DETECTION OF APPLE STEM USING HYPERSPECTRAL IMAGING

ElMasry, G.[1]; Radwan, S.M.[1]; El-Amir, M.S.[1]; Wang, N.[2]

ABSTRACT

The critical part of the development of automatic grading and sorting machines of apple is the detection of defects and stem. Stem must be recognized prior to any further processes. Hyperspectral imaging technique in visible and near infrared (VIS/NIR) region with the range of 400 to 1000 nm was established for stem recognition in four apple cultivars. The studied cultivars were Granny Smith (green), Golden Delicious (yellow), Red Delicious (dark red), and McIntosh (multicolour). The spectral responses of stems and surrounding normal surface were collected and analyzed. The optimal wavelengths that gave the highest contrast between stem and its surrounding normal surface were selected depending on Variables Importance in Projection (VIP) scores extracted from Partial Least Square (PLS) analysis. These optimal wavelengths were 730, 770 and 850 nm. Then, Maximum Likelihood Classification (MLC) method was used to classify apple images based on whether or not they included a stem and to recognize its length and location. The experimental results showed that the system successfully classified a test set consisting of 200 images and a validation set consisting of 240 images at the optimal wavelengths.

Keywords: Apple, Hyperspectral imaging, Stem detection, and Maximum Likelihood Classification.

INTRODUCTION

Detection of defects, blemishes, stem, and calyxes as well as distinguishing them from normal surface is an essential part to develop automatic handling system of apple. In classical image processing related to classification of commodities like fruits and vegetables, there is high probability of misclassification due to colour similarity between defects and stem/calyx ends (Bennedsen and Peterson, 2004 and Li et al., 2002).

[1] Agricultural Engineering Department, Faculty of Agriculture, Suez Canal University, Ismailia, Egypt.

[2] Department of Bioresource Engineering, McGill University, Macdonald Campus. 21,111 Lakeshore Road, Ste-Anne-de-Bellevue, Quebec. Canada. H9X3V9

Because the image processing system was not able to distinguish between defects and the stem or calyx ends, there is a need for an orienting system to ensure that apples were oriented with the stem/calyx out of sight of the camera before they were presented to the imaging system. With incorrect orientation, there was a risk that these would be wrongly classified as defective and hence lost. (Bennedsen and Peterson, 2004).

An impressive amount of previous work using traditional image processing techniques has been dedicated to stem/calyx identification and discriminate them from defects (Crowe and Delwiche, 1996a, 1996b; Li et al., 2002; Penman, 2001; Ruiz et al., 1996; Unay and Gosselin, 2005; Yang, 1996; and Ying et al., 2003). However, no method has so far been found that provides satisfactory classification at a reasonable recognition rate. The earliest research by Brown and Segerlind, (1975) and Reid (1976) demonstrated that the reflectance properties of stem and apple skin ripe, unripe, bruised, unpeeled, peeled and bruised peeled exhibit different reflectance properties within varieties. These reflectance differences suggest that selected wavelengths could be used to discriminate between apple defects and flesh or stem and calyx ends when suitable detectors have been designed. They reported that the preferred wavelength band for stem and calyx detection is 725-800 nm. Several researchers have tried to address this problem with varying success, Kleynen et al. (2005) detected stem/calyxes using a correlation pattern matching algorithm based on a similarity template matching method. Wen and Tao (1999) developed a near-infrared machine-vision system for automating stem/calyxes and defect inspection in apple using monochrome CCD camera attached with a 700nm long-pass filter. They separated stem-ends and calyxes from defects by means of the histogram density feature of the blob. They found that the off-line tests of 200 stem-ends and 200 calyxes showed that it is very difficult to distinguish stem-ends/ calyxes from true defects, especially when stem/ calyxes appear at the edge of fruit. The recognition rates for stem and calyxes of their work were 83.50% and 86.00%, respectively.

Compared with traditional machine vision inspection, which provides limited information for distinguishing the stem and calyxes, multispectral detection provides richer information in multiple images of different spectrum sensing results of the same object. In this direction to address the

stem/calyx recognition problem, Cheng et al. (2003) proposed a near-infrared (NIR) and a midinfrared (MIR) dual-camera vision system. The NIR and MIR cameras had, respectively, a sensitive spectrum range from 700 to 750 nm and from 7500 to 13500 nm. The NIR camera could identify both the stem/calyx portion of the apple and the true defects, while MIR camera could detect only the stem and calyxes on refrigerated fruits. Li et al. (2002) reported that stem and calyxes were separated from defects by using fractal features and artificial neural network. The method was tested on forty samples of "Fuji" apples. Defects and stem/calyxes were visually well segmented from the sound part of the fruit using band-pass optical filters of 840 nm attached to the outside lens of the camera. Regarding the stem/calyx recognition, the accuracy of the neural network classifier was over 93%. Due to the interference between stem and some defects Throop et al. (2003) and Bennedsen and Peterson (2005a) developed experimental machine vision system to identify surface defects on apples using two optical filters at 740 and 950 nm, respectively. There was confusion between defects and stem/calyx area because defects appeared as dark areas, and so did shadows and parts of the stem/calyx area. They presented an approach to locate the defects and eliminate other dark areas. The design of this sorter distinguished itself by incorporating an orienting system, which aimed at orienting the apples with the stem–calyx axis perpendicular to the image-capturing camera and hence out of sight of the camera. This design eliminates the need for distinguishing between defects and stem/calyx, which is a critical part of the development of automatic apple sorters. Bennedsen and Peterson (2005b) reported that with up to 5% of the apples incorrectly oriented there was a risk that these would be wrongly classified as defective and hence lost. Thus, they suggest adding more cameras to the system to acquire three sets of images for each apple: one covering the stem region, another the calyx region, and a number of images covering the circumference as the apple rotates in front of the camera. Kleynen et al. (2005) stated that errors of classification come from a bad segmentation of the defects or from a confusion with the calyx and stem-ends because defects like russet and recent bruises present a color similar to the healthy tissue. They used image at 800 nm because this wavelength band was not

influenced by natural variations of the skin color and offered the best contrast between the stem/calyxes and the rest of the fruit.

