DETECTION OF DEFECTS ON SELECTED APPLE CULTIVARS USING HYPERSPECTRAL AND MULTISPECTRAL IMAGE ANALYSIS

P. M. Mehl, K. Chao, M. Kim, Y. R. Chen

ABSTRACT. *Apple defects cause food safety concerns touching the general public and strongly affect the commodity market. Because accumulations of human pathogens are usually observed on surface lesions, detection of lesions is essential for assuring quality and safety. This article presents the application of hyperspectral image analysis to the development of multispectral techniques for the detection of defects on three apple cultivars: Golden Delicious, Red Delicious, and Gala. Two steps were performed: (1) hyperspectral image analysis to characterize spectral features of apples for the specific selection of filters to design the multispectral imaging system and (2) multispectral imaging for rapid detection of apple contaminations. Good isolation of scabs, fungal, soil contaminations, and bruises was observed with hyperspectral imaging using either principal component analysis or the chlorophyll absorption peak. This hyperspectral analysis allowed the determination of three spectral bands capable of separating normal from contaminated apples. These spectral bands were implemented in a multispectral imaging system with specific band pass filters to detect apple contaminations. In this preliminary work with 153 samples, good separation between normal and contaminated apples was obtained for Gala (95%) and Golden Delicious (85%). However, separations were limited for Red Delicious (76%).*

Keywords. *Food safety, Fruit, Machine vision, Spectroscopy.*

The increasing occurrence of food borne diseases and the difficulty of treating them makes it desirable to ensure as close as possible to zero contamination level. To reach this goal, various techniques have been proposed and some of them are still under investigation, such as biosensors, optical sensors, and biofilms to test the safety and quality of fruits and vegetables. Apples are one of the most important fruit commodities in the U.S. markets, with widespread processing applications, such as cider. Contamination of dropped apples leads to high incidence of contaminated of non–pasteurized cider resulting in possible illness outbreaks (Tamblyn et al., 1999; Center For Disease Control, 1998).

Several techniques have been reviewed for investigating quality defects on apples (Abbott et al., 1997). One of the most recent technologies is hyperspectral imaging (Gat and Subramanian, 1997). The technology combines imaging, spectrometric, and radiometric techniques to provide a spectrum for each pixel of a captured image. Hyperspectral imaging is capable of showing differences between the spectra of image pixels associated with the normal or abnormal parts of the samples (Mehl et al., 2001). However, hyperspectral–imaging technology currently cannot be directly implemented in an online system for automated detection of defects, diseases, and contaminations on commodities, because the time requirements for online hyperspectral image acquisition and analysis are too great.

One of the analytical methods applied to hyperspectral imaging is the classical multivariate analysis technique of principal component analysis. This technique is commonly used for various applications in multivariate analysis or chemometrics (Malinowski and Howery, 1980). This technique has been used in hyperspectral–imaging techniques for remote sensing analysis (Richards, 1994) and also in food safety for classification of poultry carcasses (Chen et al., 1998). Principal component analysis has been shown recently to be capable of determining the presence of bruises on Red Delicious apples when utilizing the near infrared spectral region (Lu et al., 1999). A recently reported analytical method for image processing of normal and contaminated apples used an asymmetric second difference method to provide good visible separations between diseased, bruised, or contaminated apples and normal apples (Mehl et al., 2001).

A hyperspectral–imaging system has been developed at the Instrumentation and Sensing Laboratory (ISL, Agricultural Research Service, United States Department of Agriculture) for detecting defects and contaminations on food commodities. The system is utilized for developing multispectral methodologies. Multispectral image analysis is a faster technique based on a discrete spectral analysis at a few wavelengths as opposed to the continuous spectral analysis performed by the hyperspectral–imaging technology. Hyperspectral–imaging analysis is capable, through multivariate analysis, of identifying the main contributing wavelength bands while preserving the information of the images and therefore optimizing theoretically the multispectral method-

Article was submitted for review in April 2001; approved for publication by the Food & Process Engineering Institute of ASAE in November 2001.

The authors are **Patrick M. Mehl**, Biomedical Engineer, Vitreous State Laboratory, The Catholic University of America, Washington D.C.; **Kuanglin Chao**, Research Scientist, **Moon Kim**, Biophysicist, and **Yud–Ren Chen, ASAE Member Engineer**, Research Leader, Instrumentation and Sensing Laboratory, USDA–ARS, Beltsville, Maryland. **Corresponding author:** Kuanglin Chao, USDA/ARS/ISL, Building 303, BARC–East, 10300 Baltimore Ave., Beltsville, MD 20705–2350; phone: 301–504–8450; fax: 301–504–9466; e–mail: chaok@ba.ars.usda.gov.

ologies (Chao et al., 2000). However, transferring these results to multispectral technology is limited to the spectral range of each channel available in the different cameras. The present analysis has been directed towards the visible spectral range for the determination of contaminations, diseases, and defects on the apples.

