholding the intensity image. The threshold value applied was  $0.75 \cdot$  the maximum gray value of the image. This value was determined by dividing the gray values of the specular region by the mean gray value of the rest of the tomato surface. A labeled mask image was calculated with one label for each tomato object, discarding the specular regions. The tomatoes and the gray reference are large objects in the image, compared to noise pixels or small artifacts in the specular region of the tomatoes. Therefore, single pixels and small objects were removed by ignoring labeled objects with an object size less than 5000. Figure 1 shows the reflectance values of the five tomatoes and the reference at 600 nm, at the day of harvest.

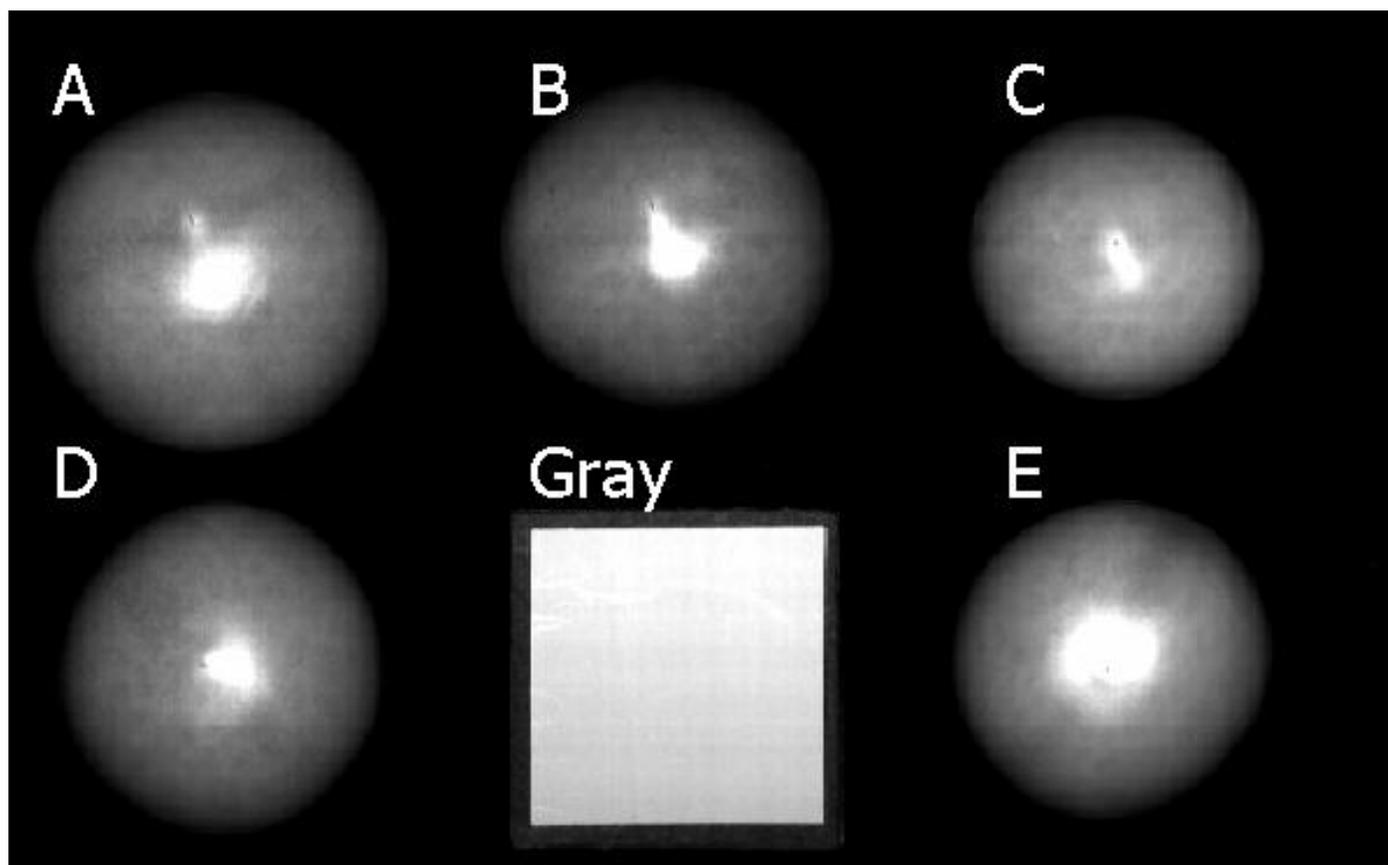


Figure 1. Image of the reflectance of the five tomatoes and the reference at 600 nm at the day of harvest.