From these studies, it appears that apple stem/calyx detection is a difficult task using conventional image acquisition devices. On the other hand, hyperspectral imaging systems provide a large amount of data which is very time consuming to acquire and to process. Thus reducing this data to such few optimal wavelengths and propose an image processing method having the potential of stem recognition could be applied in such a way in industrial applications. Thus, the overall objective of this research was to ascertain the capability of hyperspectral imaging to identify stem in images of apples and develop a proper wavelength selection method to form multispectral images in which stem could be recognized in various cultivars.

MATERIAL AND METHODS

1. Apple Samples

Apples of four cultivars Granny Smith (green), Golden Delicious (yellow), Red Delicious (dark red), and McIntosh (multicolour) were obtained from local retail stores. Abnormal apples that might contain any defects such as bruises, diseases, and contaminations were excluded to eliminate any similarities between these defects and the stem. The fruit were stored at 3°C until the time of testing, then they removed from the storage 24h before image acquisition to be equalized with surrounding temperature. Hyperspectral image of the whole fruit was acquired with the stem in a wide range of orientations. Furthermore, the hyperspectral images were captured again after 1, 3, 7 days for the same fruits to detect the difference in the spectral features of the stem compared with the normal surface. Then the images were calibrated as described below and divided into two classes, the first class consisting of 200 images (fifty images from each cultivar) was used as a calibration set to be used in extraction the optimal wavelengths. The second class consisting of 240 images was used for validation and classification routines.

2. Hyperspectral Imaging System

2.1. Component of hyperspectral imaging system

The developed hyperspectral imaging system as schematically shown in (Fig. 1) is composed of four components: (1) illumination unit which

consists of two 50W halogen lamps adjusted at angle of 45º to illuminate the camera's field of view, (2) a fruit holder surrounded by white nylon tent to diffuse the light and provide a perfect lighting condition, (3) spectrograph (ImSpector V10E, Optikon Co., Canada) coupled with a standard C-mount zoom lens, and (4) CCD camera (PCO-1600, Pco. Imaging, Germany). The optics, spectrograph and the camera, has high sensitivity from 400 to 1000 nm and the exposure time was set at 200 ms throughout the whole test. The camera-spectrograph assembly is provided with motor to move this unit through the camera's field of view to scan the fruit line-by-line. This setup resulted in spatial-spectral images with spatial dimension of 400x400 pixels and 826 spectral bands from 400 to 1000 nm. The system was controlled by PC supported with Hypervisual Imaging Analyzer® (ProVision Technologies, Stennis Space Center, USA) for image acquisition and for controlling camera, motor and spectrograph.

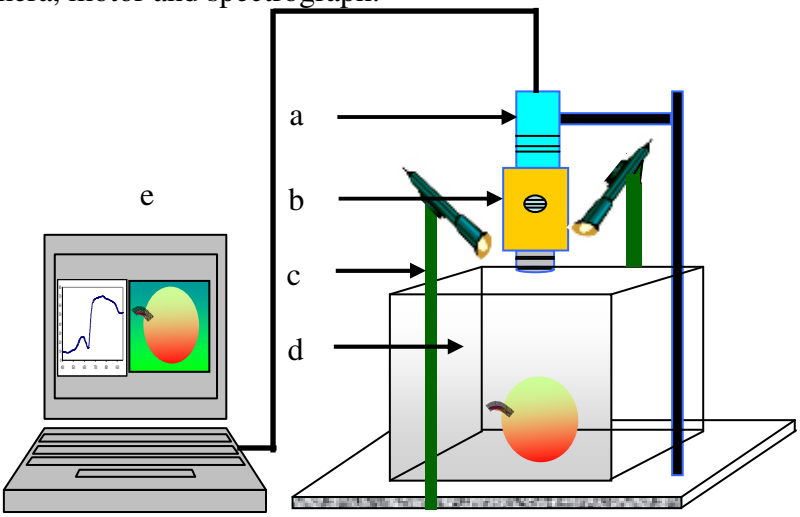

Fig. 1 Hyperspectral imaging system: (a) CCD camera, (b) Spectrograph with a standard C-mount zoom lens, (c) Halogen lighting unit, (d) White nylon tent, and (e) PC supported with image acquisition software.

2.2. Calibration of hyperspectral images

Hyperspectral images were corrected by comparison with a white reference (Teflon white board with 99% reflectance) taking into account the dark current of the CCD detectors. The corrected image (R) is then defined using the following expression:

$$R = \frac{R_o - D}{W - D} \times 100$$

Where R_o is the recorded hyperspectral image, D the dark image recorded by turning off the lighting source with the lens of the camera completely closed with a black cap, and W is the white reference image. The calibration, processing and analysis of the images were conducted using Environment for Visualizing Images (ENVI 4.2) software (Research Systems, Inc., Boulder Co., USA).

2.3. Extraction of spectral features of stem and normal surfaces

The spectral signature of any object in the hyperspectral image is defined as the pattern of reflection, absorbance, and emitting of electromagnetic energy in distinctive manner at specific wavelengths due to the differences of products' molecular compositions. This signature uniquely characterizes and identifies any given material over a sufficiently broad spectral band. So, due to the difference in molecular composition between stem and normal surface, it is required to identify the special characteristics of the stem in terms of its spectral responses at different wavelengths in the electromagnetic range from 400 to 1000 nm. Corrected images were used to extract such information about the spectral properties of stem/calyx ends as well as of the normal surface. This spectral information will be used for selecting the optimal wavelengths and classification purposes. About 2000 pixels were manually and randomly selected to collect the average reflectance spectrum of normal surface from each image for each cultivar. Also, the reflectance spectrum of the stem area were collected using the same manner using about 200-500 pixels depending on stem size.