The objective of this study was to develop a simple multispectral detection system utilizing only three channels in the visible spectral range. The system utilized an adaptable red, green, and blue (RGB) camera where the three channels can be limited with relatively narrow bandpass filters. Three apple cultivars (Red Delicious, Golden Delicious and Gala) were chosen for their wide popularity on the market and their differences in the visible spectral range. Red Delicious apples present a dominant red color with a yellowish minor color. Golden Delicious apples are yellow with an occasionally reddish minor area on their sunny side. Gala apples have a mixture of yellow, red, and orange colors. These differences in the visible spectrum are recognized and incorporated within the multivariate analysis done for the classification of the normal and contaminated apples. Detection of apple contaminations can be implemented at two stages after harvesting: either before processing (e.g., fungal treatment, waxing, etc.) or after processing. In this study, the goal was to detect contaminations at the initial stages of post–harvest handling, before further processing.

Materials And Methods

Apples

Three apple cultivars were selected for the study due to their shape and spectral differences: Red Delicious, Golden Delicious, and Gala. Normal apples (without defects) were harvested fresh from trees or taken from barrels before any treatment, sorting, or processing at the farm or the processing plant (Rice Fruit Co., Gardners, Pa.). Abnormal apples (with defects including bruises, diseases, and contaminations) were also collected from trees or barrels, or from the ground. Apples were placed on trays, bagged in plastic to avoid dehydration, and boxed for transport to ISL. Upon reaching the laboratory, apples were stored in a cold room at 1–4°C under constant relative humidity. The apples were measured by the hyperspectral–imaging system after reaching room temperature. Approximately 20 normal apples of each cultivar were randomly selected from a large set of collected apples without any visible defects. Abnormal apples were visually categorized for the type of disease, defect, or contamination. The inventory of these apples is listed in table 1 for each cultivar.

Hyperspectral Image Acquisition

Sets of apples were scanned using a hyperspectral–imaging system (Kim et al., 2001). The equipment consisted of a CCD camera system (SpectraVideo™, PixelVision Inc., Beaverton, Oreg.) equipped with an imaging spectrograph SPECIM ImSpector version 1.7 (Spectral Imaging Ltd., Oulu, Finland). The ImSpector had a fixed–size internal slit to define the field of view for the spatial line and a prism/grating/prism system for the separation of the spectra along the spatial line. The sample was scanned line by line with an adjustable rate while illuminated by the two line sources when passing through the field of view for the camera.

Table 1. Number of samples and cultivars for multispectral image measurement.

	Gala	Red Delicious	Golden Delicious
Bruise	6	7	6
Scab	7	6	14
Side rot	6	6	
Fungus	7		
Black pox			4
Gloeosporum			7
Sooty blotch			3
Soil contamination	4	6	5
Normal	20	19	20
Total	50	44	59

The image acquisition and recording were performed with a Pentium–based PC using a general purpose imaging software (PixelView™ 3.10, PixelVision Inc., Beaverton, Oreg.). A C–mount set with a focusing lens and an aperture diaphragm allowed for focusing and aperture adjustments. For that purpose, the circular aperture was opened at its maximum and an external slit was adjusted with micrometer actuators to optimize the light flow and the resolution.

The light source consisted of two 21V, 150W halogen lamps powered with a regulated DC voltage power supply (Fiber–Lite A–240P, Dolan–Jenner Industries, Inc., Lawrence, Mass.). The light was transmitted through two optic fibers to a line light illuminator. The sample was placed on a conveyor belt with an adjustable speed AC motor control (Speedmaster, Leeson Electric Motors, Denver, Colo.).

Spatial and spectral calibration were performed prior to experiments using, respectively, spatial resolution grids and an Oriel lamp set with known spectra (Hg–Ne, A, Kr, He, and Ne lamps). A 256–pixel line was available along the spectral direction for the spectrograph. Spectral calibration provided the expression with the pixel number (p): $\lambda_\pi = 3.68 \times p + 420.87$ nm with $R^2 = 0.998$ and λ_π being the wavelength at the channel p using a two–binning process to shorten the calculation time. For the present experiments, only the visible spectrum was considered (75 pixels in the 400– to 700–nm range). The spatial resolution was found to be between 0.75 and 0.5 mm along the X or Y (scanning) axis.

Determination of the Band Pass Filters

Hyperspectral–imaging analysis was performed on normal and defective or contaminated apples of the three chosen cultivars. Reflectance intensities were calculated using flat–field correction. White field or 100% reflectance was obtained by recording the reflected light from a flat shield of white Teflon (W) and the dark field or 0% reflectance was obtained by recording the reflected light with the lens of the camera completely closed (D). The normalized reflectance (R%) expressed in % corresponding to the recorded reflected light from the sample image (I) was then defined using the following expression:

$$R\% = 100 \times \frac{I - D}{W - D} \quad (1)$$

The analysis used the product 100 × R%, rather than only R%, to increase the visualization of separation during image processing. Reflectance spectra for normal and abnormal

