

## ABSTRACT

Title of Document: Multispectral Method for Apple Defect  
Detection using Hyperspectral Imaging System

Tao Tao, Master of Science, 2011

Directed By: Dr. Gang Qu, Associate Professor

Hyperspectral imaging is a non-destructive detection technology and a powerful analytical tool that integrates conventional imaging and spectroscopy to get both spatial and spectral information from the objects for food safety and quality analysis. A recently developed hyperspectral imaging system was used to investigate the wavelength between 530nm and 835nm to detect defects on Red Delicious apples. The combination of band ratio method and relative intensity method were developed in this paper, which using the multispectral wavebands selected from hyperspectral images. The results showed that the hyperspectral imaging system with the properly developed multispectral method could generally identify 95% of the defects on apple surface accurately. The developed algorithms could help enhance food safety and protect public health while reducing human error and labor cost for food industry.

MULTISPECTRAL METHOD FOR APPLE DEFECT DETECTION  
USING HYPERSPECTRAL IMAGING SYSTEM

By

TAO TAO

Thesis submitted to the Faculty of the Graduate School of the  
University of Maryland, College Park, in partial fulfillment  
of the requirements for the degree of  
Master of Science  
2011

Advisory Committee:

Dr. Gang Qu, Chair  
Dr. Moon Kim  
Dr. Robert Newcomb  
Dr. Chun Chieh Yang

© Copyright by  
Tao Tao  
2011

## **Acknowledgements**

I would like to thank the Environmental Microbiological and Food Safety Laboratory (EMFSL), Agricultural Research Service (ARS), United States Department of Agriculture and the Department of Electrical and Computer Engineering (ECE) at the University of Maryland, College Park (UMD) for sponsoring this work. I greatly acknowledge my advisor Dr. Gang Qu (ECE, UMD), who has supported my graduate studies. I am very thankful to Dr. Moon S. Kim (ARS, USDA), for his scientific discussions and support as well as educating me about the world of scientific research; Dr. Chun-Chieh Yang (ARS, USDA), who has trained me about specific details on research and the English writing style. I would also thank Dr. Robert Newcomb (ECE, UMD) and Dr. Yang Tao (UMD), who have given many kind suggestions on my thesis. The research work at the EMFSL introduced me to the field of spectroscopy sensing for imaging techniques on food safety and helped me to learn valuable background knowledge on this thesis work. I would also like to acknowledge other colleagues at the EMFSL, in particular Dr. Jianwei Qin, for discussing the data with me. I am also thankful to the research support staffs Ms. Diane Chan and Mr. Frank Gwozdz. Finally, I would like to express my appreciation to all the members of the embedded system laboratory at UMD.

# Table Content

Acknowledgements .....	ii
Table Content .....	iii
List of Tables.....	v
List of Figures .....	vi
Chapter 1 Introduction .....	1
1.1 Introduction.....	1
1.2 Thesis Organization.....	2
Chapter 2 Literature Review .....	5
2.1 Background .....	5
2.1.1 Importance of Food Safety and Quality in Fresh Produce.....	5
2.1.2 Fresh Sorting and Grading .....	6
2.1.3 Quality Inspection using Non-destructive Methods.....	7
2.1.3.1 Apple Inspection .....	8
2.1.3.2 Inspection on Other Fruits, Meat and Vegetables.....	9
2.2 Machine Vision Systems.....	11
2.2.1 Machine Vision Introduction and Future Trend.....	11
2.2.2 Machine Vision as Real-Time System.....	11
2.3 Hyperspectral and Multispectral Imaging and Sensing Techniques.....	12
2.4 Line-Scanning Hyperspectral Imaging System.....	14
2.5 Hyperspectral and Multispectral Imaging Data Analysis Methods .....	15
2.5.1 Band Ratio Method .....	15
2.5.2 Other Classification Methods with Hyperspectral Data.....	17
2.5.2.1 Principal Component Analysis.....	17
2.5.2.2 Other Classification Methods.....	19
Chapter 3 Apple Defect Detection using Hyperspectral Analysis .....	21
3.1 Experimental System and Sample Materials.....	21
3.1.1 Hyperspectral Imaging System and Image Data Acquisition .....	21
3.1.2 Sample Materials.....	22
3.2 Hyperspectral Data Analysis Methods for Apple Defect Detection .....	23
3.2.1 Reflectance Spectrum Analysis.....	23
3.2.2 Band Selection.....	29
3.2.3 Threshold Selection.....	30
Chapter 4 Results .....	36

4.1 Relative Intensity Method.....	36
4.2 Band Ratio Result.....	37
4.3 Band Ratios and Relative Intensity Combination.....	39
4.4 Discussion.....	41
Chapter 5 Conclusion.....	44
Chapter 6 Future Study.....	46
Appendix A.....	47
Appendix B.....	51
Reference.....	60

## **List of Tables**

Table 1. Classification groups of the 169 sample apples. ....	23
Table 2. Sample pixels for calibration of different classification groups .....	23
Table 3. The defects detection result using three different thresholds on selected training samples.....	33
Table 4. The defects detection result of several methods on defective apples.....	39

## List of Figures

Figure 1 Hyperspectral Cube. This cube presents the data as a volume, composed of the spatial resolution(x, y) and the number of contiguous spectral bands ( $\lambda$ ). (Kim, 2001)..... 13

Figure 2 The critical components of the line-scan imaging platform (Kim, 2009) ..... 14

Figure 3 Reflectance spectra of sample image pixels from different surface of apples. 24

Figure 4 The combined method scheme of  $R_1$  &  $R_2$  &  $I$  method ..... 27

Figure 5 The typical defect apple image at 779nm wavelength..... 29

Figure 6 Average intensity value of sample pixels at 779nm wavelength of defect area, good surface on defect apples and normal apple surface. .... 31

Figure 7 (a) (b) Average of band ratio value of the sample apple image pixel..... 32

Figure 8 The binary image using relative intensity method at 779nm waveband..... 36

Figure 9 Image result of band ratio between 577 and 643nm. The first row is apple image at 577nm; the second row is apple image at 643nm; the binary images of ratio (577nm/643nm) results is shown in the third row. .... 37

Figure 10 Image result of band ratio between 676 and 709nm. The first row is apple image at 657nm; the second row is apple image at 676nm; binary images of ratio (657nm/676nm) results is shown in the third row..... 38

Figure 11 Detection results of the entire evaluation apples using ' $R_1$ || $R_2$ ' and ' $R_1$  &  $R_2$  &  $I$ ' method. .... 40

Figure 12 Image result of combined method using band ratio and intensity. The first row is apple image at 779nm; the second row is binary image result of proposed combined method. .... 41



Figure 13 Image results of band ratio combined method and proposed band ratio and intensity combined method on good apples; the first row is the apple image at 779nm, the second row is binary image detection result of combined $R_1$ and $R_2$ ; the third row is binary image result of the proposed combined method. ....	42
Figure 14 Four apple image results that misclassified as good one using ‘ $R_1$ & $R_2$ & I’ method.....	43
Figure 15 Loading plots of the average five selected ‘Red Delicious’ Apples.....	47
Figure 16 Principal component results by PCA methods .....	49
Figure 17 Combination of normalized images (upper), and after thresholding.....	50

# Chapter 1 Introduction

## 1.1 Introduction

The United States produces over 9.48 billion pounds of apples annually (Knopf, 2010), and 148.7 million boxes of apples are packed for the fresh fruit market. (USDA ERS, 2009). In the apple packing industry, fruits are checked either visually by human or sorting machines for quality control before shipping to consumers. In a typical apple packing house, workers are placed along an apple conveyor to inspect the passing apples and remove the ones that are injured, rotten, diseased, bruised or with other defects. After defect inspection, apples are transferred to another line for cleaning, waxing and drying. The final grading lines sorts the apples by their size, color, and shape, and then pack them into different boxes based on the grade.

Some of the sorting procedures have already been automated; however, some others are still carried out manually. After working at the packing line for many hours, even the most experienced workers may feel tired, thus their efficiency may reduce dramatically. Because many apple defects such as bruises, cuts, sooty blotch, and other physical damage cannot be easily found by eyes, finding a reliable, accurate, and efficient apple defect sorting system will be very valuable and has been very challenging(Xing and Baerdemaeker, 2005).

Automated defect detection system has the advantages of overcoming the drawbacks of manual inspections in reducing the subjectivity and reducing tediousness. Optical

imaging techniques are one of the major non-destructive fruits inspection methods that are used in machine vision systems. Spectral reflectance imaging has been widely used for assessing quality aspects of agricultural products (Kavdir and Guyer, 2002). However, the development of a practical system for automated defect detection sorting still encounters limitations in accuracy.

Hyperspectral line-scan imaging techniques combine conventional imaging techniques and spectroscopy to acquire both spatial and spectral information from an object, which show discrete spectral features of the target objects. However, hyperspectral image data is too complicated for processing because it has high volume of data; therefore, practical multispectral techniques are used in real life applications. Moreover, the combination of multiple bands with spectral information leads to better classification than using single band. In this research, we developed a novel approach that combines band ratio method and relative intensity method for defect detection.

## **1.2 Thesis Organization**

The overall objective of this research is to develop multispectral algorithm derived from hyperspectral line-scan images for automatic defect apple inspection. The optimal band selection and band combination are applied to differentiate the defect area from the normal apple skins, the calyx and the stem end. The main goals of this study are:

1. To evaluate the potential of a hyperspectral imaging in detection of the defects on apples by the use of a visible/NIR technique.
2. To select the optimal band wavelengths to use for a band ratio algorithm, then to compare the single band ratio and intensity classification results with the combined classification results.

Experiments and empirical studies were conducted during the period of 2008 -2011 in the Environmental Microbial and Food Safety Laboratory to demonstrate that the proposed method meets these objectives.

This thesis contains six chapters. Chapter 1 gives the introduction to the existing apple sorting limitations, hyperspectral imaging system use for apple sorting and defect detection; lists the general objectives and methods to solve the existing manually apple defects detection drawbacks; Chapter 2 reviews literatures to our research work and identifies how current approaches can be improved; Chapter 3 elaborates our multiple-band spectral combination method with visible and near-infrared bands information for apple defect detection; Chapter 4 discusses the results of apple defects detection experiments; Chapter 5 gives the conclusion of the thesis results derived by our experiments and researches; Chapter 6 is the future study plan. The appendix reports the results of Principal Component Analysis (PCA) on hyperspectral image and the image results using proposed method

At EMFSL, I have also conducted a study on principal component analysis on hyperspectral image. This study focuses on apple defects detection, the key contribution is that I found that PCA method is more complex in computation and time-consuming in processing, even though its defect detection rate is less than the proposed method in this thesis. This helps to choose the simple and effective classification method for future online processing.

## **Chapter 2 Literature Review**

### **2.1 Background**

#### **2.1.1 Importance of Food Safety and Quality in Fresh Produce**

Fresh fruits and vegetables provide a variety of health benefits to our daily life. They contain vitamins, minerals, and many elements that help prevent illnesses such as cancer, heart disease, and stroke. A daily diet of fresh produce is highly recommended by health and nutrition authorities. For the past decade, American's consumption of fresh vegetables and fruits has been increased every year. Consequently, this increase has no doubt raised the public concerns regarding the potential safety and quality issues of the fresh produce.

Quality and safety are among the most important criteria for the evaluation of consumable fresh fruits and vegetables. Generally, quality includes external factors such as appearance (size, shape, color, gloss, and consistency), texture (firmness, crispness and toughness), and flavor (sweetness, sourness). These quality aspects are important because consumers are susceptible to produce contamination and cross-contamination that may occur during the packing process. Poor quality of the fresh produce will have a negative effect on peoples' health; thus, the development of effective fruit inspection technologies to ensure the quality of fruits and vegetables is essential for competition in the marketplace (Wikipedia).

### **2.1.2 Fresh Sorting and Grading**

The quality of fruits is affected by various factors such as conditions of growing, storage, and handling. In a typical apple processing factory, workers are employed to inspect and remove defective apples, such as those with rots, bruises, injuries, and other defects must be removed at the first stage to prevent cross-contamination and reduce subsequent processing cost. . At the final stage, apples are sorted according to their size, color and shape, and then packed into boxes according to their grades. In some large packing house operations, sorting machines using machine vision technologies are installed to sort apples into different grades based on weight, size, shape, color, defects and other parameters. Storage can keep freshness of apples such as Red Delicious, and Golden Delicious for several months (Tao, Buchanan, Song, et al. 2002).

“The United States Department of Agriculture defines the quality standard of apples into several grades: “U.S. Extra Fancy”, “U.S Fancy”, “U.S. No.1”, “U.S.No.1 Hail” and “U.S. Utility”. The “U.S. Extra Fancy” consists of apples of one variety which are mature but not overripe, clean, fairly well formed, and are free from decay, internal browning, internal breakdown, soft scald, scab, freezing injury, visible water core, and broken skins. These apples are also free from injury caused by bruised, brown surface discoloration, smooth net like russeting, sunburn or spray burn, limb rubs, hail, drought spots, scars, disease, insects, or other defects.“U.S. Fancy” consists of apples of one variety (except when more than one variety is printed on the

container) which are mature but not overripe, clean, fairly well formed, and free from decay, internal browning, internal breakdown, soft scald, freezing injury, visible water core, and broken skins. “U.S. No. 1” consists of apples which meet the requirements of U.S. Fancy grade except for color, russeting, and invisible water core. “U.S.No.1 Hail” consists of apples which meet the requirements of U.S No. 1 grade except that hail marks where the skin has not been broken and well healed hail marks where the skin has been broken, are permitted, provided the apples are fairly well formed. “U.S. Utility” consists of apples of one variety (except when more than one variety is printed on the container) which are mature but not overripe, not seriously deformed and free from decay, internal browning, internal breakdown, soft scald, and freezing injury. The apples are also free from serious damage caused by dirt or other foreign matter, broken skins, bruises, brown surface discoloration, russeting, sunburn or spray burn, limb rubs, hail, drought spots, scars, stem or calyx cracks, visible water core, bitter pit or Jonathan spot, disease, insects, or other means.” (USDA, 2002)

### **2.1.3 Quality Inspection using Non-destructive Methods**

It is essential for food factory and large market to reduce disease-causing hazards before fresh produce reaches the consumer. In order to guarantee a safe and wholesome product, fruit and vegetable producers are using various methods to improve safety and quality of their products. Since pathogens from fresh products cannot be completely removed using current washing/sanitization methods, the most effective way to minimize food safety risks is to identify and remove contaminated



raw materials from the product stream, prior to processing or fresh-cut preparation, using noninvasive on-line inspection methods that can identify fecal contamination and reduce human errors (Tao, Buchanan, Song, et al. 2002).