2.4. Wavelength selection method

The topic of wavelength selection is concentrated on establishing a formal methodology, which enables optimal utilization of specific spectral bands. In general, the selected wavelength(s) should reduce the data dimensionality while preserving the most important information contained in the lower dimensional data space. Any wavelength that preserves the most energy among the hyperspectral data, carries most spectral information and maintains any valuable details among tested objects is considered an optimal wavelength. This optimal wavelength depends on the behaviour of spectral curves of the materials under study and the difference beneath

them. However, there is not a standard method to select the optimal wavelengths from the whole spectrum. Therefore, to establish consistent multispectral imaging systems, several essential spectral bands are first sought through a variety of strategies, such as general visual inspection of the spectral curves and correlation coefficients (Keskin et al., 2004), analysis of spectral differences from the average spectrum (Liu et al., 2003), stepwise regression (Chong and Jun, 2005), principal component analysis (Kim et al., 2002; Mehl et al., 2004; and Xing and De Baerdemaeker, 2005), and others (Hruschka, 2001).

In this study, to determine the optimal wavelengths, partial least squares (PLS) analysis was conducted for normal surface and stem spectra using SAS® statistical software. PLS is implemented to transfer a large set of highly correlated and often collinear experimental data into independent latent variables or factors. When applied to spectra, the aim of PLS analysis is to find a mathematical relationship between a set of independent variables, the X matrix ($N_{samples} \times K_{wavelengths}$), and the dependent variable, the Y matrix ($N_{samples} \times 1$). The spectrum source (normal surface and/or stem) represented the dependent variable (Y); meanwhile, the 826 wavelengths represented the independent variables or the predictors (X). Typically, most of the variance can be captured with the first few latent variables while the remaining latent variables describe random noise or linear dependencies between the wavelengths/predictors.

The PLS algorithm according to (Geladi and Kowalski, 1986a, 1986b; Haaland and Thomas, 1988a, 1988b; and Osborne et al., 1997) determines a set of orthogonal projection axes $\boldsymbol{W}$, called PLS-weights, and wavelength scores $\boldsymbol{T}$. Then, regression coefficients β are obtained by regressing Y onto the wavelength scores T as follow:

$$\hat{Y} = XW_a^* \beta = T_a \beta \qquad (1)$$

With $\boldsymbol{W}^*_{\hat{}} = (\boldsymbol{W}\,(\boldsymbol{P}` * \boldsymbol{W})^{-1})$,

Where, $\hat{Y}$ is the predicted surface type (normal surface or stem), a is the number of PLS factors and $\boldsymbol{P}`$ is the wavelength loadings. To facilitate the analysis, the surface type (Y) was presented in binary codes, so that the spectra of normal surfaces were denoted by zeros meanwhile the stem spectra were represented by ones.

The relative importance of wavelengths in the model with respect to surface type (Y) could be reflected by new scores called variables importance in projection (VIP) scores according to the following formula:

$$VIP_k = \sum_{j=1}^{a} (w_{jk}^2 . SSR_j) \frac{L}{SST} \qquad (2)$$

Where SSR is the residual sum-of-squares, SST is the total sum-of-squares of Y variable, and L is the total number of the examined wavelengths (826 spectral bands). Variables Importance for the Projection (VIP) score of each wavelength could be considered as selection criteria. Wavelengths with higher VIP score are considered more relevant in classification (Bjarnestad and Dahlman, 2002). In general, predictors/wavelengths could be classified according to their relevance in explaining Y as: VIP > 1.0 (highly influential), 0.8 < VIP < 1.0 (moderately influential) and VIP < 0.8 (less influential) (Olah et al., 2004). In this study, all wavelengths above threshold of 1.0 were considered optimal wavelengths to be used for further classification processes

2.5. Classification algorithm

The method presented in this work represents an attempt to mimic the way in which humans evaluate and classify objects. The idea is to base the detection on spectral signature of each pixel using limited number of wavelengths as opposed to classical image processing which normally involves identification of the objects by segmentation.

A supervised approach using Maximum Likelihood Classification (MLC) was used for this task. MLC assumes that the statistics for each class in each band are normally distributed and calculates the probability that a given pixel belongs to a specific class. Each pixel is assigned to the class that has the highest probability (i.e., the maximum likelihood). Only the images at optimal wavelengths (selected from VIP procedure) were used as the basis for this approach using ENVI software. The method depended on calculating the spectral characteristics of stem at only optimal wavelengths in terms of minimum, maximum and average reflectance as well as its covariance matrix, and then saving these data as the unique stem library to represent the stem class. Keeping the spectral signature of stem class in one hand and the reflectance from the whole fruit at the same optimal

wavelengths on the other hand, the pixel that carry the same spectral signature like stem are marked according to its highest probability. ENVI calculates the following discriminant functions for each pixel in the image as:

$$g_i(x) = \ln p(\omega_i) - \frac{1}{2}\ln|\Sigma_i| - \frac{1}{2}(x-m_i)^T \Sigma_i^{-1}(x-m_i)$$

Where:

i : Class number, 1 or 2 (stem or normal surface),

x : The spectral reflectance of any pixel in the image at optimal wavelengths,

ω_i : The spectral data matrix of each pixel belonging to class i,

$p(\omega_i)$: The probability of class having ω_i to be occurred in the image and is assumed the same for the other class (equal to 1/2).