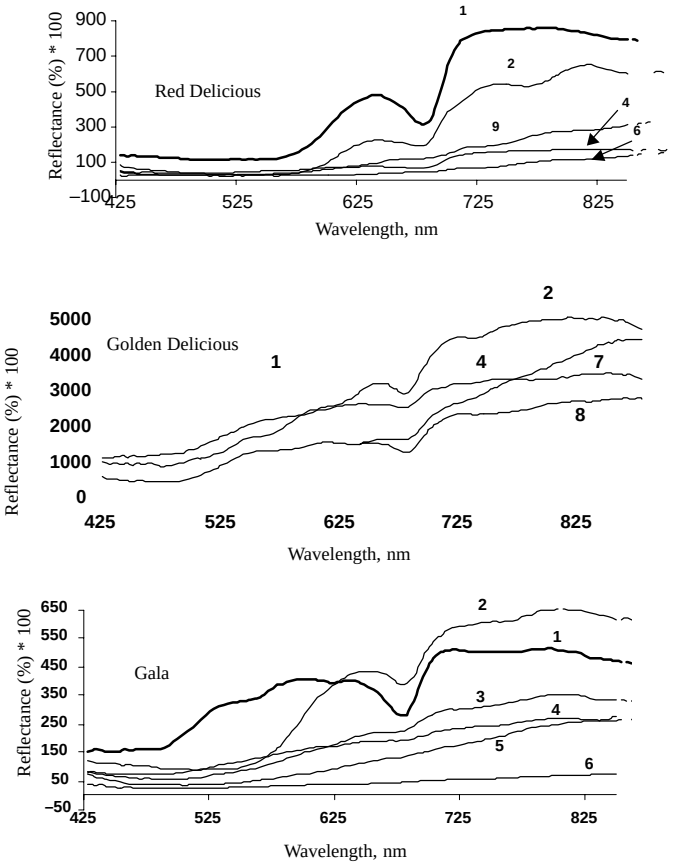

Figure 1. Example reflectance Spectra for parts of apples: 1) normal, 2) bruise, 3) gloeosporum, 4) soil contamination, 5) side rot, 6) scab, 7) black pox, 8) sooty blotch, and 9) insect bite.

surfaces were recorded for comparison and examples are shown in figure 1. Obvious differences exist in the visible spectral range, especially for the chlorophyll absorption peak at around 700 nm.

Principal component analysis was performed utilizing the hyperspectral imaging of normal and abnormal apples. Each principal component was an image that consisted of a weighted sum of the images at 120 wavelengths according to the formula:

$$E = \sum_{i=1}^{120} w_i I_i \qquad (2)$$

where I_i is the image at wavelength number i, and w_i is the weight (E is also called an eigenvector. The different principal components or eigenvectors have different sets of weights.). The first 10 principal components were visually examined, and the one that provided the best visual contrast between normal and abnormal apples was selected. The corresponding wavelength weight distribution was determined. This analysis was performed using ENVI 3.2 software from Research Systems, Inc. (Boulder, Colo.). Figure 2 shows an example of the weight distribution for the principal component giving the best discrimination between normal and bruised Golden Delicious apples. It is obvious that several band–pass filters can be selected to reconstruct the principal component in the visible spectral range. Similar distributions were obtained with the Red Delicious and Gala apples.

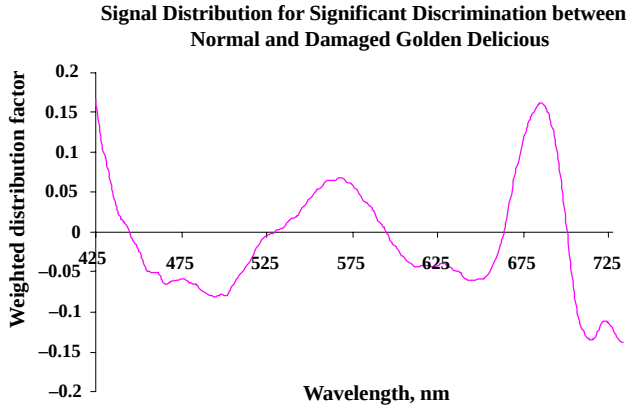

Figure 2. Weight distribution factor for the best eigenvector corresponding to the best discrimination between normal and bruised "Golden Delicious" apples. The data have been calculated using hyperspectral–imaging analysis.

The selection of filters must also match the channels of the RGB camera. The RGB camera described below has its own restriction within the visible spectral range for the three channels (red, green and blue) that are defined by the prisms' hardware along the optical path. The choice of the wavelength bandpass filters must therefore be in agreement with the spectral range of each of the channels of the camera. For each defect, the weight distribution for the principal component giving the best discrimination was plotted. From these plots, the three wavelengths that showed the best potential of discriminating all the defects were chosen. The first filter chosen had a band pass of 705 ± 40 nm to cover the last two bands for the eigenvector definition in the red channel. The other two filters were chosen to be at 575 ± 20 nm and 460 ± 20 nm in the green and the blue channels, respectively. The ±20 nm bandpasses have been used at ISL as a good balance between the need for enough light and the narrowness of the absorbing bands. In the 705–nm region, the sensitivity of the CCD falls off so a wider band pass allows more signal. These band–pass filters were installed and secured in the three–channel common aperture camera for multispectral imaging measurement.