Non-destructive inspection methods have been widely used in research to monitor quality and safety attributes of fresh produce. The ability of detecting and classifying fecal contamination and physical damage in fresh produce could highlight produce with a high risk of contamination and alert producers before the product reaches consumers.

#### **2.1.3.1 Apple Inspection**

A method that incorporates a near-infrared (NIR) camera and a mid-infrared camera was developed for simultaneous imaging the fruit being inspected (Wen and Tao, 2000). The NIR camera is in the wavelength range of 700-1000 nm and the MIR camera is in the range of 3.4-5 $\mu$ m. The final image was obtained by subtracting the MIR image from the NIR image. This image shows true defects such as bruises, rots, and limb rubs. A 98.86% recognition rate for stem-ends and a 99.34% recognition rate for calyxes were achieved using such dual-camera NIR/MIR machine vision defect sorting system. However, the high cost of NIR cameras still limits its practical implementation for apple packing house applications to date.

Another research group (Kim, Chen, Mehl, et al. 2001) developed a laboratory-based hyperspectral imaging system which is capable of performing both reflectance and fluorescence imaging. It uses dual illumination sources where fluorescence emissions

are measured with ultraviolet (UV–A peaked at 365nm) excitation. The spectra imaging range is from 430nm to 930nm with spectral resolution of approximately 10nm. This system provided the capability of assessing quality and safety attributes of food products. One group (Kim, Chen, Lefcourt, et al. 2007) has integrated an online line-scan system with a commercial apple-sorting machine and used it to evaluate and to inspect apples with fecal contamination and defects. An NIR band ratio (BR) classification method is used to achieve 99.5% apple defect classification accuracy with a false positive rate of only 2%. Besides, the NIR processing regime overcomes the presence of stem/calyx on apples that typically has been a problematic source for false positives in the detection of defects.

A systematic approach with Partial Least Squares (PLS) regression as well as stepwise Discrimination Analysis (DA) based on a spectral region between 400 and 1,000nm was used by ElMasry (2007). Three effective wavelengths in the near infrared region (750, 820, 960 nm) were selected to realize multispectral imaging tests. In comparison with other similar research, the results of this investigation indicated that this technique could be used to detect bruises on apple surfaces in the early stage of bruising.

#### **2.1.3.2 Inspection on Other Fruits, Meat and Vegetables**

Liu (Liu, Chen, Wang, et al. 2006) suggests the large spectral differences between good-smooth skins and chilling injured skins of cucumbers occurred in the 700 to 850nm visible/NIR region. Results revealed that using either a dual-band ratio

algorithms ( $R_{811/756}$ ) or PCA model from a narrow spectral region of 733 to 848nm could detect chilling injured skins with a success rate of over 90%.

Light scattering is related to the structural characteristics of fruit and hence is potentially useful for estimating fruit firmness. A hyperspectral imaging system was used to acquire 153 scattering profiles from 'Red Haven' and 'Coral Star' peaches between 500 and 1000nm. A combination of 10 or 11 wavelengths was used to obtain best predictions of fruit firmness, and it was obtained with values for  $r^2$  of 0.77 and 0.58 for 'Red Haven' and 'Coral Star' peaches (Lu and Peng, 2006).

Hyperspectral imaging in the visible and near-infrared (400-1000nm) regions was tested for nondestructive determination of Moisture Content (MC), Total Soluble Solid (TSS), and acidity in strawberry (ELMasry, Wang, Elsayed, et al. 2007). By using the Partial Least Squares analysis, a correlation coefficient was calculated to predict MC, TSS, and acidity using Multiple Linear Regression (MLR) model. A classification accuracy of 89.61% was achieved.

Spectral Angle Mapper (SAM) supervised classification method for hyperspectral poultry imagery was performed for classifying fecal and ingesta contaminants on the surface of broiler carcasses. The SAM classifier using reflectance of hyperspectral data with 512 narrow bands from 400nm to 900nm to get the overall mean accuracy was 90.13%. (Park, Windham, Lawrence, et al. 2007).

## **2.2 Machine Vision Systems**

### **2.2.1 Machine Vision Introduction and Future Trend**

Machine vision is the ability of a computer to “see”. A machine-vision system employs one or more video cameras, analog to digital conversion, and digital signal processing. As an integrated mechanical-electronic-optical-software system, machine vision has many applications, such as quality assurance, sorting, material handling, robot guidance, and calibration. Machine vision processes target at “recognizing” the actual objects in an image and assigning properties to those objects by understanding what they mean (Giraldo, 2006).

One major challenge to machine vision that uses in natural product inspection is the product variability because natural products have different shapes, sizes, textures and colors in the image data, and also the defects or attributes of the products vary in term of the environment conditions. Therefore, all the factors related to the mechanical, electronic, optical and software should be considered in the design of a machine vision system (Giraldo, 2006).

### **2.2.2 Machine Vision as Real-Time System**

The speed of machine vision system used for fruits and vegetables inspection becomes more important in a real-time machine vision system. Real-time capability includes two aspects: at high speed and in real-time. The meaning for high speed is that the vision system can take image data at a high rate. Real-time means the system will be timely (the results will be provided when they are needed). In a typical

machine application where the part stops for viewing and is not advanced until the vision system has made its decision, there is a signal value to the vision system making a decision by a certain time so that it is not the slowest element of the production process. If the vision system is earlier than this time, its value is not increased because the production line cannot go any faster. If the vision system is always on time, but very occasionally a little slower, its value might be diminished only slightly, however, as the vision system becomes slower on average or more frequently, it becomes a bottleneck, and its value declines eventually becoming negative, such a system would be called soft real time. Although high-speed and real-time are technically different criteria, they are most often found together (Perry West, Automated Vision System, Inc).

### **2.3 Hyperspectral and Multispectral Imaging and Sensing Techniques**

Hyperspectral imaging collects and processes information from across the electromagnetic spectrum. Unlike other methods that only capture one single spectrum image (monochromatic) or multiple discrete spectral images (multispectral), hyperspectral imaging records a volume of data that contains a complete spectral range for each point in the sample image. As a result, hyperspectral data offers more detailed information about the sample object.

Hyperspectral sensing is used in a wide array of real-life applications. Although originally developed for mining and geology, it has now spread into fields such as ecology and surveillance. This technology is continually becoming more available to

the public, and has been used in a wide variety of ways. Although the cost of acquiring hyperspectral images is typically high, hyperspectral remote sensing is used more and more for monitoring the development and health of agricultural products (Wikipedia).

A Hyperspectral Imaging System (HIS) generally includes an illumination device, a Charge-Coupled Device camera (CCD), a spectrograph, and a computer that controls data acquisition and processes the data. The spectral region selected depends on the characteristics that the system can identify. A few optical wavelengths are selected to provide useful information about the characteristics of target of interest. These selected wavelengths are used to form the multispectral imaging analysis.

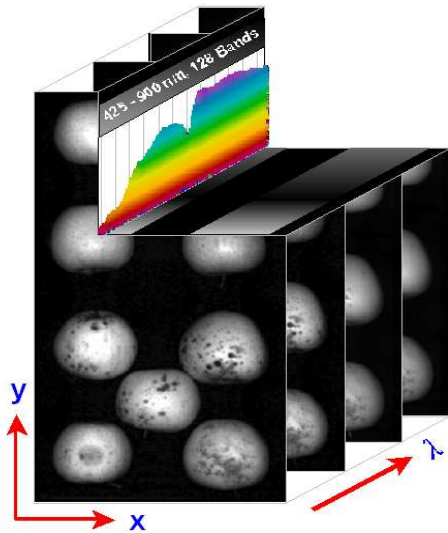


Figure 1 Hyperspectral Cube. This cube presents the data as a volume, composed of the spatial resolution( $x, y$ ) and the number of contiguous spectral bands ( $\lambda$ ). (Kim, 2001)

## 2.4 Line-Scanning Hyperspectral Imaging System

The USDA Environmental Microbial and Food Safety Laboratory (Kim, Chao, Chan et al. 2009) developed a new spectral-imaging technique that combines the advantages of spectroscopy and machine vision in addressing food quality and safety problem. Figure 2 shows the critical components of our line-scan imaging platform. The EMCCD camera is suitable for imaging with very short exposure time, while its hardware performs image binning, waveband selection, and fast pixel readout.

The line-scan imaging platform was used for apple inspection, and also for freshly slaughtered chickens that was tested on processing lines in commercial poultry plants, which helps to develop a multi-tasking inspection system.

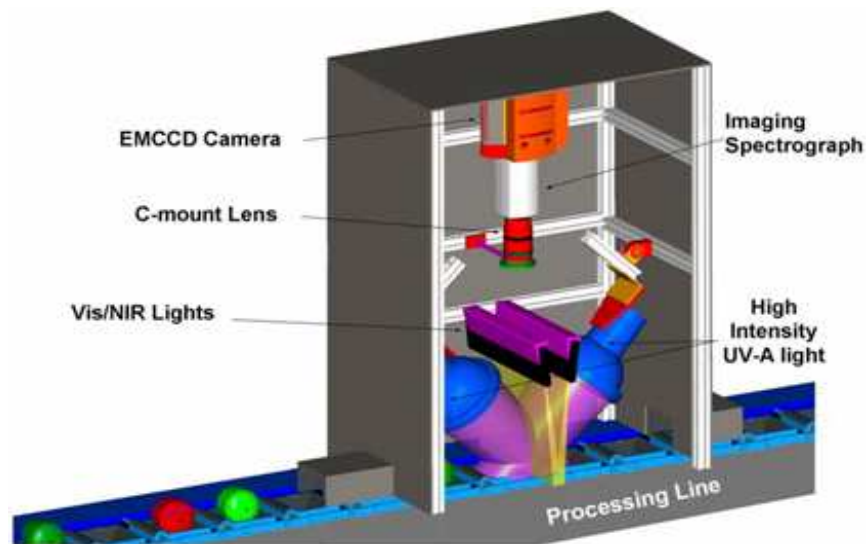


Figure 2 The critical components of the line-scan imaging platform (Kim, 2009)

## **2.5 Hyperspectral and Multispectral Imaging Data Analysis Methods**

Hyperspectral imaging techniques are widely used in food processing and inspection field recently. Many classification approaches have shown their potentials in food safety area such as fecal contamination detection, bruise detection and defects identification in the literature.

To analyze the hyperspectral image data, because the hyperspectral imaging data provide large amount of information, thus we need to extract the useful information that are needed in experiment. A general strategy contains two steps, which are: feature extraction and pattern classification. The most useful classification approaches such as BR, PCA, PLS, SAM, General Discriminant Analysis-- Fisher Discriminant Analysis (FDA) and Linear Discriminant Analysis (LDA) and Artificial Neural Network (ANN) are employed in the literature.

Different mathematical operations provide ground to further enhance the detection in the spectral images. Linear and non-linear combinations of images such as ratios can be used to increase the intensity difference between the object of interest and the background.

### **2.5.1 Band Ratio Method**

Calibrations and data correlation for hyperspectral images is complex due to a multitude of spectral bands. The hyperspectral images were first corrected to remove the effect of dark current of the CCD imager, followed by correction using a standard



white reference measurement, thus the reflectance image was calculated using the following equation (1):

$$I = (I_0 - D) / (W - D) \quad (1)$$

Where  $I_0$  is the sample hyperspectral image,  $D$  is the dark current image, and  $W$  is the white reference image. The corrected images will be the basis for hyperspectral image analysis to the extracted spectral information for band ratio algorithm to detect the defects of apple data.

Kim (Kim, Chen, Cho, et al 2007) found a two-reflectance band ratio to be an efficient multispectral image fusion method for detecting defects on apples. Based on the spectral response, an NIR two-band ratio may provide the greatest difference in ratio values between the normal apple surfaces and defect portions (Liu, Chen, Kim, et al 2007). Large spectral differences between good-smooth skins and chilling injured skins occurred in the 700 to 850nm visible/NIR region. Simple spectral band algorithms was attempted to discriminate the ROI spectra of good cucumber skins from those of chilling injured skins. Results revealed that using the dual-band ratio algorithm ( $R_{811/756}$ ) can detect the chilling-injured skins with a success rate of over 90%.

A band ratio method was applied for fecal contamination detection. Region of interest spectral features of fecal contaminated areas showed a relatively low reflectance compared to normal skin area. Larger spectral differences between contaminated and un-contaminated skins occurred in the 675-950nm visible/NIR region, a dual band

ratio ( $R_{725/811}$ ) algorithm could be used to identify fecal contaminated skins (Liu, Chen, Kim, et al, 2007).

## 2.5.2 Other Classification Methods with Hyperspectral Data

### 2.5.2.1 Principal Component Analysis

In hyperspectral imaging processing, the primary goal for researchers is to extract the useful information among the large amount of data volume, and to evaluate the features from the hyperspectral data. By evaluating these features, they select the most differentiable ones compared with the rest. It is also needed to do dimension reduction through transforming the data to a new set of axes in a new coordinate. The PCA (Campbell, 2002) and LDA (Fukunaga, 1990) are the most commonly used methods to do this transformation.

The PCA method can be formulated by the following equation. The scatter matrix can be expressed as:

$$S_T = \sum_{k=1}^n (x_k - \mu)(x_k - \mu)^T \quad (2)$$

Where  $S_T$  is the covariance matrix,  $\mu$  is the corrected sample data's mean vector,  $x_k$  is  $N^2$  dimensional sample vector and  $n$  is the total number of collected spectral band images. Since the covariance matrix is square, we can calculate the eigenvectors and eigenvalues. In PCA, the projection  $W_{opt}$  is chosen to maximize the determinant of the total scatter matrix of the projected samples. That is:

$$W_{opt} = \arg \max_n |W^T S_T W| = [w_1 w_2 \dots w_m] \quad (3)$$

Where  $\{w_i | i = 1, 2, \dots, m\}$  is the set of  $N$ -dimensional eigenvector of  $S_T$  corresponding to the  $m$  largest eigenvalues.

By checking the weight coefficient matrix (known as loading), six wavebands were selected among all the spectral bands of the sample data combined with normalized principal component images to detect the defects of apples (Xing and Baerdemaeker, 2005).