$|\Sigma_i|$: The determinant of the covariance matrix of the data in class ω_i,

Σ_i^{-1} : The inverse of the covariance matrix of the data in class ω_i, and

m_i : The mean vector of the spectral reflectance for all pixels belonging to class i.

In fact, this method can be implemented in such a way that the system can be updated without modification, simply by providing the system with the new signature of different classes, which is then used to train the classification.

The resultant classified image usually needs further processing to better define the stem area, such as the elimination of isolated pixels and small zones. For this purpose, the binary image containing only the class stem was obtained by considering the class stem as the main object and merging the rest of the other pixels with the background pixels. Then morphological operation and a spatial filter were applied over this binary image.

RESULTS AND DISCUSSIONS

1. Spectral Characteristics of Stem and Normal Surface

Figures (2a-d) show the reflectance spectra in the visible and near infrared (VIS/NIR) range between 400 and 1000 nm for normal surfaces and stems of different cultivars after 0, 1, 3, and 7 days. In spite of cultivar, the reflectance curves of stems were rather smooth across the entire spectral region and almost have the same features during the whole period especially

in the range from 400 to 900 nm. The presence of water in the fruit gave rise to the characteristic absorption bands that appear as localized minima, so that samples containing higher moisture contents had lower reflectivity across their spectra. In spite of cultivar and its background color, the absorption curves of the normal surfaces had three broadband valleys around 500, 690 nm, and 960 nm in addition to small valley at 840 nm. The absorption valleys around 500 and 690 represent carotenoids and chlorophyll pigments which represent the color characteristics in the fruit (Abbott et al., 1997); meanwhile, absorption valleys in the NIR at 840 and 960 nm represent sugar and water absorption bands. The deeper the valley the higher the content of this attribute in the fruit. Therefore, cultivar Granny Smith (Green background color) has the deepest valley at the chlorophyll absorption band (at 590 nm) meaning that it has the highest chlorophyll content as shown in fig. 2a. Meanwhile, the cultivar Golden Delicious (Yellow background color) has the lowest content of chlorophyll due to its shallow valley at the chlorophyll absorption band compared with the other cultivars. Within normal surface, more variability of relative reflectance from the same variety was found. This result was possibly caused by the variability of shape and size of fruit samples, even though calibration process extremely decreased this variability.

The visual inspection of spectral curves shown in figures (2a-d) reveal that most of the reflectance of stems was very low compared with the normal surface and it is occasionally overlapped with the normal surface in the visible region in case of Granny Smith and Golden Delicious cultivars. Therefore, the detection of stem in case of these cultivars could be performed using the ordinary RGB camera by conventional machine vision technique that utilizes only the visible range of the electromagnetic spectrum. In case of Red Delicious and McIntosh cultivars, distinguishing stems from normal surface is rather difficult in the visible region due to intersections among their reflectance spectra in the visible region because stems and normal surface have the same color. However, it was observed that the reflectance of stems changed drastically and became more significance without any over-lapping taking place in the near infrared region between 700 to 900 nm for all cultivars. The most optimistic thing is that the reflectance of stem did not changed over time and the same pattern

of the reflectance characteristics was observed for stem after 0, 1, 3 and 7days, which had much lower reflectance than normal tissue over the spectral region between 700 and 900 nm. Accordingly, the reflectance characteristic curves indicates that the spectral region between 700 and 900 nm would be appropriate for stem detection because there was no confusion between normal surface and stem spectra in this region.

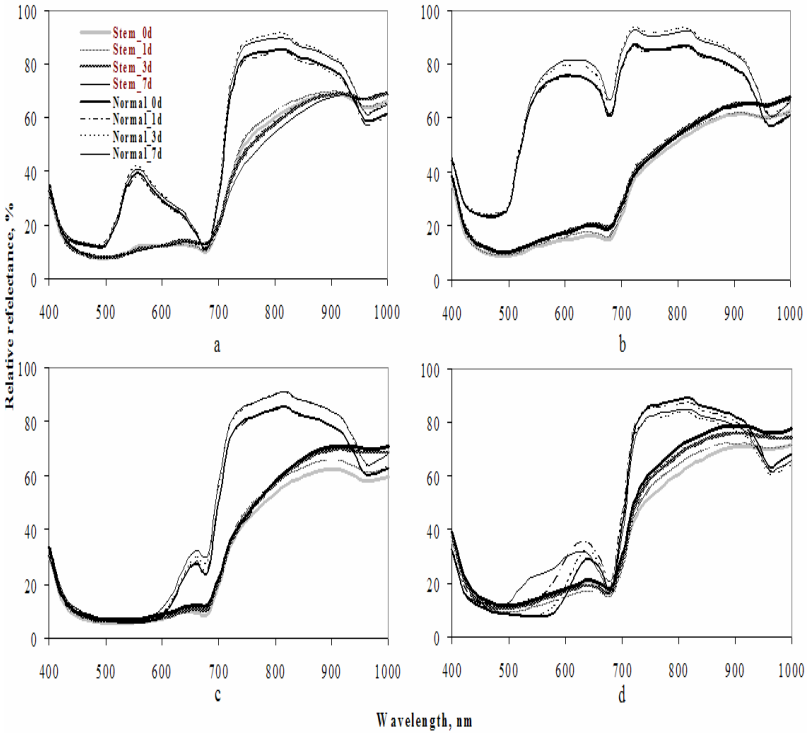

Fig. 2 Spectral feature of normal surface and stem at 0, 1, 3, 7 days for the cultivars (a) Granny Smith, (b) Golden Delicious, (c) Red Delicious, (d) McIntosh

2. Selection of Optimal Wavelengths

Variable Importance for Projection (VIP) scores extracted from PLS analysis of spectral data of stem and normal surface are depicted in fig.4a. If a predictor (wavelength) has a high regression coefficient (in absolute value) and a high value of VIP (Bjarnestad and Dahlman, 2002), then it is a prime candidate for stem discrimination from the normal surface. In general, a VIP threshold value of 1.00 was used as a criterion for corresponding highly influential predictors/optimal wavelengths.