Multispectral Image Acquisition

The multispectral imaging system consisted of a three–channel common aperture camera (TVC3, Optec, Milano, Italy), an illumination chamber, and a computer equipped with a frame grabber (XPG–1000, Dipix, Ontario, Canada). A darkened chamber was built with a box with a round–open inlet (7 cm diameter) on the top, through which the camera was mounted facing downwards. The illumination was provided using a pair of fiber–optic lights (QDF5048, Dolan–Jenner Industries, Mass.) equipped with an AC regulated 150W quartz–halogen illuminator (PL841, Dolan–Jenner Industries, Mass.). The dual lights were mounted 12 cm apart and were covered with plastic light diffusers, 30 cm from the object. The walls of the chamber were built with optical grade black acrylic to avoid uncontrolled light reflections. Each sample apple was placed on a white Teflon sheet mounted to a laboratory jack for vertical position control. Image size was 728 × 572 pixels, with each pixel representing a sample area of 0.05 mm^2.

The three–channel system was a prism–based system with three spectral range cut–offs. Three broadband images can be simultaneously acquired with the following spectral ranges: blue channel (B) with 430 nm < B < 495 nm, green channel (G) with 495 nm < G < 605 nm and red channel (R) with 605 nm < R < 900 nm. The three band–pass filters selected from the hyperspectral–imaging analysis were manually placed between the prisms and the three CCD's, respectively, to obtain the spectral specificity needed to optimize contrast between normal area and defects on apples.

Multispectral Analysis for Classification

Multispectral image reflectance parameters between normal and abnormal (defects and diseases) apples were compared based on statistical significance tests. The accompanying classification was then used in a SAS 6.12 (SAS Institute Inc., Cary, N.C.) discriminant analysis procedure (PROC DISCRIM) using the selected features from the significance tests. This procedure was done to find a distinct subset of features that could separate the normal and abnormal apples. Linear discriminant functions were used for data with approximately multivariate normal distributions within classes. The classification accuracy was evaluated by the cross–validation method (Lachenbruch and Mickey, 1968).

Results and Discussion

Principal Components Analysis (PCA) for Processing Multispectral Images

The determination of principal components for normal Gala apples is reported as an example in figure 3. The images recorded from the three–channel camera were pre–processed using flat–field correction as with the hyperspectral–imaging analysis.

The next step was to define a mask to eliminate the background and its further contribution during the statistical analysis. This mask was calculated using the corrected images in the blue channel followed by segmentation and morphological operations such as opening/closing with a very large window, and erosion to eliminate the edge effects. After the opening/closing procedure, it was observed that the stems were eliminated from the edge of the apples. Dougherty (1992) discussed morphological operations in detail. These morphological operations include erosion and dilation, and combinations of these operations. Erosion removes pixels from an image. A complementary operation known as dilation can be used to add pixels. The combination of an erosion followed by a dilation is called an opening (opens up spaces between just–touching features). The combination of a dilation followed by an erosion is called a closing, which is used to fill missing pixels within narrow gaps between portions of a feature.

The mask was then applied to each of the three–channel images after flat–field correction. The principal component analysis was performed on these three masked images to calculate the eigenvectors of normal apples. The first principal component band (PC1) was the one containing the most information, as shown in figure 3, and was the one to select. The weight distributions or coordinates of the eigenvectors in the RGB vector base were then determined for a random set of normal apples (5 to 10 apples).

The barycentric coordinates were initially defined as the coordinates equal to the means of the measured coordinates. The barycentric coordinates of the eigenvectors corresponding to the PC1 of the normal apples were calculated for at least five apples. These barycentric coordinates were those that provided the most similar results within the set of normal apples. The mean and standard deviation corresponding to these barycentric coordinates were calculated for each normal set of the apple cultivars and are reported in the table 2. Assuming that these barycentric eigenvectors were denoted Eb for each cultivar, the distance (I) of the sample to these barycentric eigenvectors was then defined as the projection of the recorded multispectral reflectance images along these eigenvectors. This projection was mathematically defined as the scalar product of the eigenvector with the image vector (Red, Green, Blue) using the expression:

$$I = Eb^{T} \begin{bmatrix} \text{Red} \\ \textit{Green} \\ \textit{Blue} \end{bmatrix} \quad (3)$$

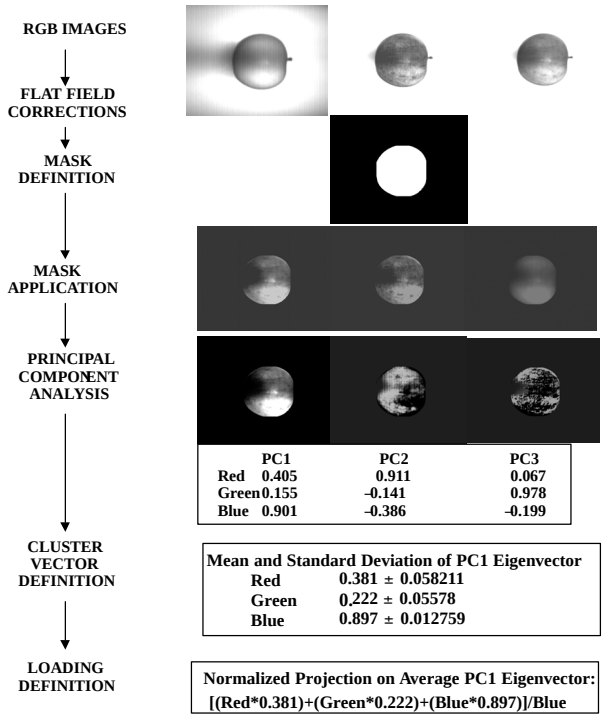

Figure 3. Data processing for determination of the principle components of the normal "Gala" apple images utilizing the three–channel camera.