An experiment using a hyperspectral imaging system for bruise detection on ‘Golden Delicious’ apples using the wavelength region between 400 and 1000nm with classification algorithms based on PCA and PLS-DA respectively were conducted by Xing’s group. Classification algorithms based on PCA and PLS-DA results were developed, and their performance with respect to the classification accuracy and feasibility to implement on-line sorting were compared (Xing, Bravo, Jancsok, et al 2005).

Separating stem-end/calyx region from true bruises is a big problem in apple defects inspection research. Based on PCA of hyperspectral images, multiple effective wavebands were selected to distinguish the stem-end/calyx region from the cheek surfaces by analyzing the contour features of the first principal component score

images (Xing, Jancsok, Baerdemaeker, 2007). A near-infrared hyperspectral imaging system with spectral region of 900-1700nm was developed; the system consisted of an imaging spectrograph attached to an InGaAs camera with line-light fiber bundles as an illumination source. Principal component analysis, band ratio, and band difference were applied in the image processing to segregate bruised cucumbers from normal cucumbers (Ariana, Lu, Guyer, 2006). Bruised tissue had consistently lower reflectance than normal tissue and the former increased over time. Best detection accuracies from the PCA were achieved when a bandwidth of 8.8 nm and the spectral region of 950–1350 nm were selected.

The results developed by (Roggo, Edmond, Chalus, et al. 2007) showed that the band ratio method and band difference method had similar performance, but better than the PCA.

#### **2.5.2.2 Other Classification Methods**

Hyperspectral image contains large amount of image, some useful information should be extracted for analysis. Therefore, a general strategy contains feature extraction and pattern classification. In addition to the most useful classification approaches such as Band Ratio, Principal Component Analysis, some other approaches such as: Partial Least Squares, General Discriminant Analysis, Artificial Neural Network and Spectral Angel Mapper are employed in the literature.

In hyperspectral imaging processing, the primary goal for researchers is to select the useful information among the large amount of data volume and evaluate the features for the hyperspectral data. By evaluating these features, they select the most differentiable ones compared with the rest. Furthermore, an integrated PCA-FDA method was developed (Cheng, 2002). ANN (ElMasry, Wang, Vigneault, et al, 2007; Hahn, Lopez, Hernandez et al 2004; Jayas, Paliwal, Visen, 2000; Ariana, Lu, Guyer, 2006) is a unique pattern recognition method used in hyperspectral image processing because of its nonlinear property. SAM (Park, Lawrence, Windham, et al. 2005, Park, Windham, Lawrence, et al 2007) is another supervised classification method by using  $n$  dimensional angle for matching pixels to the reference spectra. This method helps to determine the angle between individual spectra and was useful for the fecal detection and ingesta contaminations. Gaussian Mixture Model (GMM) is another classical method which uses the multivariate normal (Gaussian) probability density model. It was also reported that Gaussian-kernel based support vector machine (SVM) performance better than PCA and FDA in classifying the walnuts shell and pulp (Jiang, Zhu, Jing, et al. 2007a; Jiang, Zhu, Rao, et al. 2007b) .

## **Chapter 3 Apple Defect Detection using Hyperspectral Analysis**

### **3.1 Experimental System and Sample Materials**

#### **3.1.1 Hyperspectral Imaging System and Image Data Acquisition**

Hyperspectral and multispectral imaging for food safety have been extensively studied at ARS. In the laboratory, a rapid line-scan imaging system has been developed for hyperspectral Vis/NIR reflectance a. The system is integrated with a commercial apple sorting machine (FMC Corp, Philadelphia, PA, USA) and tested by inspecting apples for defects and fecal contamination.

The hyperspectral line-scan imaging system utilizes an electron-multiplying charge-coupled-device (EMCCD) imaging device (Andor iXon Du 860). An imaging spectrograph (ImSpector V10E, Spectral Imaging Ltd., Oulu, Finland) and a C-mount lens (XENOPLAN 1.4/23mm Compact, Schneider, USA) are attached to the EMCCD. The instantaneous field of view (IFOV) is limited to a thin line by the spectrograph aperture slit (50  $\mu\text{m}$ ). Through the slit, light from the scanned IFOV line is dispersed by a prism-grating-prism device and projected onto the EMCCD. Therefore, for each line-scan, a two-dimensional (spatial and spectral) image is created with the spatial dimension along the horizontal axis and the spectral dimension along the vertical axis of the EMCCD.

The line-scan imaging system used two 150-w quartz-tungsten lamps for reflectance imaging. The reflectance at wavelengths shorter than 450 nm is not used because of its poor signal-to-noise ratio; the very low irradiance in that portion of the spectrum is an attribute of the quartz halogen light sources. (Kim, Chen, Cho, et al. 2007)

### **3.1.2 Sample Materials**

The apples used in this study were 169 Red delicious apples, randomly selected from the Rice fruit Company (Gardners, PA). Among them, 89 of them are normal good apples, and the rest 80 apples have visible defects. These apples are “tree-run” apples—no wax or other coating was applied to the apple post-harvest. They are stored in a 4 °C cold room in the Environmental Microbiological and Food safety laboratory. Beltsville Area research Center, Agricultural Research Service, United States Department of Agriculture.

We use the aforementioned hyperspectral line-scan imaging system to inspect each of the 169 apples. After the line-scan image data is obtained, we randomly split the data into two groups: the calibration group and the testing group. The calibration group randomly selected 10 normal apples and 15 defect apples, the remaining 79 normal apples and 65 defect apples were in the testing group as shown in Table 1.

Table 1. Classification groups of the 169 sample apples.

Condition	Total Samples	Samples for Calibration	Samples for Testing
Normal	89	10	79
Defect	80	15	65

The hyperspectral image data in the calibration group is used to determine the optimal wavebands to generate the multispectral imaging algorithm based on the selected optimal wavebands. For the 10 normal samples, we randomly select 2817 pixels from each data. For the 15 apples with defects, we randomly select 3901 pixels from the normal skin area and 794 pixels from the defect area skin for evaluation (Table 2).

## 3.2 Hyperspectral Data Analysis Methods for Apple Defect Detection

### 3.2.1 Reflectance Spectrum Analysis

From the preliminary trial-and-error study, wavebands containing a little information were discarded. Only 64 channels were used in this research study, ranging from 535nm to 831nm, with an increment of approximate 4.7nm.

Table 2. Sample pixels for calibration of different classification groups

Condition	Selected Calibration Area	Number of pixels
Normal Apple	Random area	2817
Defect Apple	Normal surface	3901
	Defect surface	794



Our combined apple defect detection method is motivated by the reflectance spectrum of the sample apples. Figure 3 shows the relative intensity of the normal skin and the defect areas for wavelength between 535nm and 831nm. The three solid curves, from top to bottom, depicts the average spectral reflectance responses over three different categories of sample areas - good apple skin, normal surface of defect apples, and defect areas, respectively. The variance of the data is shown by the six dashed curves, which are the average plus or minus one standard deviation for each of the three solid curves.

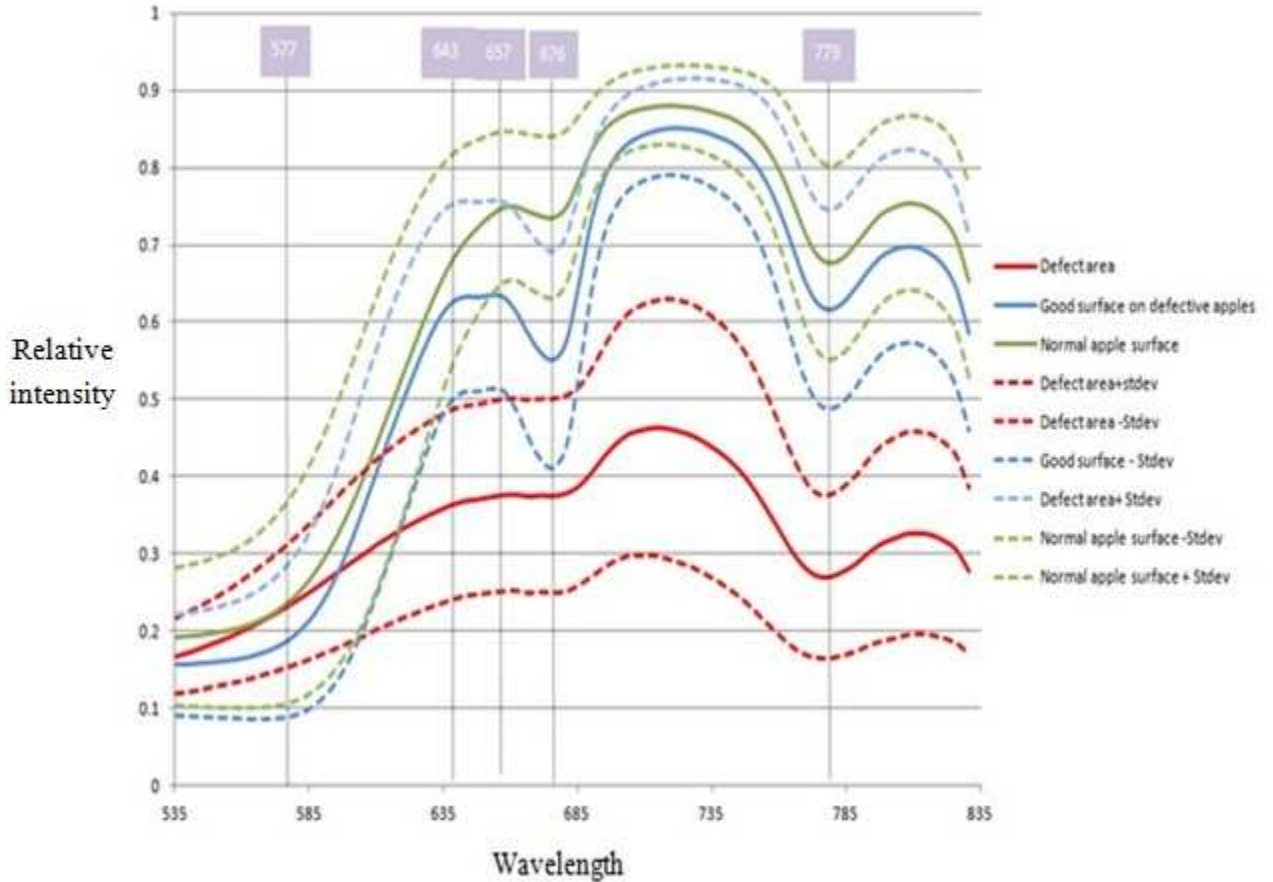


Figure 3 Reflectance spectra of sample image pixels from different surface of apples.

First, we observe that in general, the intensity level of good apple is the highest, followed by the normal surface of defect apples and the defect area has the lowest

intensity. This suggests that if the intensity level ( $I_x$ ) at certain wavelength  $X$  can be used to detect whether the apple has defect. From Figure 3, when the wavelength  $x$  is about 685nm or higher, the waveband of the average plus one standard deviation for defect area becomes so low that it does not even overlap with the curve of average minus one standard deviation for the normal surface of defect apples. The curves in Figure 3 reveals that the waveband at  $x=779\text{nm}$  may give the best result, where the intensity level of defect area has the largest difference from other surfaces. Using  $I_{779}$  for classification can successfully and effectively identify defect spots. However, it is unable to distinguish good apples and the normal surface of defect apples, which means that it will have a relatively high false negative ratio (that is fail to detect the defect apples when there is no response from the defect areas due to, for example, a bad position of the apple). Furthermore, our study also shows that this method has limitations in identifying the calyx, stem-end, and the apple boundary; the detail will be more discussed in chapter 4.

Second, it is observed that the pace the intensity level changes as wavelength increases are also different. For example, between 577nm and 643nm, the intensity level of both the normal surface area on defect apples and the surface of good apples increase much more sharply than that of the defective areas; between 657nm and 676nm, the surface reflectance of both good apples and the normal area of defect apples decrease while that of the defect area almost remains as a constant. These suggest us that we can measure how the wavebands change over different wavelength and use this to detect defects. For a wavelength interval  $[x, y]$ , we define the band

ratio  $R_{x/y} = I_x/I_y$ , where  $I_x$  and  $I_y$  are the intensity levels at wavelength  $x$  and  $y$ , respectively. The above observation from Figure 3 implies that the band ratio  $R_{577/643}$  for good apples and normal surface of defect apples will be noticeable smaller than that for the defect area; and the band ratio  $R_{657/676}$  for good apples and normal surface of defect apples will be larger than that for the defect areas. Our empirical study also shows that the band ratio  $R_{577/643}$  often misclassifies the apple calyx and stem-end as defect; it sometimes even reports good normal surface as defect; and interestingly, the other band ratio  $R_{657/676}$  is very effective in identifying calyx and stem-end in addition to its defect detection ability.

The flow chart of this combined apple defects detection method is shown in Figure 4. The apples were first inspected by relative intensity method, if the apple is detected as good apple, then it is the good apple; if the apple is detected as defect apple, then it will be further inspected by band ratio  $R_1$ , if the apple is detected as defect one using band ratio  $R_1$ , then it will be sent using band ratio  $R_2$  for further inspection, otherwise, the apple will be recognized as the good one.