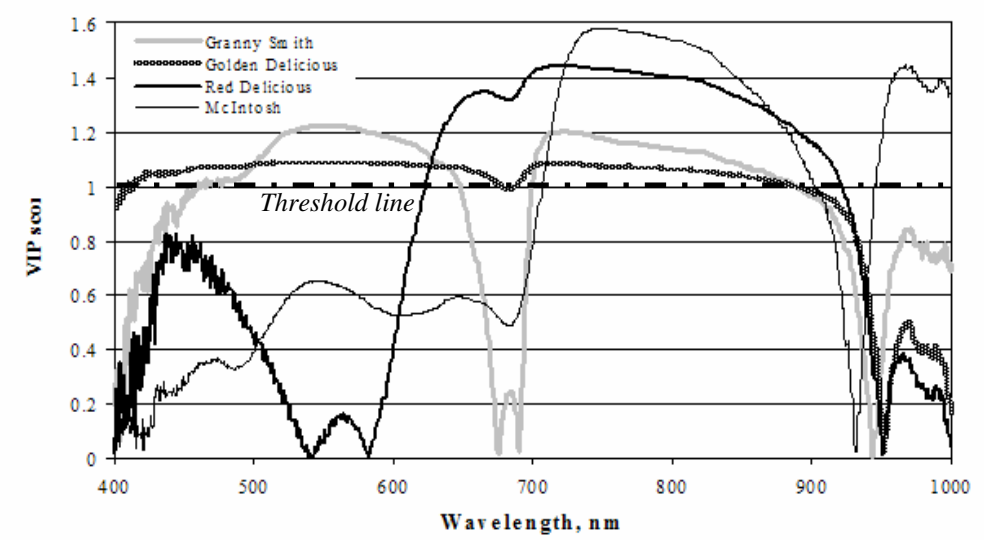

Fig. 3 Variables Importance in Projection (VIP) scores resulting from PLS analysis of spectral reflectance data of stem and normal surface of different cultivars.

The wavelengths in the broadband 460-650 and 700-890 nm are optimal wavelengths for stem detection in case of Granny smith cultivar. Meanwhile, the wavebands 410-895, 620-920, and 705-905 nm are optimal wavelengths for stem detection in Golden Delicious, Red Delicious and McIntosh respectively. Accordingly, figure (3) reveals that, depending on VIP threshold at 1.00, the wavelengths in the broadband of the NIR region from 705 to 890 are all effective in projection and then in stem discrimination in all cultivars. Within this range, wavelengths in the range from 730 to 850 nm corresponded to the highest VIP scores for all cultivars. Based on the previous spectral data analysis and the highest values of VIP scores, three wavelengths 730, 770, and 850 were chosen for stem detection purposes. The selected wavelengths are in the NIR region and can be used for stem detection despite the cultivar. The obvious advantage of working in the NIR range is that problem caused by color variations on normal surfaces can be circumvented

3. Stem Detection Algorithm

The stem detection algorithm is divided into three stages. The first stage is called the primary stage and it is composed of creating a binary mask to produce an image containing only the fruit, avoiding any interference from

the background that could reduce discrimination efficiency. This was done using Matlab (Release 14, The MathWorks Inc., Natick, MA, USA) by thresholding which determines the threshold value based on a histogram of the grey values in the image. Image at 500 nm was used for this task because the fruit appeared opaque compared with the background and can be segmented easily by thresholding. The thresholded image showed the position of the apple in the image as a black area on a white background. This black and white image was used as a template to remove the background in the original image and to focus all image processing operations only in the black area representing the fruit and avoiding any discrepancy with the background region. The second stage included two primary steps, the first and the most important one is the extraction of the spectral features of stem and normal surface, and then selecting the optimal wavelength depending on the VIP scores as described before. The second step in this stage is to reduce the 826 spectral images for each sample to form a set of image at only three optimal wavelengths (730, 770, 850 nm) using the binary mask to separate the fruit area as the main object in the image. These three spectral images will be the base for the next image analysis processes. In these three images, stem's pixels were generally darker than the normal tissue's pixels. In most cases the simple thresholding was not be able to identify all of the area representing the stem, due to variations in the grey level within the stem area and the surrounding surface. So, the solution for this problem is to look for an alternative method to overcome this discrepancy utilizing the spectral signature of the stem in the three image altogether.

The third stage and the final one comprised three steps as follows:

(a) Extracting the spectral signature of the stem at only the optimal wavelengths. The minimum, maximum and average reflectance characteristics of stem for all calibration set fruits (200 images) were recorded and its covariance matrix was calculated. All of these statistics were saved in a file to represent the unique spectral signature of the stem. This saved file will be used for the next step of classification.

(b) Conducting the Maximum Likelihood Classification (MLC) as described above utilizing the reflectance from for each pixel in the three images and from the saved unique spectral signature of the stem. According to this

approach each pixel has a spectral vector associated with it and pixels are segmented into the class they most closely resemble spectrally. Each pixel is assigned to the class that has the highest probability (maximum likelihood) according to the discriminant function for each pixel as described earlier. In case there were no stem and/or calyxes in the image, the resulting classification image will be blank indicating that the tested fruit did not present any stem at all.