Table 2. Principal component analysis: main eigenvector statistics.

Apple	Channel	Barycentric Coordinate	(Mean / Std)
Gala	Red	0.381	0.0582
	Green	0.222	0.0558
	Blue	0.897	0.0128
Red Delicious	Red	0.591	0.0940
	Green	0.167	0.2140
	Blue	0.724	0.0790
Golden Delicious	Red	0.561	0.0072
	Green	0.587	0.0073
	Blue	0.583	0.0117

To further analyze the presence of defects or contaminations on the apples, multispectral images of the sample apples were projected along these barycentric eigenvectors, similar to estimating the loading of the images along this vector. A normalization was simultaneously provided by the most contributing channel to the images, which in this experiment was the blue channel image. For each of the apple cultivars, a distance or loading function was therefore defined by:

$$\text{Red Delicious: } I = [(\text{Red} \times 0.591) + (\text{Green} \times 0.167) + (\text{Blue} \times 0.724)]/\text{Blue} \quad (4)$$

$$\text{Golden Delicious: } I = [(\text{Red} \times 0.561) + (\text{Green} \times 0.587) + (\text{Blue} \times 0.583)]/\text{Blue} \quad (5)$$

$$\text{Gala: } I = [(\text{Red} \times 0.381) + (\text{Green} \times 0.222) + (\text{Blue} \times 0.897)]/\text{Blue} \quad (6)$$

where Red, Green, and Blue were the reflectance intensities of the sample image recorded in the red, green and blue channels, respectively. The analysis for the classification of the apples using the multispectral analysis was then ready to proceed.

Elimination of Apples with Scabs

Examination of figure 1 suggests that the presence of scab is expected to be obvious and its detection can be performed simply, without a highly time–consuming analysis. A direct segmentation (fig. 4) from the principal component images (blue channel value between 0.5 and 1) provided a 87, 100, and 57% determination of scab presence for contaminated Red Delicious, Golden Delicious, and Gala apples, respectively. The accuracy of the scab detection was improved with an additional normalization step. The 460–nm image was used to generate a mask. The 575– and 705–nm masked images were ratioed to the 460–nm image before further processing. The new masked 575– and 705–nm images together with the original masked 460–nm image were called the "normalized" images. The original principal component image was then generated from the normalized images. Proceeding with a subsequent threshold (threshold = 0.9) method resulted in a 100% separation between normal and scab–contaminated Red Delicious, Golden Delicious, and Gala apples without any misclassification. Application of these methods allowed us to eliminate the apples with scabs from the set of contaminated apples. We will therefore not consider apples with scab contamination any further since they can be easily detected without additional analysis.

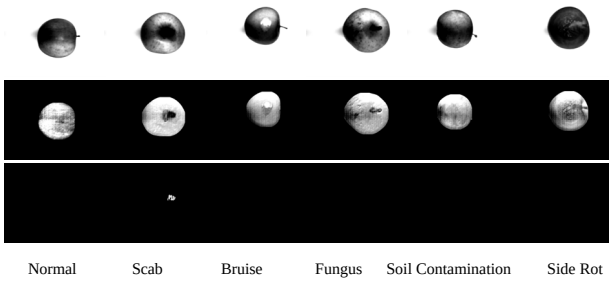

Figure 4. Detection of scab using direct threshold method for "Gala" apples: the first row illustrates the flat–field images from the blue channel; the second row illustrates the image result from application of loading function; the last row illustrates the threshold results for scab detection.

Multispectral Image Analysis for Other Contaminations and Diseases

Six parameters using the distance functions defined above in equations 4, 5, and 6 were used: a) minimum and maximum reading values to define the range of intensity (range), b) the mean value over the pixels of the sample (mean), c) the associated standard deviation (std), d) the coefficient of variation defined as the ratio of standard deviation/average intensity (std/mean), e) the normalized average defined as the ratio of average intensity/range of intensity (mean/range), and f) the normalized standard deviation defined as the ratio standard deviation/range of intensity (std/range). The range was useful for compensating for the size of the apple. The mean and standard deviation were obviously needed since the presence of defects, diseases, or contaminations will lower the reflectance intensity and therefore lower the mean and increase the standard deviation over the apple image. These variations of the mean and standard deviation with the presence of defects, diseases, or contaminations will be non–linearly amplified by calculating the coefficient of variation, defined as the ratio of std/mean. It was therefore expected that this coefficient of variation would increase from normal to abnormal apples. A normalization of the previous statistical functions with the intensity range also tended to cluster the data more closely for the normal and for the abnormal apples.

Variability of Multispectral Reflectance Parameters

Table 3 summarizes significance t–test results for comparing mean values of reflectance parameters from the normal and abnormal conditions, including bruises, side rot, fungus, and soil contamination of Gala apples. Four of the six reflectance parameters (mean, std, std/range, std/mean) allowed separation of the normal apples from all those with abnormal conditions. Two parameters (range, mean/range) did not allow separation of normal apples from those with fungus or soil contamination.