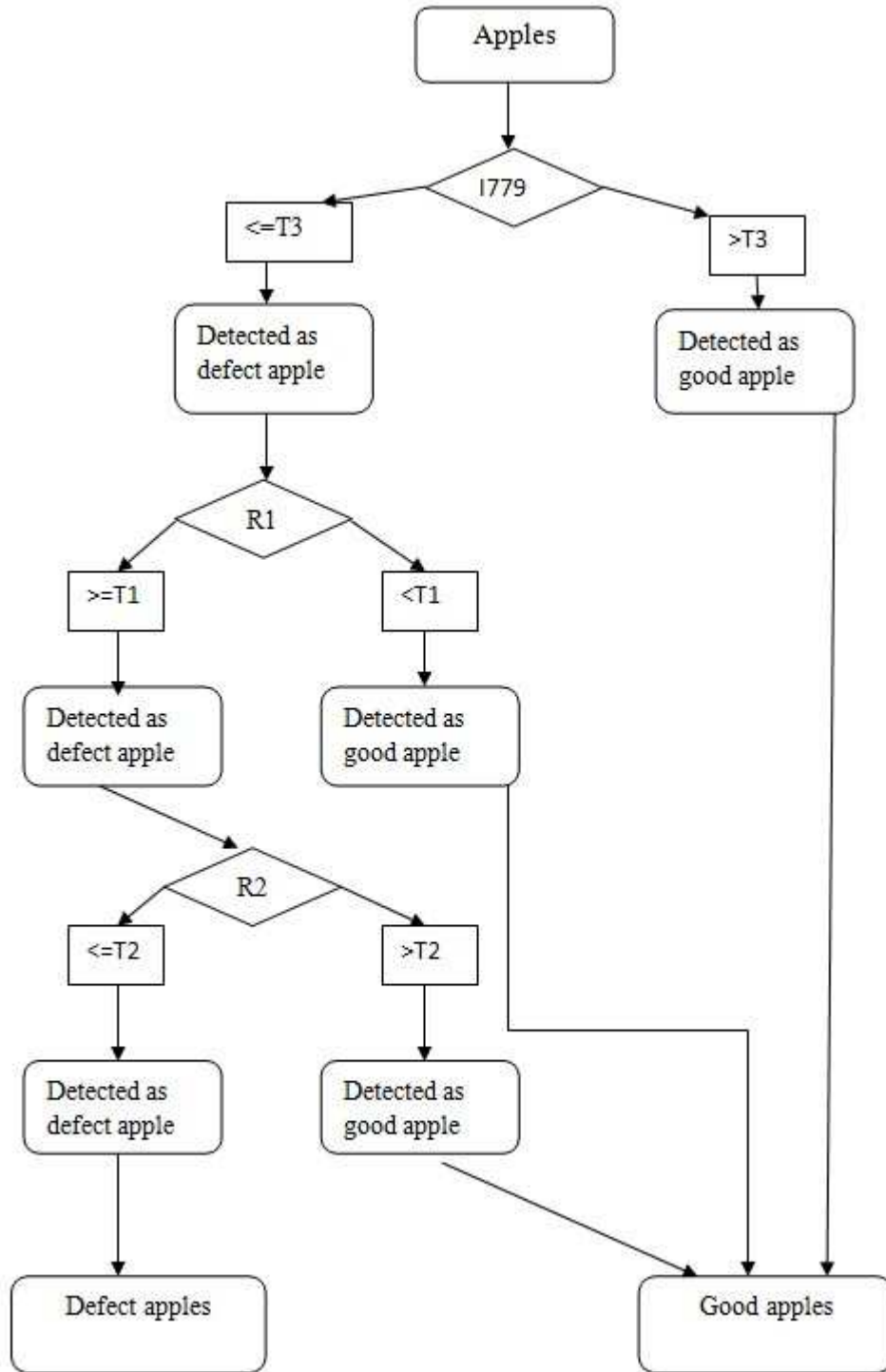


Figure 4 The combined method scheme of  $R_1$  &  $R_2$  & I method

To test the proposed combined classification method has a better classification rate, I tested three other combined methods, they were band ratio  $R_{577/643}$  and  $R_{657/676}$  using logic AND operation ( $R_1 \&\& R_2$ ) and logic OR ( $R_1 \parallel R_2$ ) operation; band ratio with logic OR operation then combined with relative intensity method using logic OR operation ( $I \parallel R_1 \&\& R_2$ ).

The four combined methods describe as follows:

$$B_1 = R_1 \&\& R_2$$

$$B_2 = R_1 \parallel R_2$$

$$B_3 = I \&\& (R_1 \parallel R_2)$$

$$B_4 = I \&\& R_1 \&\& R_2 \quad (4)$$

Where  $B_1, B_2, B_3, B_4$  are binary image results of four combined detection methods,  $R_1$  is the binary image ratio of  $R_{577/643}$  ( $R_{577/643} = I_{577}/I_{643}$ ) with threshold  $T_1$ ,  $R_2$  is the binary image of  $R_{657/676}$  ( $R_{657/676} = I_{657}/I_{676}$ ) with threshold  $T_2$ , symbol ‘&&’ is the logic AND operation, symbol ‘||’ is the logic OR operation. ‘I’ is the relative intensity of the apple image at 779nm.

To get a binary image result, different thresholds are applied to band ratio and relative intensity method. In this research, we define  $T_1$  as the threshold value for  $R_{577/643}$ ,  $T_2$  as the threshold value for  $R_{657/676}$  and  $T_3$  is the threshold value for  $I_{799}$ , respectively. To successful recognize a normal apple, all pixels in a normal apple should have  $R_{577/643}$  value less than  $T_1$ ,  $R_{657/676}$  and  $I_{779}$  value large than  $T_2$  and  $T_3$  respectively. For successful detection of defect area, at least one pixel in defect apples should have

$R_{577/643}$  value higher than  $T_1$ ,  $R_{657/676}$  and  $I_{779}$  value should lower than  $T_2$  and  $T_3$  respectively.

### 3.2.2 Band Selection

Representative reflectance spectra from 535nm to 831nm extracted from the hyperspectral images for normal apple surface on the both good apples and defective apples are shown in Figure 3. Reflectance of Red Delicious apples exhibits relatively high red reflectance at around 640nm, and also at around 676nm is the characteristic absorption of chlorophyll II  $\alpha$  in the red region (620nm-750nm). The reflectance spectra of defects, calyx and stem regions are visually dark in appearance Figure 5.

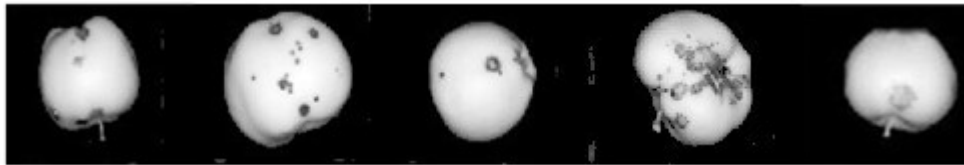


Figure 5 The typical defect apple image at 779nm wavelength

Differences between the normal apple surface response and the defect areas are clearly shown from the spectrum reflectance in the green region (495nm-570nm). In the red region, the defect area does not show the chlorophyll  $\alpha$  absorption phenomenon. The reflectance spectra between around 750nm and 830nm show the same characteristic information between normal apple skin and defect area.

Reflectance information at 577nm and 643nm are selected for the band ratio because the intensity values between these wavelengths of normal surface have much difference than that value of the defect area; the wavelengths at 657nm and 676nm

are also selected for band ratio because the intensity value of the normal surface increased much higher than that of the defect area; the intensity value of normal surface and defect area at 779nm is selected due to their big intensity value difference.

### 3.2.3 Threshold Selection

As stated in the spectrum analysis section, the threshold values were selected for band ratio method and the relative intensity method to get a binary result image. In this research, we compared three different thresholding methods. Here, we defined  $T_\alpha$  as the first threshold method, which uses the average defect sample pixel intensity or ratio value (plus or minus standard deviation  $\sigma_d$ ) as its value;  $T_\beta$  used the average normal pixel intensity value or ratio value (plus or minus standard deviation  $\sigma_n$ ) as its threshold value;  $T_\varepsilon$  is the average value of both  $T_\alpha$  and  $T_\beta$ .

The three threshold method can be expressed as in equation (5):

$$\begin{aligned}
 T_\alpha &= (I_d \text{ or } R_d) \pm \sigma_d, \\
 T_\beta &= (I_n \text{ or } R_n) \pm \sigma_n, \\
 T_\varepsilon &= (T_\alpha + T_\beta) / 2
 \end{aligned} \tag{5}$$

Where  $I_d$  and  $I_n$  are the intensity value of the defect area and normal apple surface on defect apples, respectively;  $R_d$  and  $R_n$  are the band ratio value of the defect area and normal apple surface on defect apples, respectively;  $\sigma_d$  and  $\sigma_n$  are the standard deviation value of each  $I_d$  and  $I_n$ , or  $R_d$  and  $R_n$ .

The average intensity value of normal apple surface, good surface on defect apples, and defect area accompanied with their standard deviation bar at wavelength 779nm are shown in Figure 6; and the band ratio value of normal apple surface, good surface on defect apples, and defect area at wavelength 779nm accompanied with its standard deviation bar are shown in Figure 7.

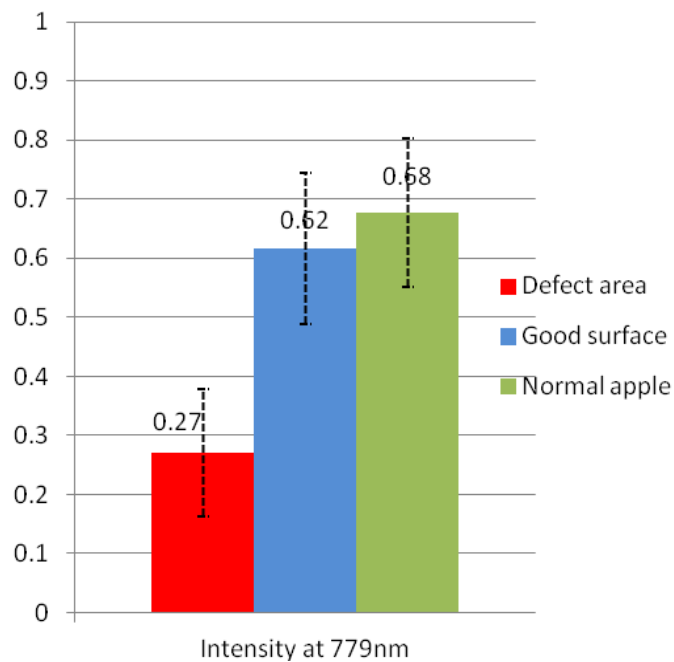
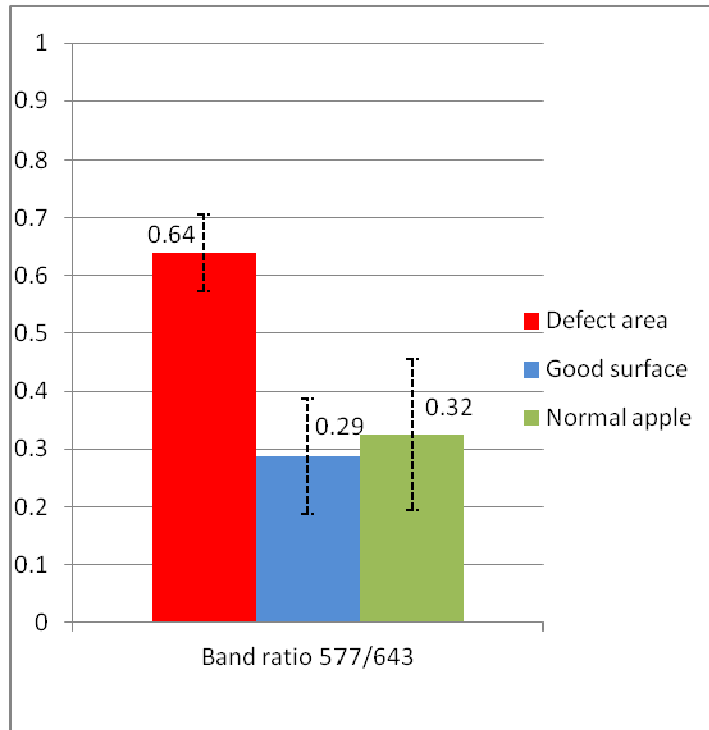
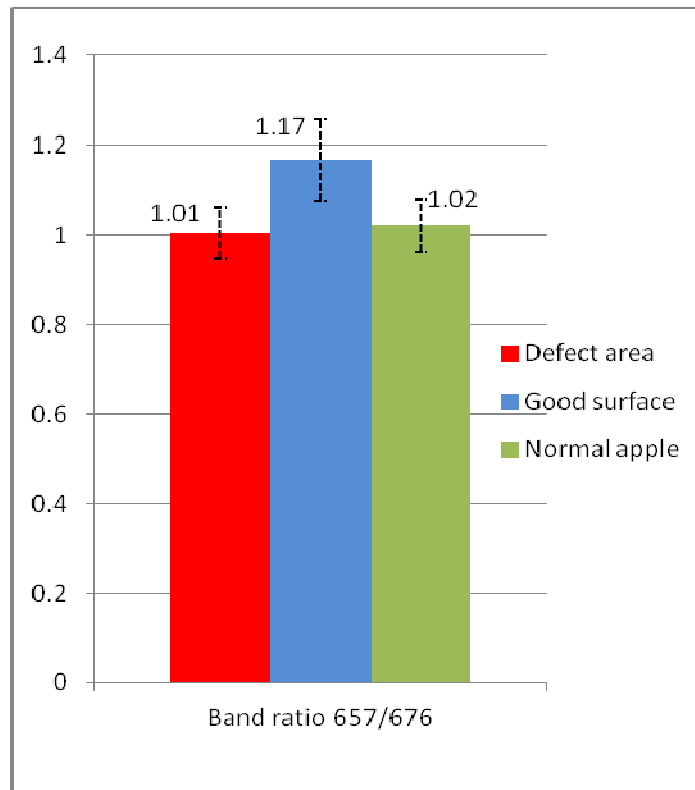


Figure 6 Average intensity value of sample pixels at 779nm wavelength of defect area, good surface on defect apples and normal apple surface.





(a)



(b)

Figure 7 (a) (b) Average of band ratio value of the sample apple image pixel

Prior to calculate the ratio values of the two band images, we usually perform a mask operation to eliminate the background information of the image. Table 3 shows the sample detection result using relative intensity method and band ratio method of  $T_\alpha$ ,  $T_\beta$  and  $T_\epsilon$ . As stated in previous methodology section, the normal surfaces on defect apples were considered very useful for threshold selection. For example, the result in table 3 showed that among three thresholding methods for the intensity method,  $T_\epsilon$  has a better classification result than  $T_\alpha$  and  $T_\beta$  methods. If we selected a good apple surface instead of the normal surface on a defect apple for the thresholding method, it will misclassify some normal surface area as defect.

Table 3. The defects detection result using three different thresholds on selected training samples

	$T_\alpha$	$T_\beta$	$T_\epsilon$
Intensity	82%	94%	100%
Ratio $R_{577/643}$	98%	95%	98%
Ratio $R_{657/676}$	68%	94%	79%

$T_\epsilon$  thresholding method was selected for the band ratio method between 577nm and 643nm from the result on Table 3. In Figure 7(a), the average ratio value of the defect area ( $R_{d1}$ ) is 0.64 with standard deviation value ( $\sigma-R_{d1}$ ) is 0.07; the average ratio value of normal apple surface ( $R_{gl}$ ) is 0.32 with standard deviation ( $\sigma-R_{gl}$ ) 0.13, and the average ratio value of the good surface on the defect apples ( $R_{n1}$ ) is 0.29 with its

standard deviation value ( $\sigma-R_{n1}$ ) 0.10. Therefore, the threshold value for  $T_1$  is defined in equation (6):

$$T_1 = ((R_{d1} - \sigma-R_{d1}) + (R_{n1} + \sigma-R_{n1})) / 2 \quad (6)$$

If the sample pixel ratio value is higher than  $T_1$ , it is classified as the defect; if it is lower than  $T_1$ , the pixel will be classified as the good area.