(c) Finally, the noise was removed by some morphological operations as median filtering, in addition to erosions and dilations to remove the separated pixels or zones that might carry the same features. If the stem is detected, its direction and lengths can be estimated for fruit quality evaluation purpose or for precise orientation in the processing lines.

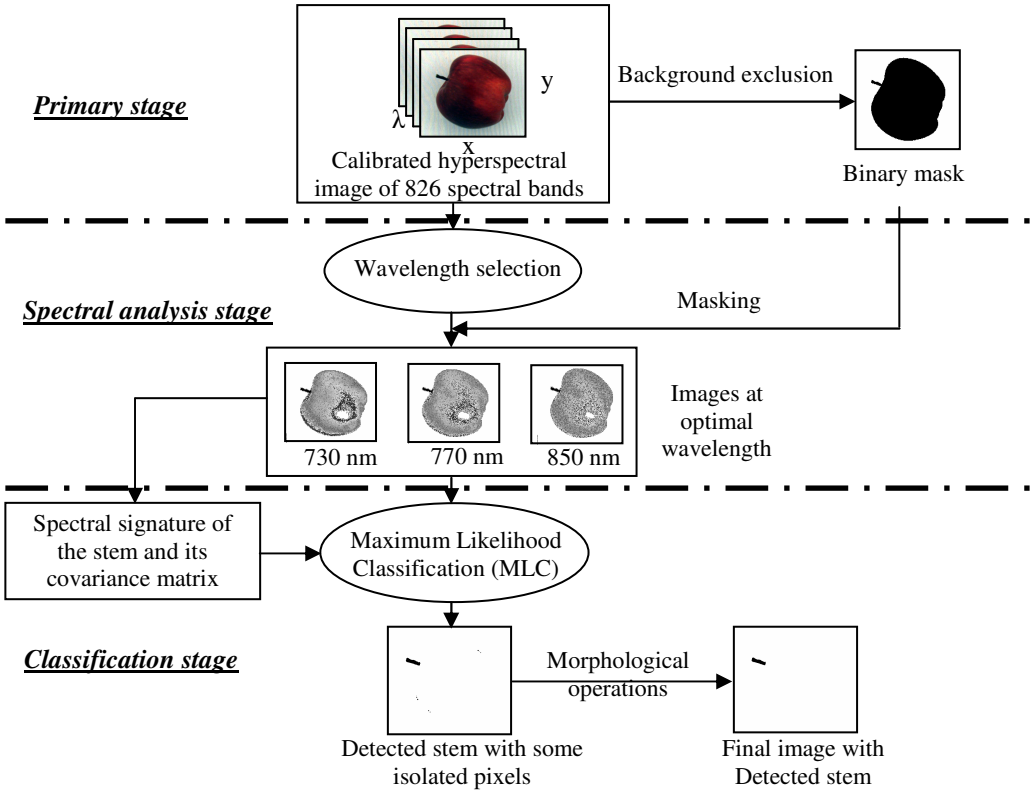

Fig. 4 Flow chart representing the sequences and stages of classification algorithm used for stem detection.

4. Validation of Stem Detection Algorithm

Due to the fact that apples are rotating during sorting and grading processes to inspect the whole surface, the previous algorithms should be validated for detecting stem that might have different lengths and different orientations to know the capability of this method for stem recognition at different positions. The problem becomes very complex when a standing stem appears as a protrusion or the small spikes of a calyx appear on the boundary of the images (Yang, 1996). The validation image set consisted of 240 images containing fruits without stem, images with short stem, images with long and thick stem, and finally images with branched stem as shown in figure (5). Images were acquired for Granny Smith, Golden Delicious, Red Delicious and McIntosh cultivars with stem distributed in different orientations (stem in the top position, stem in the right position, and stem in the down left position). It is important to indicate that each tested image is actually consisted of three sub-images at the selected optimal wavelengths (730, 770, 850 nm). Because the selected optimal wavelengths are in the NIR region, the color variations between varieties are neglected in this region. Figure (5) demonstrates the outcomes of the proposed algorithm for stem detection of different lengths and orientations. The first row of each group represented the original NIR image at optimal wavelengths (only one image of the three sub-images is presented for simplicity to visualize the fruit before applying the classification) and the second row showed the final result of maximum likelihood classification algorithm. It is obvious that the algorithm had great capabilities for stem recognition at different orientations and different lengths and thickness due to its high efficiency for identifying and isolating most of the pixels that carry stem signature. The potential usage of this process is so important in industrial application because it is desired to exclude all fruits without stem because absence of stem opens a way for possible invasion of infections during transport and storage (Ruiz et al., 1996) as well as excluding that ones having very sharp or stiff stem that might puncture adjacent fruits in the final package.

The obtained results for the validation set images were 100% of succession in stem detection. There were very few pixels in the stem area were misclassified, but these few pixels did not affect the performance of classification algorithm.