For the Red Delicious apples, significance test results for comparing mean values of reflectance parameters from the normal and abnormal conditions including bruises, side rot, and soil contamination are listed in table 4. In case of comparison between normal versus side rot and normal versus soil contamination, the average reflectance intensity was useful. However, the reflectance parameters of standard deviation, range, and average/range were not significant for comparing normal versus side rot and normal versus soil contamination. For the comparison between normal versus bruised and normal versus side rot, the reflectance parameters of standard deviation/range and standard deviation/average were significant.

For the Golden Delicious apples, only two reflectance parameters (std and std/average) were significantly different between normal and all those apples with abnormal conditions (table 5). However, no significant difference was found between normal and all those apples with abnormal conditions when average reflectance intensity was used. The

Table 3. Significance test results of normal vs. abnormal conditions of Gala cultivar for six reflectance parameters.

	Average Intensity[a][b]	Standard Deviation	Range (max – min)	Average Range	Standard Deviation	
					Range	Average
Bruise	4.6127****	–3.1763**	–2.1037*	4.3532***	–5.0461****	–3.1145**
	(0.0001)	(0.0041)	(0.0461)	(0.0002)	(0.0001)	(0.0047)
Side Rot	–7.6661****	–5.7089****	–1.8155*	2.2521*	–8.6264****	–6.2521****
	(0.0001)	(0.0001)	(0.0819)	(0.0337)	(0.0001)	(0.0001)
Fungus	7.9336****	–3.4105**	0.9846	–0.0334	–6.2881****	–4.1046***
	(0.0001)	(0.0022)	(0.3343)	(0.9736)	(0.0001)	(0.0004)
Soil	9.1043****	–7.7708****	–1.4272	2.2156*	–9.5509****	–9.2665****
	(0.0001)	(0.0001)	(0.1676)	(0.0374)	(0.0001)	(0.0001)

[a] $H_0 : \mu1 = \mu2$; H1 : $\mu1 > \mu2$. ; Values in each column represent t–values and the asterisks indicate the significant levels (0.1 = * , 0.01 = ** , 0.001 = *** , 0.0001 = ****).

[b] Values in the parentheses are the P values for performing a significant test to reject the null hypothesis ($\mu1 = \mu2$).

Table 4. Significance test results of normal vs. abnormal conditions of Red Delicious cultivar for six reflectance parameters.

	Average Intensity[a][b]	Standard Deviation	Range (max – min)	Average Range	Standard Deviation	
					Range	Average
Bruise	–0.0888	–2.5489*	–2.4421*	2.6691*	–1.9598*	–3.0643**
	(0.9299)	(0.0176)	(0.0223)	(0.0134)	(0.0617)	(0.0053)
Side Rot	2.8004*	–1.1347	0.5685	0.54361	–2.9020**	–2.4466*
	(0.0102)	(0.2682)	(0.5752)	(0.59119	(0.0080)	(0.0225)
Soil	2.7924*	0.5685	1.1206	–0.4326	–0.4275	–0.0464
	(0.0103)	(0.5752)	(0.2739)	(0.6693)	(0.6729)	(0.9633)

[a] $H_0 : \mu1 = \mu2$; H1 : $\mu1 > \mu2$. ; Values in each column represent t–values and the asterisks indicate the significant levels (0.1 = * , 0.01 = ** , 0.001 = *** , 0.0001 = ****).

[b] Values in the parentheses are the P values for performing a significant test to reject the null hypothesis ($\mu1 = \mu2$).

Table 5. Significance test results of normal vs. abnormal conditions of Golden Delicious cultivar for six reflectance parameters.

	Average Intensity[a][b]	Standard Deviation	Range (max – min)	Average Range	Standard Deviation	
					Range	Average
Bruise	0.2050	–2.9603**	–0.5611	0.5736	–2.4086*	–2.9317**
	(0.8393)	(0.0068)	(0.5799)	(0.5715)	(0.0241)	(0.0073)
Black Pox	–0.5386	2.2927*	0.5301	–0.5159	2.2758*	2.5379*
	(0.5956)	(0.0318)	(0.6014)	(0.6111)	(0.0329)	(0.0187)
Gloesporum	0.3921	1.8160*	2.6259*	–2.1678*	–0.7051	2.2370*
	(0.6983)	(0.0786)	(0.0145)	(0.0398)	(0.4872)	(0.0275)
	0.3504	3.4513**	–1.1653	1.0181	1.7046	3.6866**
Sooty Blotch	(0.5602)	(0.0021)	(0.2569)	(0.3202)	(0.1031)	(0.0098)
	–0.6346	3.4513**	2.3832*	–2.1678*	0.9662	–3.6315**
Soil	(0.5319	(0.0022)	(0.0258)	(0.0399)	(0.3440)	(0.0014)

[a] $H_0 : \mu1 = \mu2$; H1 : $\mu1 > \mu2$. ; Values in each column represent t–values and the asterisks indicate the significant levels (0.1 = * , 0.01 = ** , 0.001 = *** , 0.0001 = ****).