The low radiation from the light source and the low sensitivity of the CCD camera combined with low reflectance values of tomatoes in the lower part of the spectrum (below 500 nm) produced a low signal-to-noise ratio (<30 dB) in this part of the spectrum. Therefore, the spectrum was reduced by discarding reflectance values below a certain threshold. The threshold value chosen was empirically determined at 2% of the maximum reflectance value of the mean spectrum of the first image of each time series. The remaining spectra consisted of 200 data values.

## Linear Discriminant Analysis

Each labeled tomato was considered as a separate ripeness stage. From the similarly labeled tomato objects, 2500 (pixel) spectra were randomly selected to form a learning set for the LDA. From the remaining spectra, 2500 spectra were randomly chosen to form the validation set. This experiment was repeated for all five tomatoes. In this experiment, we assume that all training pixels of a tomato image, captured at a certain day, belong to one ripeness class. To ensure that the training and the test data were exactly of the same ripeness stage, the pixels were randomly taken from the same tomatoes.

To compare the power of spectral images with that of RGB images for ripeness classification, LDA was performed on the RGB images, the original (raw) spectral images, the color-constant spectral images, and the normalized spectral images.

After learning the LDA of each spectrum or RGB triple, the test sets were classified.

## Results

Scatter plots were made of randomly selected test sets of 100 pixels, from the original RGB triplets (figure 2), the LDA mapping to two canonical variables for the RGB images (figure 3) and the LDA mapping to two canonical variables for the raw spectral images (figure 4) of one tomato (figure 1, tomato B). From these plots, it is clear that the RGB domain representations of the tomato show considerable overlap at the different time stages. The LDA mapping of the RGB images also show this overlap. The overlap in the LDA domain of the spectral images is considerably reduced. Due to the reduction of variation within classes, the classes in the LDA domain are much more concentrated.

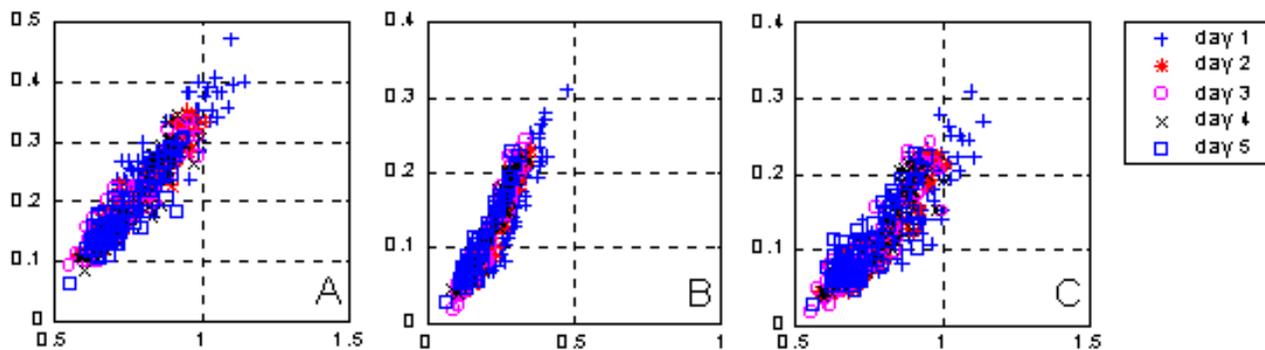


Figure 2. Scatter plot of the RGB points of the color images. Depicted are: (A) red versus green, (B) red versus blue, and (C) green versus blue. Classes 1 to 5 represent the ripeness stages of tomato B during the five days after harvest, respectively.