$T_\alpha$  threshold method was selected for band ratio between 657nm and 676nm according to the results in table 3. Figure 7(b) showed that average ratio value of the defect area ( $R_{d2}$ ) is 1.01 with standard deviation ( $\sigma-R_{d2}$ ) value 0.05; the average ratio of normal apple surface ( $R_{g2}$ ) is 1.02 with standard deviation ( $\sigma-R_{g2}$ ) is 0.06, and the average ratio value of the good surface on the defect apples ( $R_{n2}$ ) is 1.17 with its standard deviation value ( $\sigma-R_{d2}$ ) 0.09. Therefore, the selected threshold value for  $T_2$  defines in equation (7):

$$T_2 = ((R_{d2} + \sigma-R_{d2}) \quad (7)$$

The average value of the defect area ( $I_d$ ) is 0.27 with standard deviation ( $\sigma-I_d$ ) is 0.11; the average intensity of normal apple surface ( $I_g$ ) is 0.68 with standard deviation

$(\sigma-I_g)$  is 0.12, and the average intensity of the good surface on the defect apples ( $I_n$ ) is 0.62 with its standard deviation value ( $\sigma-I_n$ ) is 0.13.  $T_g$  threshold method is selected for the intensity method; therefore, the threshold value  $T_3$  defines as:

$$T_3 = ((I_d + \sigma-I_d) + (I_n - \sigma-I_n)) / 2 \quad (8)$$

Where,  $I_d$  and  $I_n$  is the average intensity value of defect area and normal surface, respectively;  $\sigma-I_d$  and  $\sigma-I_n$  is the standard deviation for average intensity value of defect area and normal surface. If the sample pixel intensity value is less than  $T_3$ , it will be classified as the defect; if it is higher than threshold value  $T_3$ , the pixel will be classified as the good apple surface.

The equation (6) (7), (8) help to find the threshold value for band ratios  $T_1$  and  $T_2$  is 0.51 and 1.01, the threshold value for intensity method at 799 nm ( $T_3$ ) is 0.47.

## Chapter 4 Results

### 4.1 Relative Intensity Method

The band difference detection result on normal apples by using the threshold value ( $T_3 = 0.47$ ) is shown in Figure 8. The band difference method checks every pixel of the image data, if the pixel value is less than the threshold, it could be regarded as a defect. The limitation for using the relative intensity method are the classification of the apple edge and also the calyx and stem-end classification, the fourth apple from left clearly showed that the calyx is misclassified as defect. Besides, the edges of apples were misclassified as the defect area, which can be shown in Figure 8. The intensity method used for inspection on 89 normal apples, which gave the 82% (73/89) correctly detection rate on good apple detection.

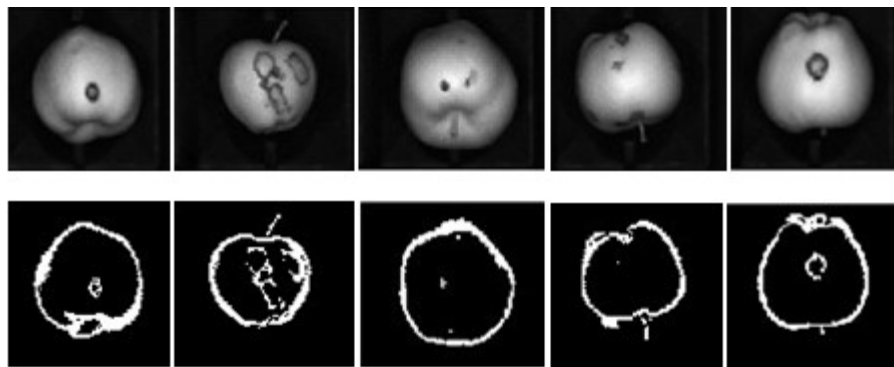


Figure 8 The binary image using relative intensity method at 779nm waveband

## 4.2 Band Ratio Result

The results in Figure 9 show that the bright spot is classified as the defect area. The apples on first two rows are the apple images at 577nm and 643nm waveband respectively, and the third row was the classifications result using band ratio  $R_{577/643}$ . The defect on the first apple (from left to right) is correctly detected from the binary image; however, the stem ends of the last four apples are misclassified as the defect spots. Therefore, the classification of the stem and calyx from the real defects is still the challenge in apple defects detection. Among total sample group and evaluation group apples, 74 of the 80 (92.5%) defective apples are correctly detected, while 48 apples with stem-ends and calyx are misclassified as defects; only 15 of the 89 (17%) normal apples are correctly detected.

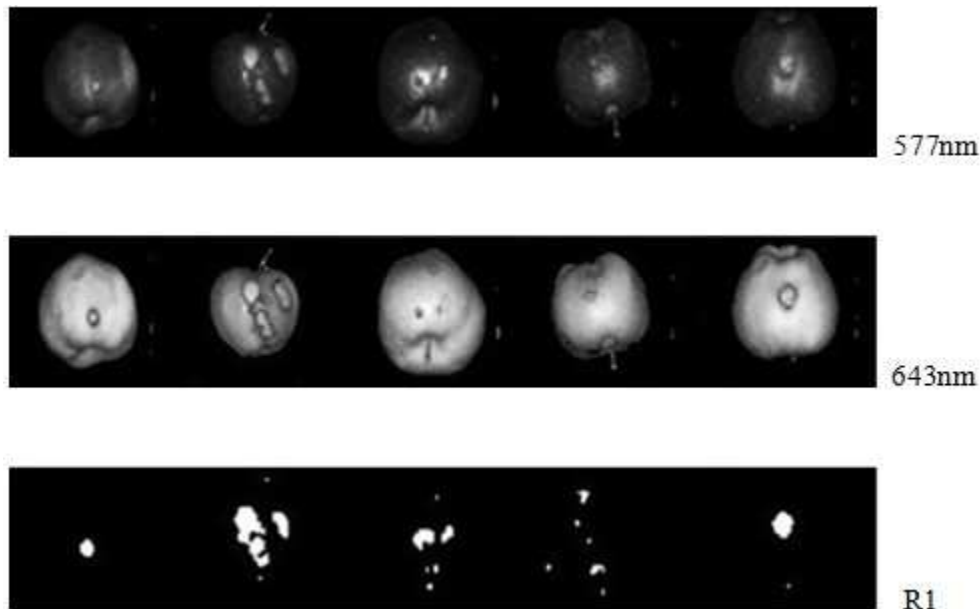


Figure 9 Image result of band ratio between 577 and 643nm. The first row is apple image at 577nm; the second row is apple image at 643nm; the binary images of ratio (577nm/643nm) results is shown in the third row.

The results in Figure 10 show that the bright spots are classified as the defect area. The upper two row images are the apple images at 657nm and 676nm waveband, respectively. The third row is the classification result using band ratio  $R_{657/676}$ . The defects on the first four apples were correctly detected from the binary image, even most of the stem-ends parts are correctly detected; however, the last apple misclassified apple edge as defect spot. The detection results using band ratio  $R_{657/676}$  show that among total sample group and evaluation group apples, 76 of the 80 (95%) defective apples are correctly detected, and most of the apples' stem-ends are also classified correctly, however normal surface on 13 defective apples (16%) are misclassified as defect spots, which is also called false-positive rate.

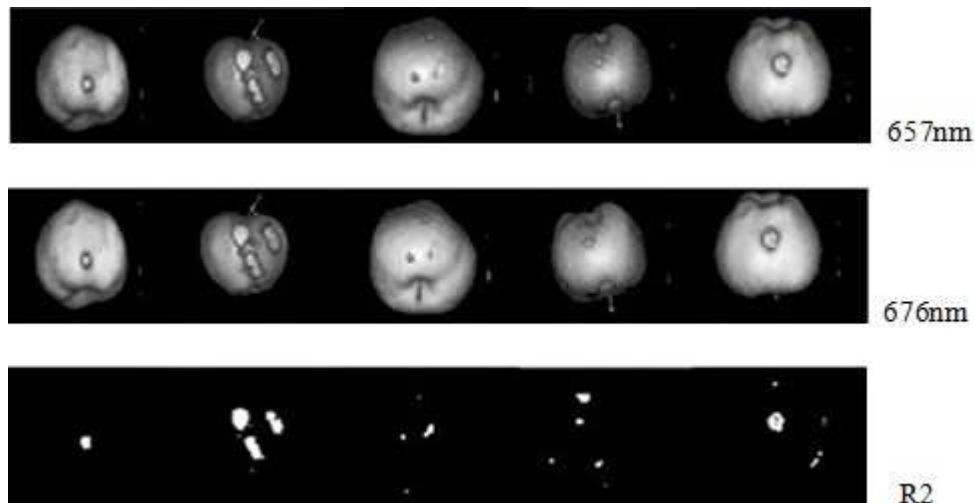


Figure 10 Image result of band ratio between 676 and 709nm. The first row is apple image at 657nm; the second row is apple image at 676nm; binary images of ratio (657nm/676nm) results is shown in the third row.

### 4.3 Band Ratios and Relative Intensity Combination

Although band ratio method and intensity method both can help to find some certain defects, they all have drawbacks in classifying the calyx, stem-ends and defects. To seek for a better defect detection rate, the combined method is applied into this research. Meanwhile, threshold value  $T_1$ ,  $T_2$  and  $T_3$  should be selected at the same time when they are applied to the combination of  $I_{779}$ ,  $R_{577/643}$  and  $R_{657/676}$ .

The detection results on 80 defect apples are shown in Table 4,  $R_1$  and  $R_2$  are the binary image result of band ratio  $R_{577/643}$  and  $R_{657/676}$  with threshold  $T_1$  and  $T_2$  respectively;  $I$  is the binary result of relative intensity method at 779nm with threshold  $T_3$ . Table 4 indicates that using  $R_1 \parallel R_2$  combined method gave a 84% defect detection rate, however, some normal surface are misclassified as defective area; by using ' $R_1 \&\&R_2$ ' method, a 82.5% defect detection rate is achieved.

Table 4. The defects detection result of several methods on defective apples

	R1	R2	I	$R_1 \parallel R_2$	$R_1 \&\&R_2$	$(R_1 \parallel R_2) \&\&I$	$R_1 \&\&R_2 \&\&I$	
Defect apples	Defect area	92.5% (74/80)	95% (76/80)	87.5% (70/80)	95% (76/80)	92.5% (74/80)	92.5% (74/80)	95% (76/80)
	Normal surface	40% (32/80)	84% (67/80)	45% (36/80)	84% (67/80)	82.5% (66/80)	85% (68/80)	96.3% (77/80)
	Whole apple	40% (32/80)	84% (67/80)	45% (36/80)	84% (67/80)	82.5% (66/80)	85% (68/80)	95% (76/80)

Figure 11 plots the final detection results on total 169 Red delicious apples compared with ' $(R_1 \parallel R_2) \&\&I$ ' and ' $R_1 \&\&R_2 \&\&I$ ' combined methods. The first method gives a



85% defect detection rate of 80 defective apples, and 67% normal apple detection rate of 89 normal apples; and the second method gives a 95% defect detection rate of 80 defect apples, and a 100% normal apple detection rate of 89 normal apples. By comparing these combined methods, ‘ $(R_1 || R_2) \& I$ ’ method has much better detection rate than other combined methods.

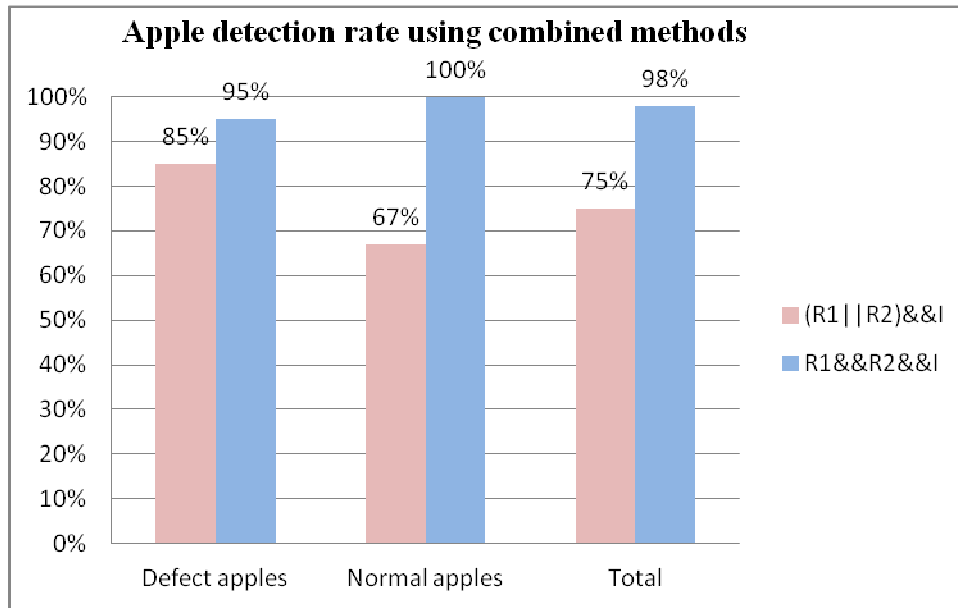


Figure 11 Detection results of the entire evaluation apples using ‘ $(R_1 || R_2) \& I$ ’ and ‘ $R_1 \& R_2 \& I$ ’ method.

In Figure 12, the binary image of combined method detection result is showed in the second row. The evaluation results show that all the 89 normal apples are classified as the good apples, which indicate a 100% correctly detection rate; 76 of 80 defect apples are also detected as defect apples, which gives a 95% correctly detection rate. The detection results using combined band ratio and intensity clearly show a better detection rate than using only band ratio method or intensity methods.