[b] Values in the parentheses are the P values for performing a significant test to reject the null hypothesis ($\mu1 = \mu2$).

reflectance parameters (range and mean/range) allowed separation of normal from those with gloeosporum and soil contamination apples. The reflectance parameter of std/range was useful to separate normal from bruised and black pox apples.

Accuracy of Cross–Validation Classification

The cross–validation method was applied for discriminant analysis to classify apples into normal and abnormal conditions. Table 6 shows the classification results of Gala using all six reflectance parameters and using the four chosen from significance tests. The results of the classification are presented in the form of a confusion matrix showing the numbers correctly classified (on the diagonal) and the numbers misclassified as the other apple type. The accuracy of the linear discriminant model was 100% to separate normal and abnormal conditions of Gala apples when six reflectance parameters were used. It shows only slight decrease in classification ability (95% for normal and 96% for abnormal) when four of the six reflectance parameters were used.

Table 6. Cross–validation results for classification of Gala cultivar using all parameters.[a]

		Predicted		
		Normal	Abnormal	Percentage Correct (%)
Actual	Normal	20 (19)	0 (1)	100 (95)
	Abnormal	0 (1)	23 (22)	100 (96)

[a] Results using the selected four reflectance parameters appear in parentheses.

Table 7. Cross–validation results for classification of Red Delicious cultivar using all parameters.[a]

		Predicted		
		Normal	Abnormal	Percentage Correct (%)
Actual	Normal	12 (17)	7 (2)	63 (89)
	Abnormal	7 (7)	12 (12)	63 (63)

[a] Results using the selected three reflectance parameters appear in parentheses.

Table 8. Cross–validation results for classification of Golden Delicious cultivar using all parameters.[a]

		Predicted		
		Normal	Abnormal	Percentage Correct (%)
Actual	Normal	14 (17)	6 (3)	70 (85)
	Abnormal	8 (4)	17 (21)	68 (84)

[a] Results using the selected two reflectance parameters appear in parentheses.

Table 7 shows the accuracy of discriminant models for classifying normal and abnormal Red Delicious apples. The accuracy was 63% for classifying normal and abnormal apples when six reflectance parameters were used. Although using all six reflectance parameters resulted in a low overall accuracy (63%), using only three reflectance parameters showed high separability (89%) for classification of normal apples.

Table 8 shows the accuracy of discriminant models for classifying normal and abnormal Golden Delicious apples. Although using all six parameters resulted in limited classification accuracy (70% for normal and 68% for abnormal), using only two parameters showed much better results (85% for normal and 84% for abnormal). This result indicates that the unused parameters added no significant information to the discrimination. The situation is similar to that of adding insignificant terms to a linear regression, which can result in a lower correlation coefficient.

Discussion of the Choice of the Filters

The 705–nm filter was obviously necessary for the determination of the attenuation of chlorophyll activity on the apple due to the presence of defects or contaminations (fig. 1). This wavelength choice for the red channel was also supported by previously reported spectrometric studies that showed the possibility of identifying bruises on apple surfaces (Upchurch et al., 1990). A more recent study using filtered imaging techniques extended the possibility to identifying a wider range of apple surface defects (Aneshansley et al., 1997). The corresponding statistical analysis showed good surface defect classification at 540, 750, 970, and 1030 nm, depending on the apple cultivars and their surface defects (Aneshansley et al., 1997). The Mahalanobis distance analysis between normal and abnormal areas of the apples was significant for wavelengths around 750 nm and also around 540 nm. The next obvious differences existed in the middle range of the visible spectrum corresponding to the green channel of the camera. The compromise for the differences between Golden Delicious, Red Delicious, and Gala apples led to a choice of 575 nm. Previous reports on the separation of normal and bruised areas on Red Delicious apples concluded that the spectral range 720 to 840 nm provided the best models for discrimination of bruised and non–bruised apples (Upchurch et al., 1990).

The detection of early frost damage and bruises for Empire apples was been optimized by spectrometry utilizing the spectral range of 450 to 490 nm with a normalization wavelength at around 700 nm (Upchurch et al., 1991). Empire apples have similar color characteristics to the Gala apples. They have a red or yellow color combined with orange stripes. These characteristics have been related to the Red Delicious and Golden Delicious apples to direct the search for optimal wavelengths for analysis and discrimination of bruised and non–bruised Empire apples. The present approach is similar with a different development by using a distance function from normal apples.

Surface defects already have been automatically detected on Golden Delicious apples using a filter at 550 ± 100 nm (Davenel et al., 1998). The consideration of such a large band–pass range suggests that the variation over that range of the reflectance intensity is sufficient from normal to surface defective apples to allow the discrimination.

Conclusions

The first conclusion is that hyperspectral–imaging analysis is a useful tool for the development of less technologically sophisticated techniques such as multispectral analysis. Hyperspectral–imaging technology can provide more details and more precise determination of the most important spectral bands that can be utilized for classification or separation of information.

Despite their color differences, it has been possible to use the same configuration combined with the same analytical methodology with distinct distance functions to perform similar results on different cultivars. This approach is the first time that simple statistical functions have been used to achieve good separations between normal and abnormal apples.