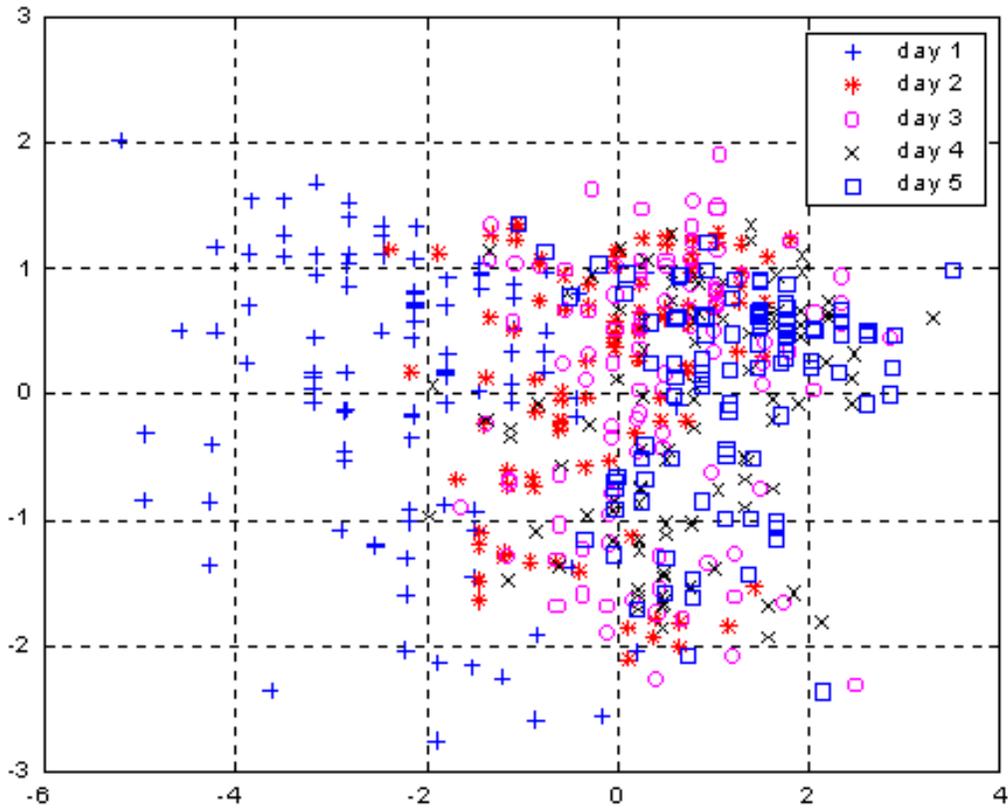


Figure 3. Scatter plot of the first and second canonical variables of the LDA analysis of the RGB images. Classes 1 to 5 represent the ripeness stages of tomato B during the five days after harvest, respectively.

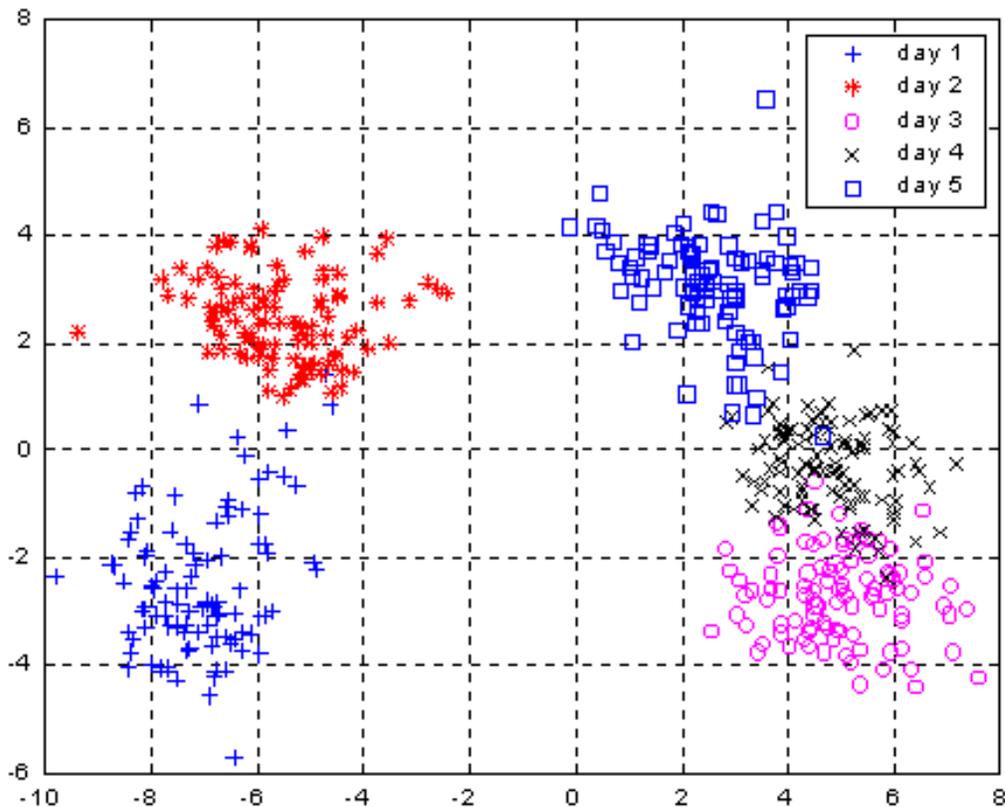


Figure 4. Scatter plot of the first and second canonical variables of the LDA analysis of the spectral images. Classes 1 to 5 represent the ripeness stages of tomato B during the five days after harvest, respectively.