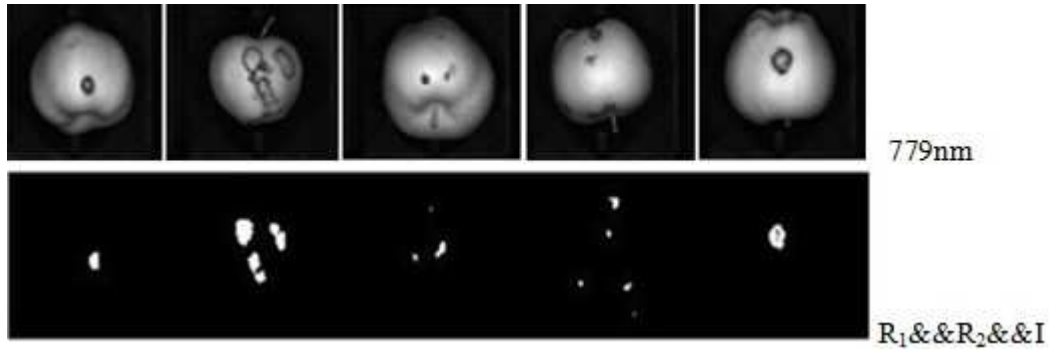


Figure 12 Image result of combined method using band ratio and intensity. The first row is apple image at 779nm; the second row is binary image result of proposed combined method.

Here, we can analyze the ‘ $R_1$  &  $R_2$  &  $I$ ’ combined classification into two detail stage. The first stage is to use single relative intensity method as the first process to classify the good apples and defect apples. If the apple is detected as good one using intensity method, the apple will be finally detected as good apple; if the apple is detected as the defective one by relative intensity method, then the apple will be further inspected using the band ratio methods.

#### 4.4 Discussion

From apple detection results in Table 4, band ratio method is good at detecting defect apples, among total 80 defect apples, band ratio  $R_1$  and  $R_2$  gives a correct defect detection rate 92.5% and 95%, respectively. However, when  $R_1$  and  $R_2$  are used on good apple classification, I find that both  $R_1$  and  $R_2$  misclassify many of good apples as defect ones, which causes high false positive rate. Results show that 59 normal apples are correctly detected and the other 30 normal apples are misclassified as defect ones by using combined method of  $R_1$  and  $R_2$ .

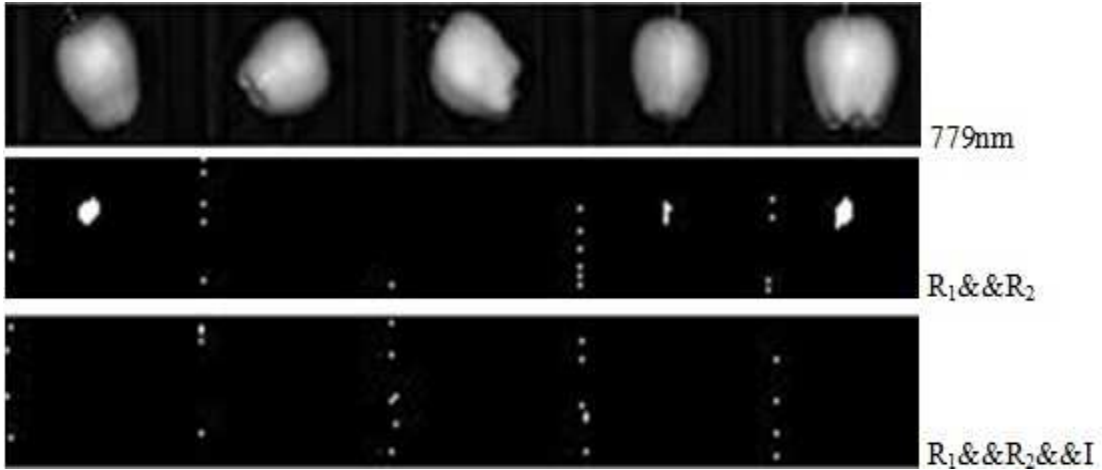


Figure 13 Image results of band ratio combined method and proposed band ratio and intensity combined method on good apples; the first row is the apple image at 779nm, the second row is binary image detection result of combined  $R_1$  and  $R_2$ ; the third row is binary image result of the proposed combined method.

From Figure 13, we can find that the first apple (from left), the fourth apple and the fifth apple are misclassified as defect apples using ' $R_1$  and  $R_2$ ' method; the detection results show that when the intensity method is added, it helps to correct this misclassification caused by ' $R_1$  and  $R_2$ ' method. Therefore, ' $R_1$  and  $R_2$  and  $I$ ' method gives a 100% classification rate on total 89 normal apples.

When using ' $R_1$  and  $R_2$  and  $I$ ' method, a 95% defect detection rate is achieved on total 80 inspected defect apples, 4 of them cannot be detected by this method. In Figure 14, the first row are the apple color images, the second row are apple images at 779nm, and the third row is the detection binary image result, we can see that four defect apples are not detected. The third apple (from left) has a minor defect spot, and the

proposed method may not has that much accuracy in detecting some small defects; however, the other three apples has large defects, but cannot be detected by the proposed method. We suspect the reason is these defects have similar surface with normal surface, thus there is no big difference between the normal surface reflectance response and defect area reflectance response along the spectrum.

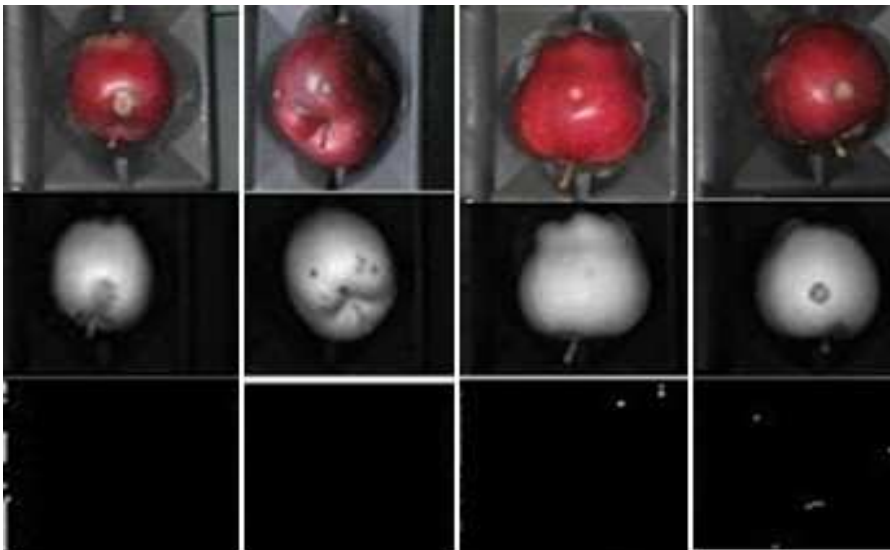


Figure 14 Four apple image results that misclassified as good one using 'R<sub>1</sub>&&R<sub>2</sub>&&I' method..

## Chapter 5 Conclusion

The study in this research selected five wavelengths at 577nm 643nm 657nm 676nm and 779nm for band ratio and relative intensity method. The requirement of only five wavebands and the simple computation of the algorithm present a potentially significant reduced time usage of the image acquisition and processing. It is an essential consideration for real-world application of non-destructive food safety inspection method.

An additional benefit of the developed combined method is that there is no need to differentiate apples from background. Since the PCA method and some other existing classification methods have the negative influenced by background information, throughout the entire imaging analysis after corrected image process, there is no pixel in the background gave false positive information by combined relative intensity and band ratio method. The studied detection results show that 95% of defect apples are correctly detected and 100% of the normal apples are detected correctly. However, the apple number in this study is count manually, it is necessary to detect and count apples automatically at the same time; this will be considered in the future research.

In this research, I implemented the relative intensity method first; however, this method has limitations, although it can find some of the defect area, it misclassified the apple edges and even calyx and stem-end as the defect spots. Another method I tested in this research is band ratio method, which is good to classify apple edge and

some calyx and stem-end. However, the band ratio method will miss some of the defect area. By thinking the advantages and disadvantages of these two classification methods, I found that the combination of the relative intensity method and band ratio method will give a better correct classification result.

For this study of focusing on the apple defect detection, it only use the reflectance image of the line-scan imaging system, we do have using the violet LED lights presented another band ratio combined method on fecal contamination detection on apples using the fluorescence imaging technique utilized UV excitation. Considering of an EMCCD camera, spectrograph, C-mount lens and even line lights shown in Figure 2 , this system could be made portable for installment on the a food processing line. For the real-world application, using such a system offers great potential as a value-added dynamic inspection system due to its capability for multi-tasking to meet a variety of inspection objectives. A multi-tasking inspection system that can meet current industry sorting needs with the added benefit of safety inspection without requiring significant modification of existing infrastructure or incurring significant costs may lead the apple industry to consider adopting voluntary measures to further enhance safe production and processing of fruits.

## **Chapter 6 Future Study**

The light source of hyperspectral imaging system used in this research is a pair of 150w quartz-tungsten lamps, the apple we have collected before did not under best uniform illumination, when we used band ratio method on good apples, this situation showed clearly. The apple surface has sharp bright area which below the fiber halogen light. I am thinking that whether we could align the illumination light and make it more uniform in the future study.

The apple hyperspectral data in our research only captured most part of the apple surface that face to the camera. In the future study, we could make the apple sorting machine to rotate the apple. Imagine that we can rotate the apple surface for 360°, and the line scanning imaging system can capture the whole surface of an apple.

## Appendix A

### Results of Principal Component Analysis on Hyperspectral Image

Principal component analysis (PCA) was performed on the reflectance spectra of total 169 Red delicious apples, and hence several principal components could be achieved. In general the first four eigenvalues calculated in PCA process have preserved more than 90% energy of the whole dataset. As a result, there is a balance between the number of selected principal components and the performance and computation time.

#### 1.1 Selection of optimal wavebands

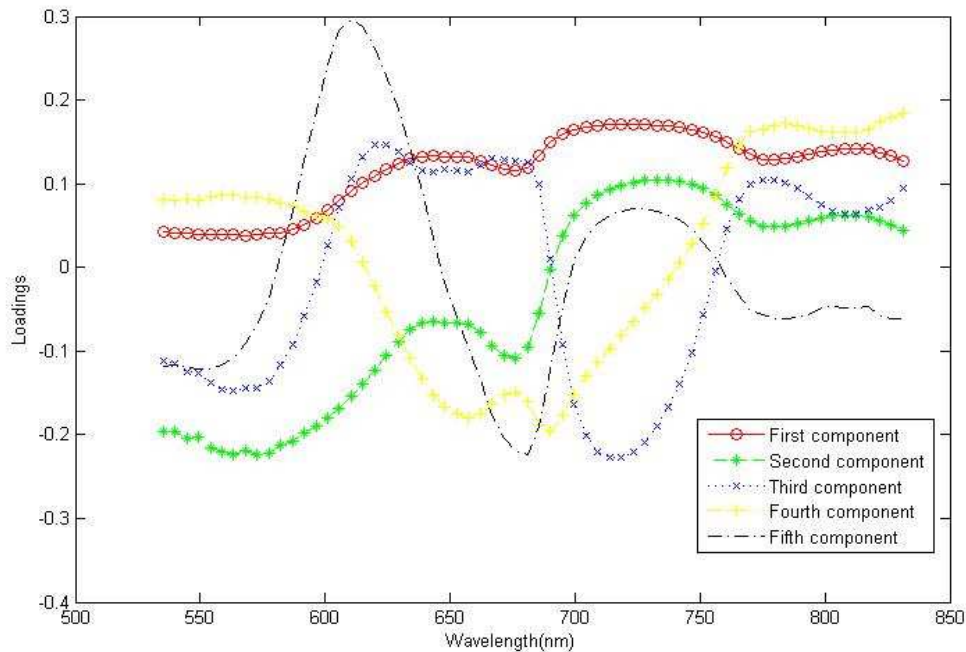


Figure 15 Loading plots of the average five selected 'Red Delicious' Apples

In Figure 15, it showed the average loadings of first four components from selected 5 Red delicious apples across the entire spectral region. Although the first principal component



accounts for most of the variance in the image, we will still analysis the second principal component image and the third and fourth principal component to because it can clearly show the difference between the defect area and normal area in single image. Thus, four wavebands are choose, are 558nm 676nm 779nm and 812nm. The following analysis was carries out on the selected four wavebands instead of whole spectral region. As we can see from Figure 16, Prin 1 scores image mainly accounts for the apple shape effect or illumination conditions; the Prin2 or Prin3 are good for discriminating defects and even bruise tissue.

A classification method to discriminate the defect from normal area was suggested by Xing (2004). Combined normalized image were used to identify the apple defects using the following equations:

$$I_{i,norm} = \frac{I_i - \min(I_i)}{\max(I_i) - \min(I_i)} \quad (8)$$

$$I_{ij} = \frac{I_{i,norm} - I_{j,norm}}{I_{i,norm} + I_{j,norm}} \quad (9)$$

Where,  $I_{i,norm}$  is the original Prin $i$  score image,  $\max(I_i)$  is the maximum value of the Prin  $i$  score image, and  $\min(I_i)$  is the minimum value of the Prin  $i$  score image;  $I_{i,norm}$  is the normalized Prin  $i$  score image;  $I_{ij}$  is the combined Prin  $i$  and Prin  $j$  normalized image. The following Figure 17 is the example of the combined normalized prin2 and prin3 images.

After using this method on total 169 Red delicious apples, among these apples, 80 of them are apples with visible defects and the rest 89 apples are normal one. The PCA

classification method result showed that 7 of the 80 defects are not detected, and 9 normal apples were mistaken as defected ones. However, we found that among the 9 mistaken defect apples, 4 of them are have light bruise which cannot clearly see by human eyes were detected by PCA classification method.