The present RGB camera is limited by its internal prism organization to spectral ranges in the visible spectrum but our results support the possibility of using specific visible bands to detect defects, diseases, and contamination on apples. Filters at 705 and 460 nm are apparently more essential for the design of the present system than 575 nm. The present multispectral analysis system is actually capable of classifying normal and abnormal apples for the three cultivars. The classification for normal/abnormal apples is found to be close to 63 and 70% for Red Delicious and Golden Delicious apples, respectively. The classification approaches 85 to 89% when choosing fewer statistical parameters. We found to our own surprise that the multicolor cultivar Gala had the best classification with a 100% separation between normal and abnormal apples. We have therefore developed a general approach to utilize the hyperspectral–imaging analysis to better design a rapid multispectral analysis system for food safety and for food quality also. Trying to compromise between different cultivars may lead to a more complex approach to the design of multispectral imaging analysis.

This work was a first attempt to use hyperspectral imaging to guide the selection of wavelengths in a multispectral camera. In a practical application, a method (such as multiple images of a rotating apple) would have to be found to bring the plane of each defect as close as possible to perpendicular-

ity to the camera–sample axis. In this work, the samples were oriented manually to provide the most contrast between normal and abnormal areas. Also, a camera that was not restricted to three specific wavelength regions would be more flexible. The RGB camera used in this work was chosen for its built–in image registration capabilities, and it could probably be modified to accommodate any set of wavelengths in the visible/near–infrared region. Also, mention should be made of the limitations of the sample set. Many more apples from a several locations would have to be added to the calibration before settling on the optimum set of wavelengths and coefficients. In particular, bruise detection here was limited to naturally occurring bruises that were clearly visible. Exploration into the least detectable bruise size and age is necessary.

REFERENCES

Abbott, J. A., R. Lu, B. L. Upchurch, and R. L. Stroshine. 1997. Technologies for non–destructive quality evaluation of fruits and vegetables. *Horticultural Reviews* 20(1): 1–120.

Aneshansley, D. J., B. L. Upchurch, and J. A. Throop. 1997. Reflectance spectra of surface defects on apples. *Proceedings of the Sensors for Nondestructive Testing International Conference, Northeast Regional Agricultural Engineering Service*, 143–160. College Park, Md.

Center For Disease Control. 1998. Preventing emerging infections diseases: a strategy for the 21st Century. *MMWR Morb. Mortal Wkly Rep.* 47: 1–14.

Chao, K., P. M. Mehl, and Y. R. Chen. 2000. Use of hyper– and multi– spectral imaging for detection of chicken skin tumors. ASAE Paper No 003084. St. Joseph, Mich.: ASAE.

Chen, Y. R., B. Park, R. W. Huffman, and M. Nguyen. 1998. Classification of on–line poultry carcasses with back propagation neural networks. *J. Food. Proc. Engr.* 21(1): 33–48.

Davenel, A., C. Guizard, T. Labarre, and F. Sevila. 1998. Automatic detection of surface defects on fruit by using a vision system. *J. Agric. Eng. Res.* 41(1): 1–9.

Dougherty, E. R. 1992. *An Introduction to Morphological Image Processing*, 17–26. Bellingham, Wash.: SPIE Optical Engineering Press.

Gat, N., and S. Subramanian. 1997. Spectral image: Technology and applications. *Hyperspectrum News Letter* 3(1) Optoknowledge Systems, Inc. Torrance, Calif.

Kim, M. S., Y. R. Chen, and P. M. Mehl. 2001. Hyperspectral reflectance and fluorescence imaging system for food quality and safety. *Transactions of the ASAE* 44(3): 721–729.

Lachenbruch, P. A., and M. R. Mickey. 1968. Estimation of error rates in discriminant analysis. *Technometrics* 10(1): 1–10.

Lu, R., Y. R. Chen, B. Park, and K.–H. Choi. 1999. Hyperspectral imaging for detecting bruises in apples. ASAE Paper No. 993120. St. Joseph, Mich.: ASAE.

Malinowski, E. R., and D. G. Howery. 1980. *Factor Analysis in Chemistry.* New York: John Wiley & Sons.

Mehl, P. M., Y. R. Chen, and M. S. Kim. 2001. Development of detection methods for defects and contaminations on apple surface using hyperspectral imaging technique. Submitted to *Optical Engineering*.

Richards, J. A. 1994. *Remote Sensing Digital Imaging Analysis: An Introduction*. Berlin, Germany: Springer Verlag.

Tamblyn, S., J. deGrosbois, D. Taylor, and J. Stratton. 1999. An outbreak of *Escherichia coli* O157:H7 infection associated with pasteurized non–commercial, custom–pressed apple cider—Ontario, 1998. Can. Commun. Dis. Rep. 25: 113–7.

Upchurch, B. L., H. A. Affeldt, W. R. Hruschka, K. H. Norris, and J. A. Throop. 1990. Spectrophotometric study of bruises on whole, 'Red Delicious' apples. *Transactions of the ASAE* 33(2): 585–589.

Upchurch, B. L., H. A. Affeldt, W. R. Hruschka, and J. A. Throop. 1991. Optical detection of bruises and early frost damage on apples. *Transactions of the ASAE* 34(3): 1004–1009.