The result of comparing the LDA classification of the ripeness stages of the RGB images against the actual stages for all five tomatoes is tabulated in table 1. The error rate is 51%. If one class difference is allowed (e.g., for class 2, classes 1, 2, and 3 are considered correct), then the error rate is 19%. The error rates for the individual tomatoes are tabulated in table 3. From this table, we see that the error rate varies from 48% to 56% with a standard deviation of 3.2%. When one class difference is allowed, these values are 15% to 24% with a standard deviation of 3.7%.

Table 1. Cross table of the number of pixels of the actual ripeness stage against the ripeness stage predicted by LDA on a test set of the RGB images (five tomatoes with 2500 pixels per class). Error rate = 51%.

LDA RGB	1	2	3	4	5	Total
Actual						
1	11985	406	59	29	21	12500
2	2545	5277	1732	1181	1765	12500
3	1245	4218	2090	1856	3091	12500

4	459	2159	1852	2600	5430	12500
5	170	668	952	2210	8500	12500
Total	16404	12728	6685	7876	18807	62500

Table 2 lists the result of the LDA classification of the raw spectral images. The error rate is 19%. If one class difference is allowed, then the error rate is reduced to 1.3%. The error rates for the individual tomatoes are tabulated in table 3. From this table, we learn that the error rate varies from 16% to 20% with a standard deviation of 1.7%. When one class difference is allowed, these values are 0.6% to 1.5% with a standard deviation of 0.4%.

Table 2. Cross table of the number of pixels of the actual ripeness stage against the ripeness stage determined with the LDA on a test set of the raw spectral images (five tomatoes with 2500 pixels per class). Error rate = 19%.

LDA Spectral Actual	1	2	3	4	5	Total
1	11737	743	20	0	0	12500
2	59	11531	811	70	29	12500
3	3	1306	9964	1094	133	12500
4	1	372	2447	5977	3703	12500
5	1	36	138	616	11709	12500
Total	11801	13988	13380	7757	15574	62500

Table 3. Error rates for individual tomatoes. Error rates for the case in which one class difference is allowed are shown in brackets.

Error Rates Tomato	RGB	Spectral
A	50 [17]	18 [1.5]
B	56 [24]	20 [1.5]
C	48 [15]	18 [1.4]
D	54 [23]	16 [1.5]
E	48 [16]	20 [0.6]
Mean	51 [19]	19 [1.3]

Std. Dev.	3.2 [3.7]	1.7 [0.4]
-----------	--------------	--------------

Tables 1 to 3 show the results for individual pixels; when we moved from pixel classification to object classification, only one tomato RGB image was misclassified. Object classification was performed by a simple majority vote (i.e., each object was assigned to the class with the highest frequency of assigned pixels). In the spectral images, no error was made. Table 4 shows the result.

Table 4. Cross table of the actual ripeness stage against the ripeness stage determined with the LDA on a test set of the raw spectral images and the RGB images (in brackets) by majority voting on the individual pixels.

LDA Spectral [RGB]	1	2	3	4	5
Actual					
1	5 [5]				
2		5 [5]			
3			5 [4]		
4			[1]	5 [5]	
5					5 [5]

When using different light sources for the training and test set, the reflectance of the reference is needed to make the spectra color-constant. Figure 5 shows the four reference spectra obtained by the reflectance of the gray reference. The intensity of the TLD58W source and the Arcadia Blue sources were relatively low compared to the halogen intensity. The greatest difference is at the emission peaks of these fluorescence tubes. Since wavelengths with low signal-to-noise ratios were excluded from the analysis, only the spectra between 487 and 736 nm were used. Because of this, the effect of the blue fluorescence of the Arcadia lamp was minimal.