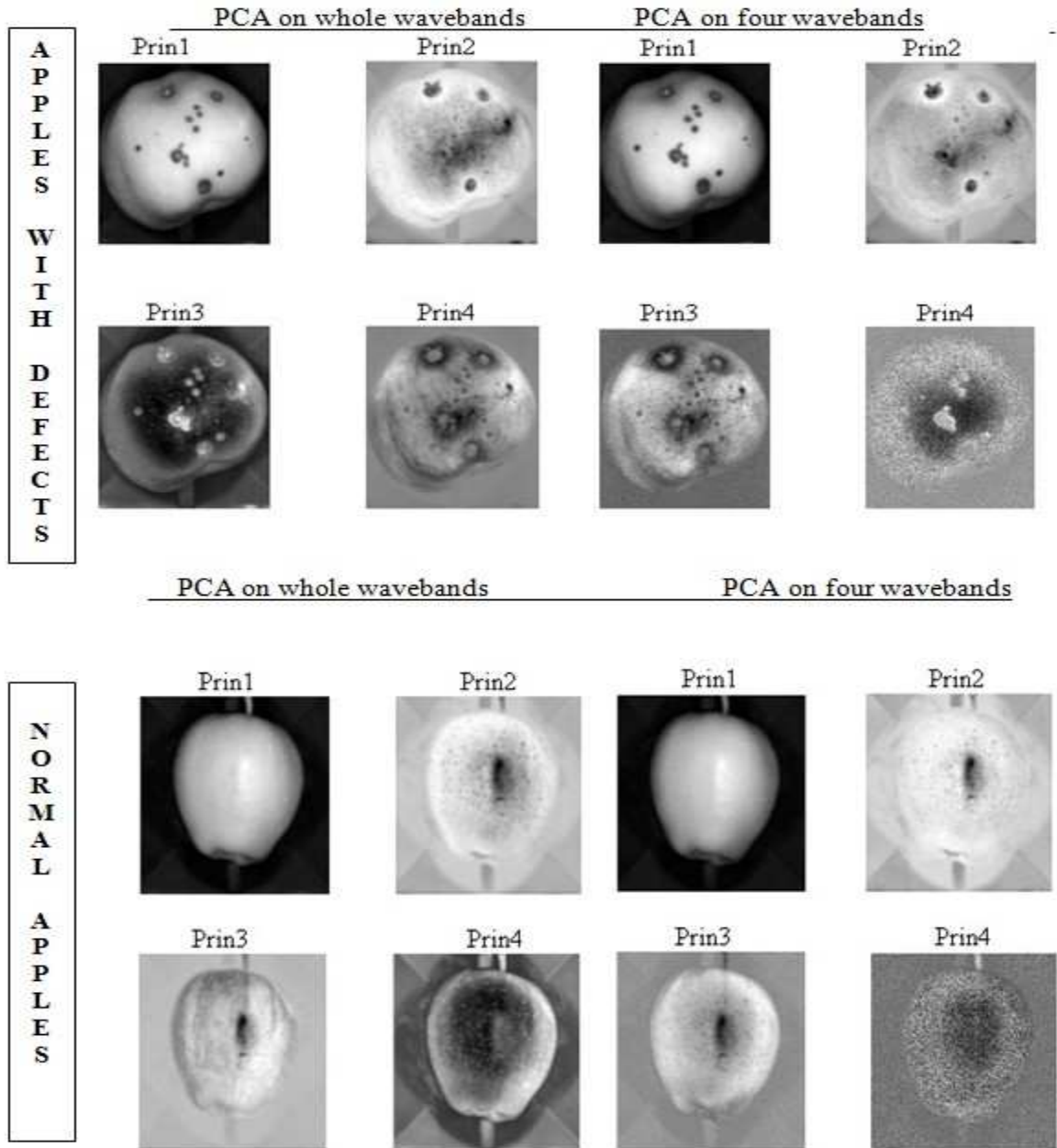


Figure 16 Principal component results by PCA methods

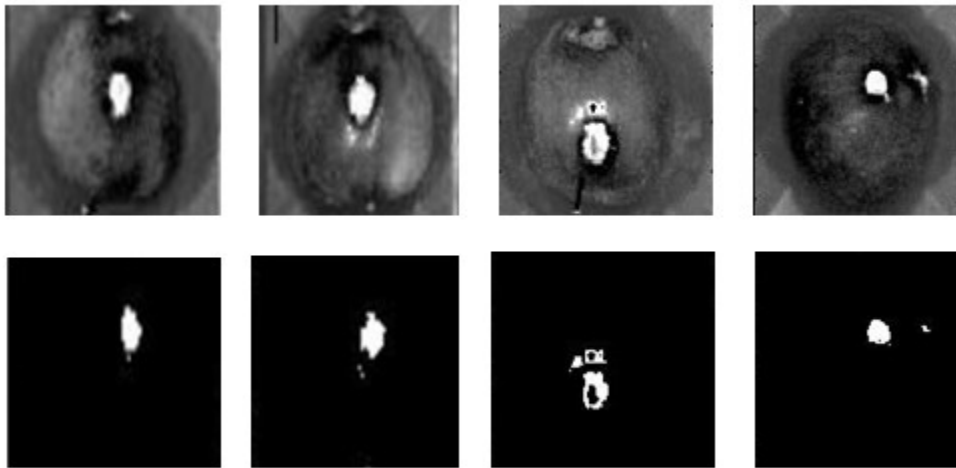


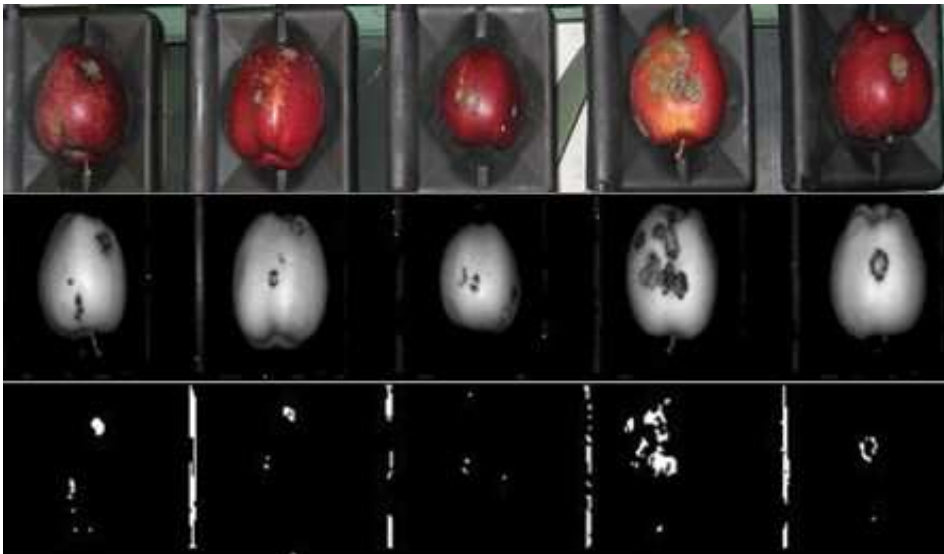
Figure 17 Combination of normalized images (upper), and after thresholding results (lower)

## Appendix B

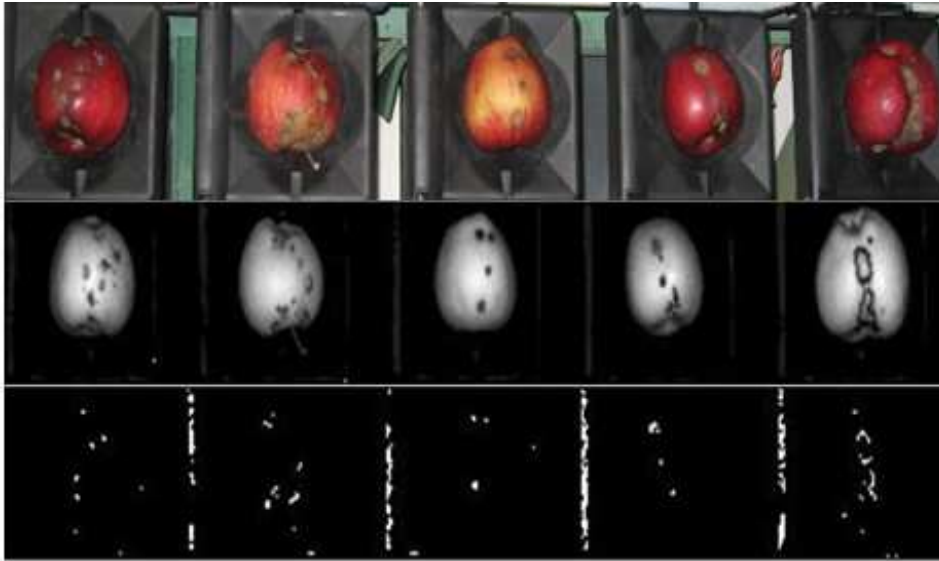
### Apple Image Results using Proposed Classification Method

The following 16 images show apple defects detection results, the first row is the color photo taken during experiment; the second row is the reflectance image at 779nm; the third row is the defect detection result using the proposed intensity and band ratio combined method in this thesis.

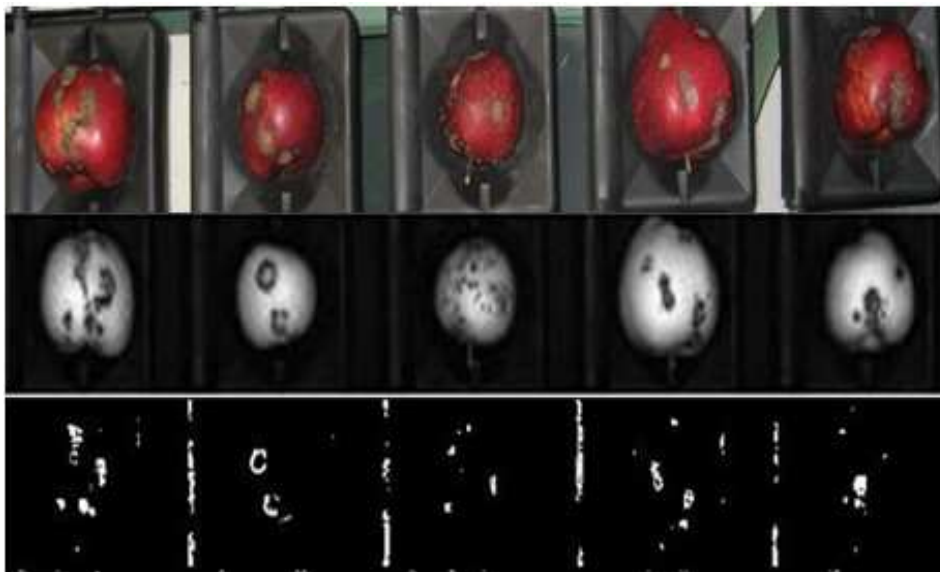
1) Defect apples 1 to 5 (Color photo, Reflectance at 779nm, Detection results)



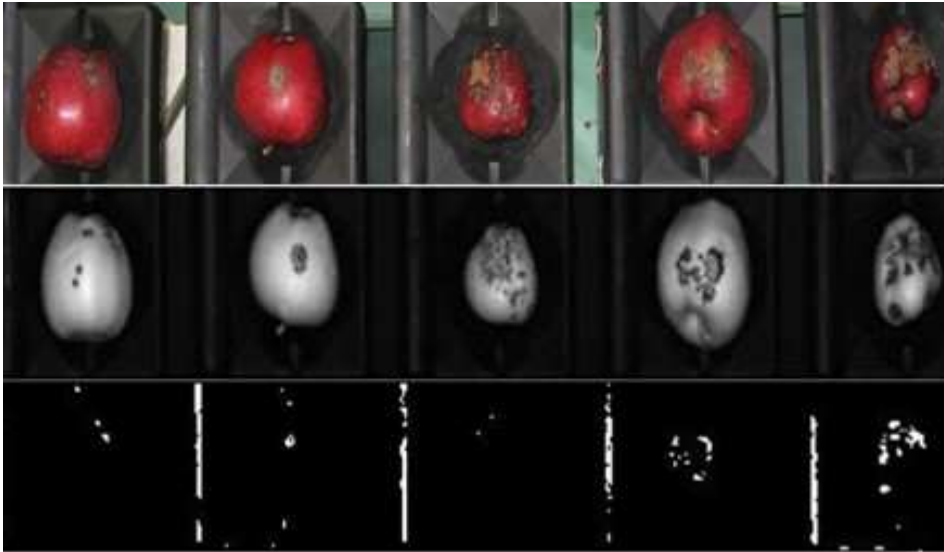
2) Defect apples 6 to 10 (Color photo, Reflectance at 779nm, Detection results)



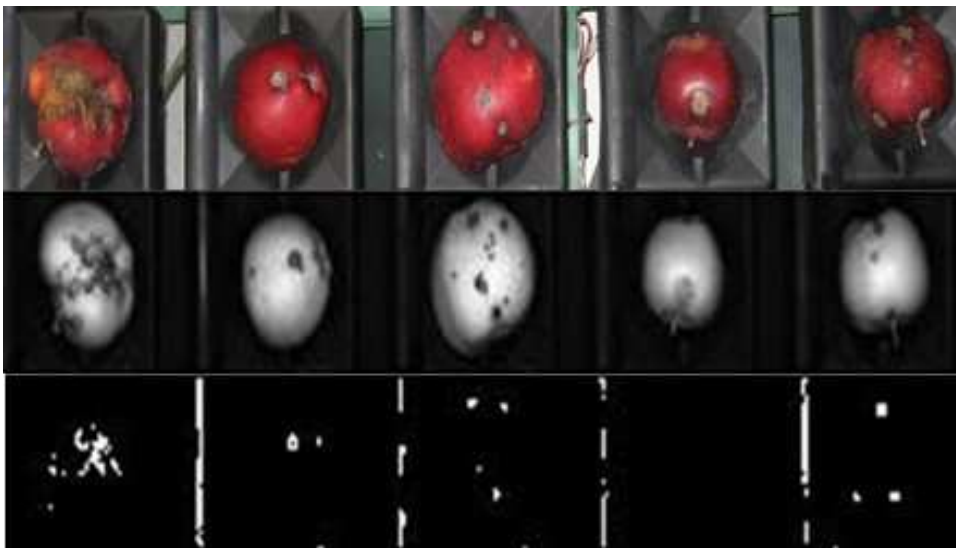
3) Defect apples 11 to 15 (Color photo, Reflectance at 779nm, Detection results)



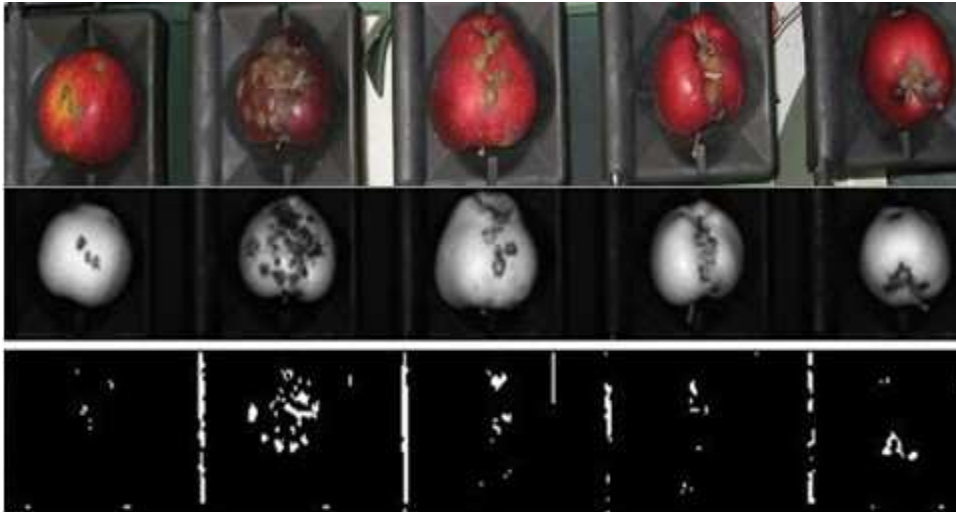
4) Defect apples 16 to 20 (Color photo, Reflectance at 779nm, Detection results)