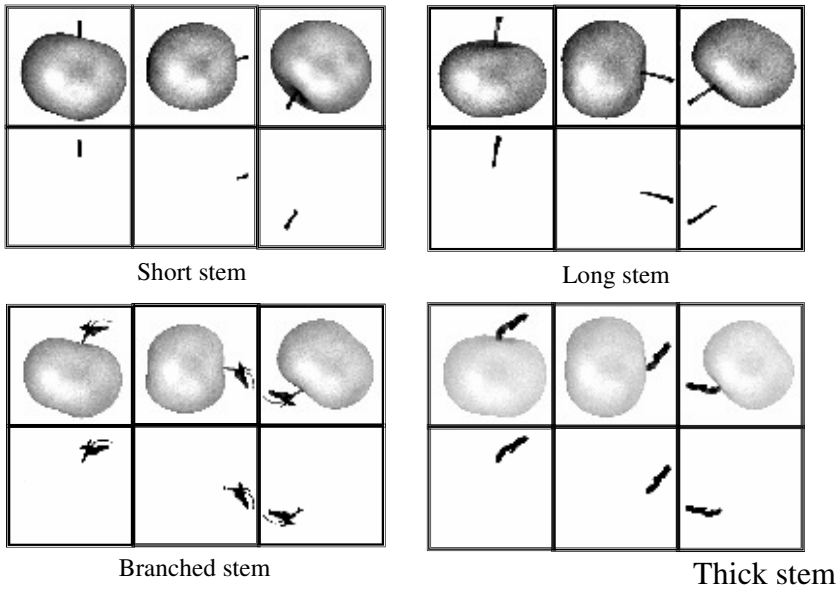

Fig. 5 Stem detection at various orientations with stem located at top, right, and down left positions for short, long, branched and thick stem. The first row of each group represented the original NIR image and the second row shows the final result of maximum likelihood classification algorithm.

CONCLUSION

Hyperspectral imaging technique in visible and near infrared (VIS/NIR) range at 400-1000 nm was built for stem detection purposes in Granny Smith, Golden Delicious, Red Delicious and McIntosh apples. Optimal wavelengths that gave the highest contrast between stem and the normal surface were selected by analyzing the spectral data collected from stem and its surrounding normal surface using Partial Least Square (PLS). Three optimal wavelengths (730, 770, 850 nm) were selected depending on Variables Importance in Projection (VIP) scores extracted from PLS. Although each cultivar has its own optimal wavelengths in both VIS and NIR regions, the common wavebands that can be used for stem detection was in the range 705-890 nm in the NIR region to alleviate the problems aroused from color variations of cultivars. In this range, stem could be discriminated from the surrounding normal surface in all tested cultivars. The spectral characteristics of the stem in terms of its minimum, maximum

and average reflectance as well as its covariance matrix are used to reflect the unique spectral signature of the stem. Thereby, Maximum Likelihood Classification (MLC) method was used to classify apple images at these wavelengths to locate the stem using its unique spectral signature. The experimental results demonstrated that the system successfully classified a set of image consisting of 220 images with and without stem. Also, the method was able to locate the stem having various lengths and thickness as well as different orientations.

Using VIP scores for wavelength selection and then MLC for classification opens a new avenue for more optimistic applications in commercial implementations for detecting various objects in the tested produces. Because the fruits used for this study did not suffer any damage, diseases, and/or injuries, further research should be done in order to discriminate between the stem region and some external defects to be satisfactory for practical implementation.

Acknowledgment

The authors gratefully acknowledge the financial support by Egyptian Government Scholarship offered to the first author Mr. Gamal ElMasry.

REFERENCES

Abbott, J.A., Lu, R., Upchurch, B.L., Stroshine, R.L., 1997. Technologies for non-destructive quality evaluation of fruits and vegetables. Hortic. Rev. 20, 1-120.

Bennedsen, B. S., Peterson, D. L., 2004. Identification of apple stem and calyx using unsupervised feature extraction. Transactions of the ASAE, 47(3): 889−894.

Bennedsen, B.S., Peterson, D.L., 2005a. Identifying defects in images of rotating apples . Computers and Electronics in Agriculture 48 (1) 92–102.

Bennedsen, B.S., Peterson, D.L., 2005b. Performance of a System for Apple Surface Defect Identification in Near-infrared Images. Biosystems Engineering 90 (4), 419–431.

Bjarnestad, S, Dahlman, O., 2002. Chemical compositions of hardwood and softwood pulps employing photoacoustic fourier transform infrared spectroscopy in combination with partial least-squares analysis. Anal. Chem. 74: 5851-5858

Brown, G.K., Segerlind. L. J., 1975. Method for detecting bruises in fruit, U. S. Patent No. 3867041, issued February 18, 1975. Commissioner of Patents and Trademarks, Washington, D.C. 20231.

Cheng, X., Tao, Y., Chen, Y.R., Luo, Y., 2003. NIR/MIR dual–sensor machine vision system for online apple stem end/calyx recognition. Trans. ASAE 46(2): 551–558.

Chong, L-G., Jun, C-H., 2005. Performance of some variable selection methods when multicollinearity is present. Chemometrics and Intelligent Laboratory Systems 78 (1) 103–112.

Crowe, T., Delwiche, M., 1996a. Real-time defect detection in fruit— part I: Design concepts and development of prototype hardware. Transactions of the ASAE, 39(6), 2299–2308.

Crowe, T., Delwiche, M., 1996b. Real-time defect detection in fruit— part II: An algorithm and performance of a prototype system. Transactions of the ASAE, 39(6), 2309–2317.

Geladi, P., Kowalski, B.P., 1986a. An example of 2-block predictive partial least-squares regression with simulated data. Anal. Chim. Acta 185 (1):19-32.

Geladi, P., Kowalski, B.P., 1986b. Partial least-squares regression: A tutorial. Anal. Chim. Acta 185 (1):1-17.

Haaland, D.M., Thomas, E.V., 1988a. Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. Anal. Chem. 60: 1193-1202.

Haaland, D.M., Thomas, E.V., 1988b. Partial least-squares methods for spectral analyses. 2. Application to simulated and glass spectral data. Anal. Chem. 1988. 60, 1202-1208.

Hruschka, W.R., 2001. Near Infrared Technology in the Agricultural and Food Industries; Data Analysis: Wavelength Selection Methods. American Association of Cereal Chemists, Edited by Phil Williams and Karl Norris.

Keskin, M., Dodd, R.B., Han, Y.J., Khalilian, A., 2004. Assessing nitrogen content of golf course turfgrass clippings using spectral reflectance. Applied Engineering in Agriculture 20(6): 851–860.