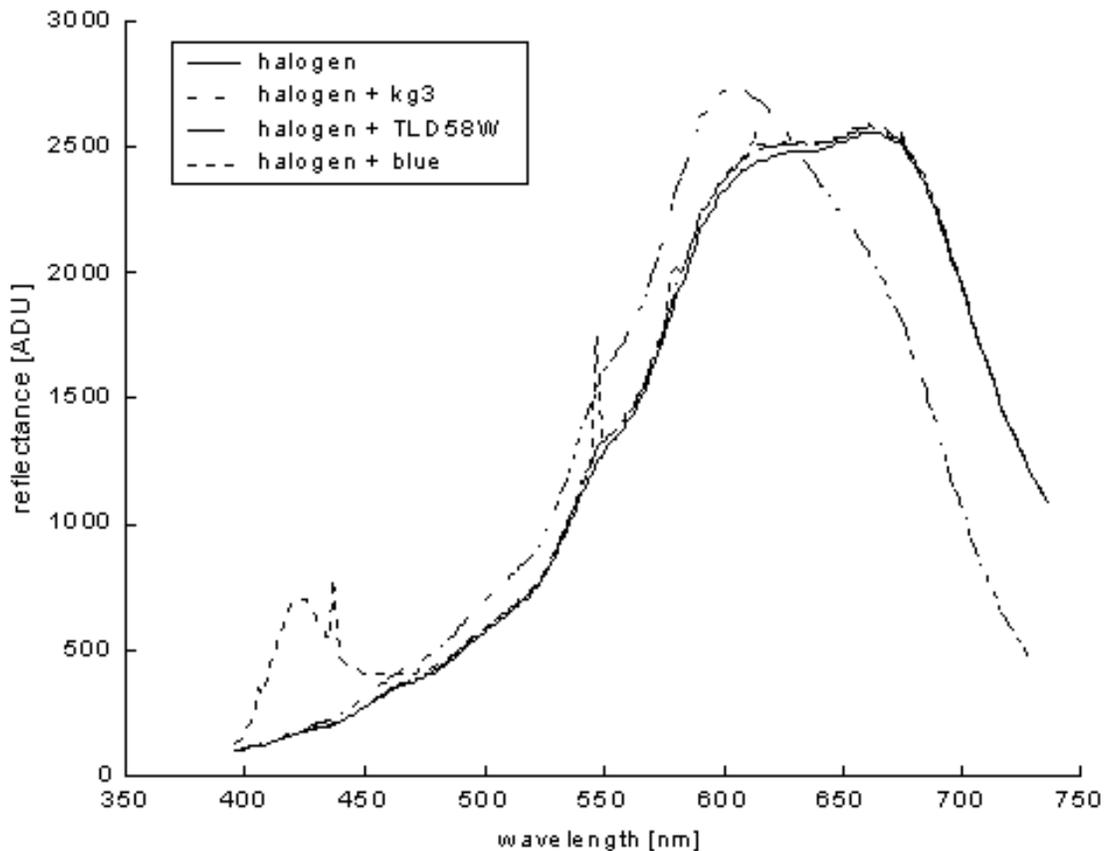


Figure 5. Spectral reflectance of the gray reference using different illumination sources.

## Discussion

An RGB color camera is frequently used for ripeness sorting. However, the results show that considerable errors may occur when classifying small differences in ripeness using RGB images. Spectral images are more suitable for classifying ripeness because they have a higher discriminating power compared to regular color images. Spectral images allow us to become independent of the light source and, after normalization, of object geometry as well.

Variations in lighting conditions such as intensity, direction, and spectral power distribution are the main disturbing factors in fruit sorting applications. Traditionally, these factors are kept constant as much as possible. This is very difficult, since illumination is sensitive to external factors such as temperature and aging. In addition, this procedure does not guarantee identical results using various machines, each equipped with different cameras and light sources. Calibration of machines is tedious and error-prone. In this article, we show that by using color-constant spectral images we become independent of recording conditions such as camera and light source, as long as the light source is regularly measured (e.g., by recording a small piece of white or gray reference material in every image).

We have used spectral imaging for a very demanding problem: comparing tomatoes with very limited maturity differences. From table 5, we can see that, although the error rate increases from 19% to 36% when using different light sources, it is still considerably below the 51% for RGB under the same light source. Nevertheless, an error rate of 36% is still very high. The main reasons for this high error rate are

the rather small differences in maturity (one-day difference) and non-uniform ripening of the tomato. If tomatoes are classified as whole objects, using majority voting of the pixels, then all tomatoes are correctly classified based on the spectral images, and only one tomato is wrongly classified using the RGB images. If one class difference is allowed for classification of the individual pixels, then the error rate is about 7% for spectral images, compared to 19% for RGB. These two factors are indications that our assumption of uniform ripening of a single tomato is not fully valid and that individual pixels of the same tomato may have slightly different maturity stages.

Table 5. Error rates for individual pixels of spectral images captured with different illumination sources, using raw, color-constant, and color-constant normalized spectra. Error rates for the case in which one class difference is allowed are shown in brackets. The training pixels were captured with halogen illumination.

Method of Illumination	Raw	Color Constant	Normalized Color Constant
Halogen	19 [1]	19 [2]	19 [2]
Kg3	80 [42]	35 [4]	36 [3]
Halogen/TLD	41 [10]	35 [7]	34 [7]
Halogen/Blue	42 [6]	36 [7]	33 [4]

Tomatoes are spherical objects with a shiny, waxy skin. Since high intensity illumination is required for spectral imaging, it is almost impossible to avoid specular patches on the tomato surface. Pixels from these specular patches do not merely show the reflection values of the tomato, but they also exhibit the spectral power distribution of the illumination source. To avoid disturbance from this effect, preprocessing the images was needed to discard these patches.

In the normalized spectral image, the color difference due to object geometry has also been eliminated. When using normalized images, the color is independent of the surface normal, the angle of incident light, the viewing angle, and shading effects, as long as sufficient light is still present and under the assumption of non-specularity. The results indicate that the normalized spectral images yield at least the same, if not better, results than the color-constant spectral images. Since tomato fruit is a spherical object, the above-mentioned effects play a role in the images. Because the training pixels were randomly taken from the whole fruit surface, the positive effect of normalization could possibly be achieved in the color-constant images using linear discriminant analysis. In situations where the training pixels are taken from positions on the tomato surface that are geometrically different from the validation pixels, it is expected that normalized spectral images would give a significantly better result than color-constant spectra only. Since the normalized images do not perform worse than the color-constant images, normalization is preferred. However, care should be taken not to include specular patches.

The accuracy of spectral imaging appeared to suffer slightly if different light sources were used. Under all circumstances, however, the results were better than those for RGB color imaging under a constant light

source. This opens possibilities to develop a sorting machine with high accuracy that can be calibrated to work under different conditions of light source and camera.

## REFERENCES

- Arias, R., L. Tung Ching, L. Logendra, and H. Janes. 2000. 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. Agric. and Food Chemistry* 48(5): 1697-1702.
- Choi, K., G. Lee, Y. J. Han, J. M. Bunn. 1995. Tomato maturity evaluation using color image analysis. *Trans. ASAE* 38(1): 171-176.
- Fukunaga, K. 1990. *Introduction to Statistical Pattern Recognition*. 2nd ed. New York, N.Y.: Academic Press.
- Herrala, E., and J. Okkonen. 1996. Imaging spectrograph and camera solutions for industrial applications. *International J. Pattern Recognition and Artificial Intelligence* 10(1): 43-54.
- Hyvonen, T., E. Herrala, and A. Dall'Ava. 1998. Direct sight imaging spectrograph: A unique add-on component brings spectral imaging to industrial applications. *SPIE Symposium on Electronic Imaging*, No. 3302-21: 165-175. Bellingham, Wash.: SPIE.
- Polder, G., and G. W. A. M. van der Heijden. 2001. Calibration and characterization of spectral imaging systems. In *Multispectral and Hyperspectral Image Acquisition and Processing, Proc. SPIE 4548*: 10-17. Bellingham, Wash.: SPIE.
- Ripley, B. D. 1996. *Pattern Recognition and Neural Networks*. Cambridge, U.K.: Cambridge University Press.
- Shafer, S. A. 1985. Using color to separate reflection components. *COLOR Res. Appl.* 10(4): 210-218.
- Stokman, H., and T. Gevers. 1999. Hyperspectral edge detection and classification. In *Proc. Tenth British Machine Vision Conference*, vol. 2: 643-651. Nottingham, U.K.
- Van den Broek, W. H. A. M., D. Wienke, W. Melssen, R. Feldhoff, T. Kantimm, T. Huth-Fehre, and L. Buydens. 1997. Application of a spectroscopic infrared focal plane array (FPA) sensor for on-line identification of plastic waste. *Applied Spectroscopy* 51(6): 856-865.
- Van der Heijden, G. W. A. M., G. Polder, and T. Gevers. 2000. Comparison of multispectral images across the Internet. In *Proc. SPIE: Internet Imaging 3964*: 196-206. Bellingham, Wash.: SPIE.
- Van Dijk, M., W. van den Berg, and W. Cadel. 2001. Product-Info Tomaat 2000/2001, Report No. PT 2001-30. Zoetermeer, The Netherlands.