Kim, M. S. , A. M. Lefcourt, K. Chao, Y. R. Chen, I. Kim, D. E. Chan (2002). Multispectral detection of fecal contamination on apples based on hyperspectral imagery: part I. Application of visible and near–infrared reflectance imaging. Trans. ASAE 45(6): 2027–2037.

Kleynen, O., Leemans, V., Destain, M.-F., 2005. Development of a multispectral vision system for the detection of defects on apples. Journal of Food Engineering 69 (1) 41–49.

Li, Q., Wang, M., Gu, W., 2002. Computer vision based system for apple surface defect detection. Computers and Electronics in Agriculture, 36 (2-3) 215-223.

Liu, Y., Windham, W.R., lawrence, K.C., and Park. B., 2003. Simple Algorithms for the Classification of Visible/Near-Infrared and Hyperspectral Imaging Spectra of Chicken Skins, Feces, and Fecal Contaminated Skins. Applied Spectroscopy 57 (12): 1609-1612.

Mehl, P., Chen, M.Y-R, Kim, M.S., Chan, D.E., 2004. Development of hyperspectral imaging technique for the detection of apple surface defects and contaminations. Journal of Food Engineering 61 (1) 67–81.

Olah, M., Bologa, C. and Oprea, T.I., 2004. An automated PLS search for biologically relevant QSAR descriptors. Journal of Computer-Aided Molecular Design 18: 437–449.

Osborne, S.D., Jordan, R.B., Künnemeyera, R., 1997. Method of wavelength selection for partial least squares. Analyst, 122: 1531–1537.

Penman, D.W., 2001. Determination of stem and calyx location on apples using automatic visual inspection. Computers and Electronics in Agriculture, 33 (1): 7–18.

Reid, W.S., 1976. Optical Detection of Apple Skin, Bruise, Flesh, Stem and Calyx. J. Agric. Eng. Res., 21(3): 291-295.

Ruiz, L.A., Molto, E., Juste; F., Pla, F., Valiente, R., 1996. Location and characterization of the stem–calyx area on oranges by computer vision. J. Agric. Eng Res., 64(1):165–172.

Throop J.A., Aneshansley D.J., Anger W.C., Peterson D. L., 2003. Quality evaluation of apples based on surface defects - An inspection station design. ASAE Paper No. 03-6161 ASAE Meeting, Las Vegas, Nevada, USA.

Unay, D., Gosselin, B., 2005. Stem and calyx recognition on 'Jonagold' apples by pattern recognition. Journal of Food Engineering (In Press).

Wen, Z., Tao, Y., 1999. Building a rule-based machine-vision system for defect inspection on apple sorting and packing lines. Expert Systems with Applications 16: 307–313.

Xing, J., De Baerdemaeker, J., 2005. Bruise detection on 'Jonagold' apples using hyperspectral imaging. Postharvest Biology and Technology 37 (1) 152–162.

Yang, Q., 1996. Apple Stem and Calyx Identification with Machine Vision. J . Agric . Engng Res . 63(2), 229–236.

Ying, Y., Jing, H., Tao,Y., Zhang, N., 2003. Detecting stem and shape of pears using fourier transformation and an artificial neural network. Transactions of the ASAE, 46(1): 157–162.

الملخص العربي

تحديد أعناق التفاح باستخدام التصوير الطيفى

م/جمال المصرى[1] د/شريف محمد عبد الحق رضوان[1] د/محمد صلاح الأمير[1] د/ نينج وانج[2]

ان الجزء الحرج من تطوير آلات أوتوماتيكية لفرز وتدريج ثمار التفاح هو مقدرة هذه الآلات على تحديد مكان العيوب والأعناق فى الثمار. ،على هذا فقد تم تنفيذ نظام تصوير طيفى يعمل فى الجزء المرئى والقريب من الأشعة التحت حمراء (VIS/NIR) من الطيف الضوئى فى المدى من ٤٠٠-١٠٠٠ ن.م بغرض تحديد والتعرف على الأعناق فى أربعة أصناف مختلفة من التفاح وهى (أخضر اللون) Granny Smith و (أصفر اللون) Golden Delicious (أحمر قاتم) Red Delicious و (متعدد الألوان) McIntosh.

ولقد تم تجميع ومن ثم تحليل الإستجابة الطيفية لكل من العنق والسطح العادى للثمرة ومنها تم تحديد الأطوال الموجية المثالية Optimal wavelengths والتى عندها يوجد تميز واضح بين السطح العادى للثمرة والعنق وذلك باستخدام العوامل المهمة (VIP) Variables Important in Projection والمستخلصة من التحليل الإحصائى للإستجابات الطيفية بطريقة أقل المربعات الجزئية Partial Least Squares (PLS) وكانت الأطوال الموجية المثالية هى ٧٣٠، ٧٧٠، ٨٥٠ وبتطبيق طريقة التصنيف بأقصى احتمال Maximum Likelihood Classification (MLC) على الصور الطيفية عند هذه الأطوال الموجية المثالية أمكن معرفة وتصنيف الثمار على أساس احتوائها على عنق من عدمه فضلاً عن تحديد طول هذا العنق وسمكه واتجاهه.

ولقد أوضحت النتائج التجريبية نجاح هذا النظام فى تصنيف مجموعة صور للاختبار تتألف من ٢٠٠ صورة وكذلك مجموعة صور أخرى تأكيدية تتألف من ٢٤٠ صورة باستخدام الأطوال الموجية المثالية فقط.

[1] قسم الهندسة الزراعية - كلية الزراعة - جامعة قناة السويس- الإسماعيلية

[2] قسم هندسة المصادر الحيوية - جامعة ماكجيل - كندا