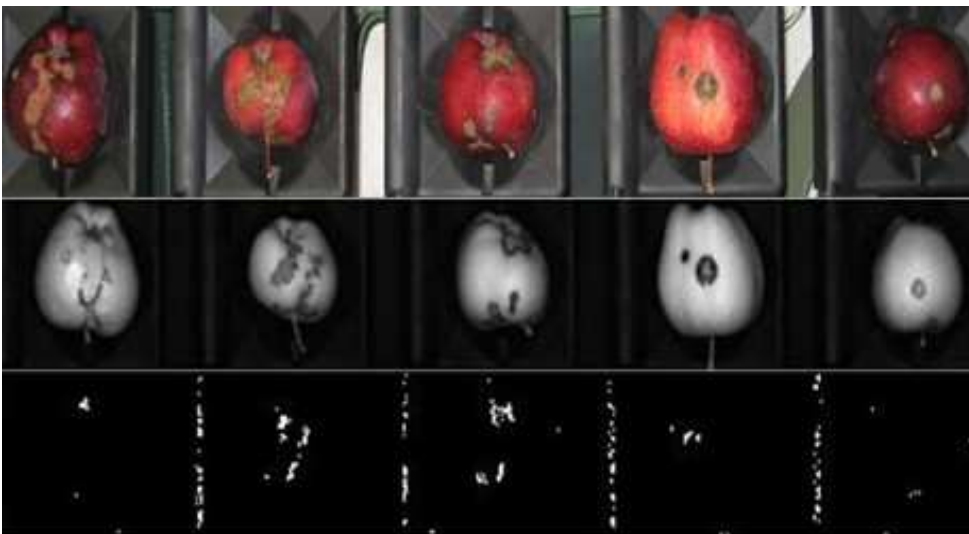
5) Defect apples 21 to 25 (Color photo, Reflectance at 779nm, Detection results)



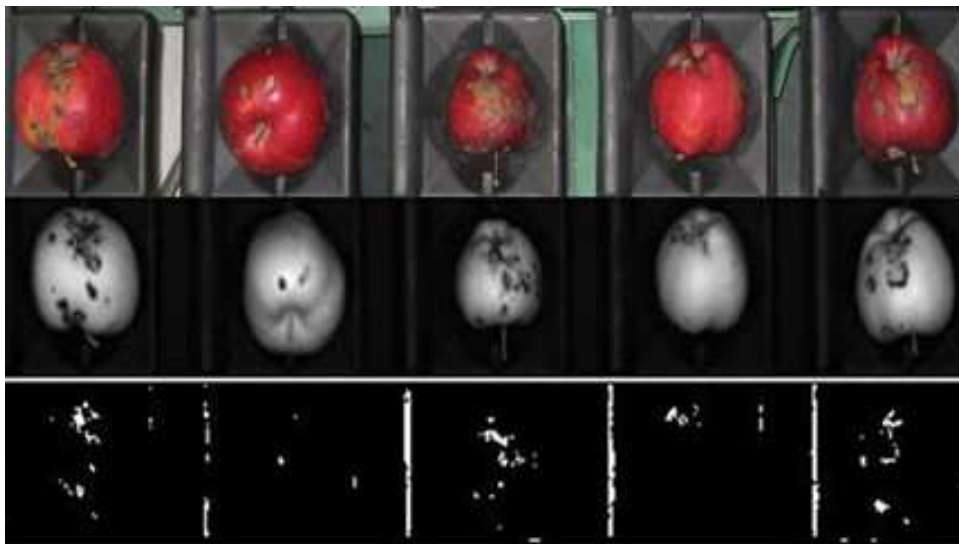
6) Defect apples 26 to30 (Color photo, Reflectance at 779nm, Detection results)



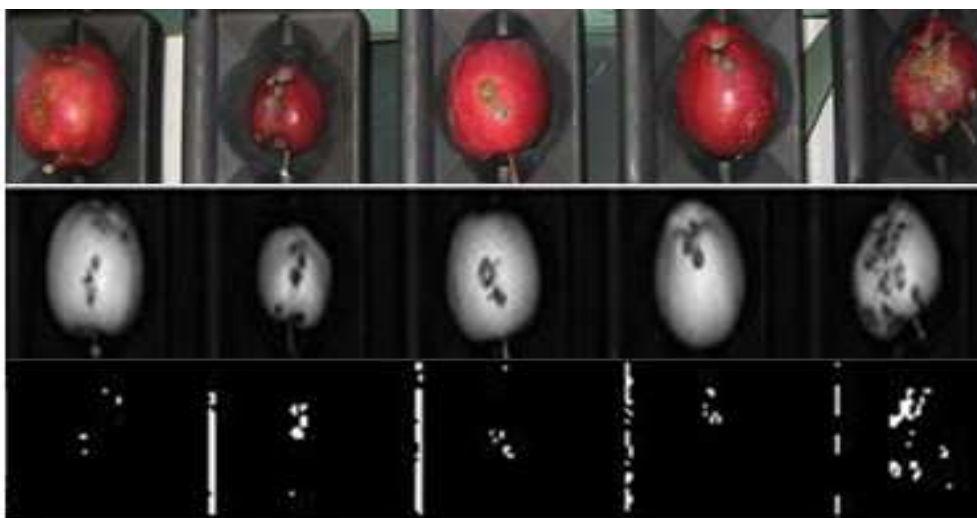
7) Defect apples 31 to35 (Color photo, Reflectance at 779nm, Detection results)



8) Defect apples 36 to 40 (Color photo, Reflectance at 779nm, Detection results)

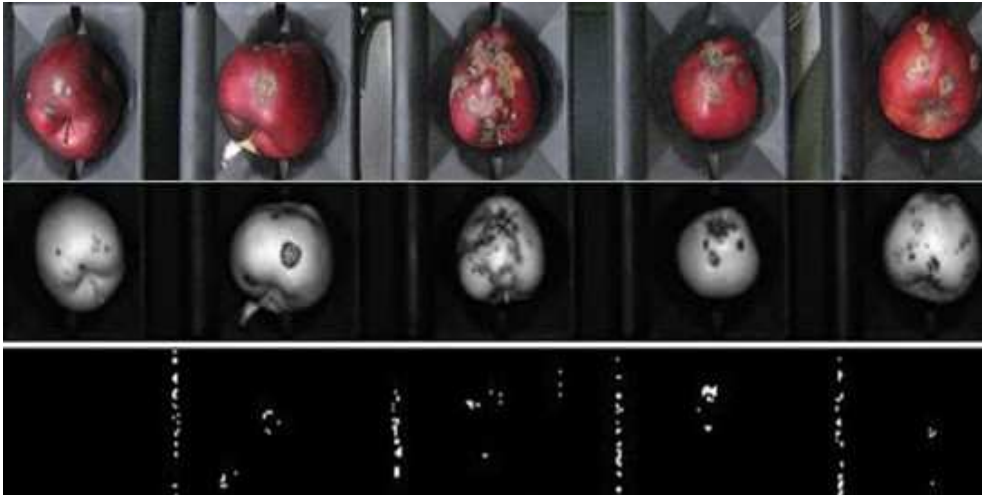


9) Defect apples 41 to 45 (Color photo, Reflectance at 779nm, Detection results)

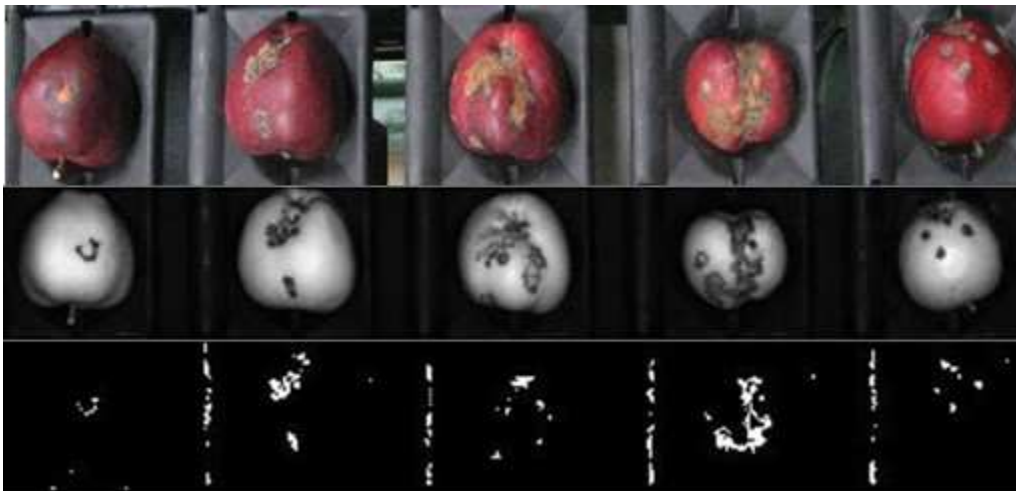




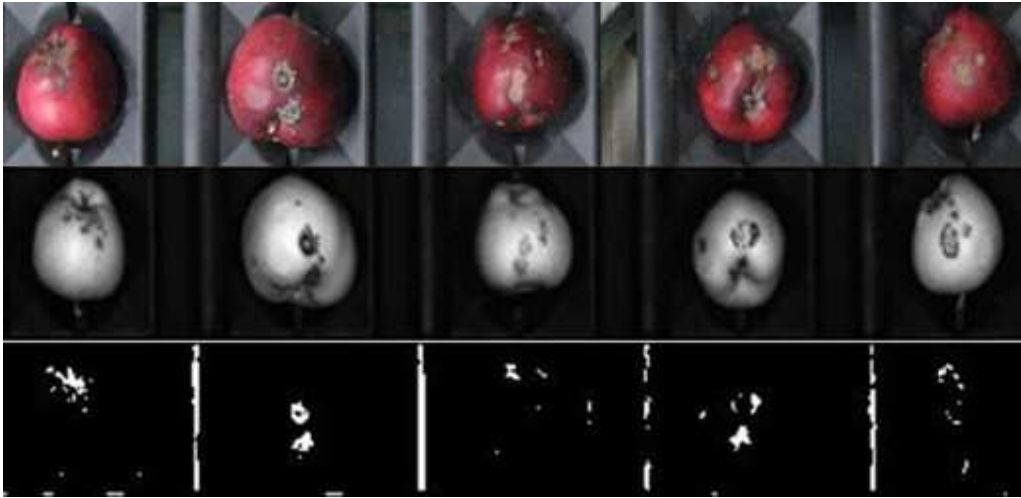
10) Defect apples 46 to 50 (Color photo, Reflectance at 779nm, Detection results)



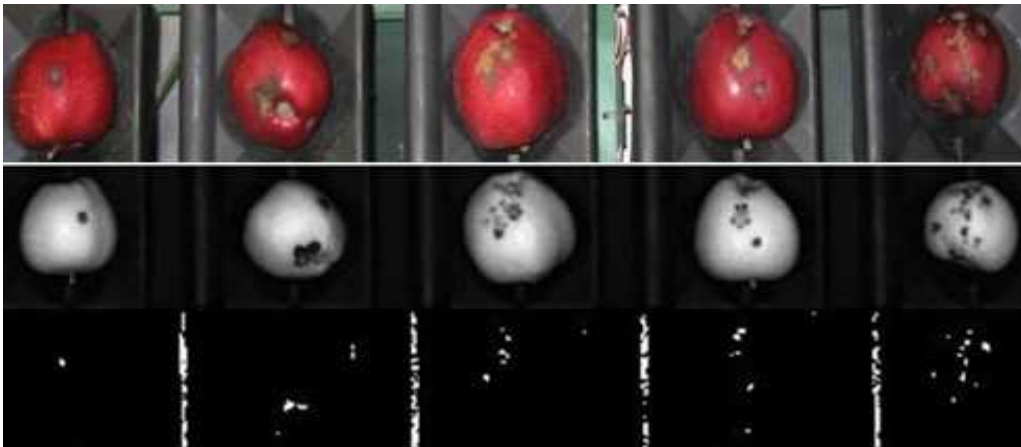
11) Defect apples 51 to 55 (Color photo, Reflectance at 779nm, Detection results)



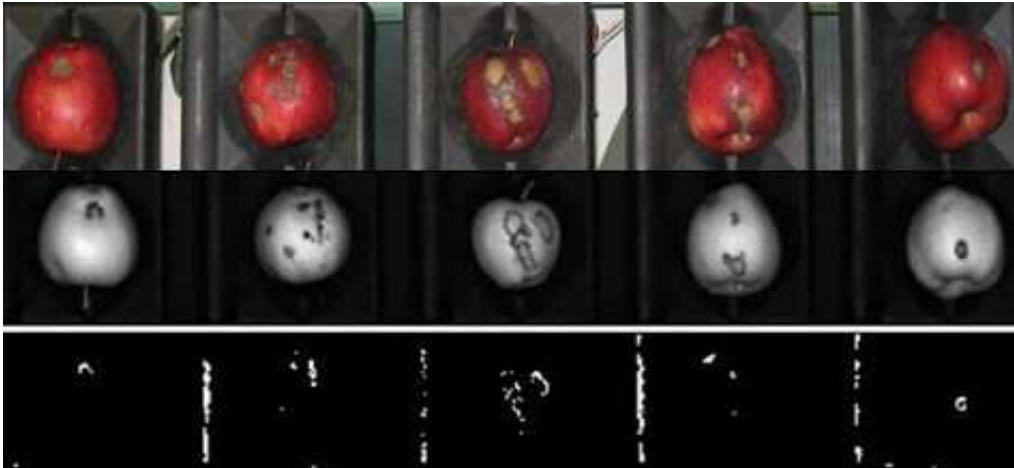
12) Defect apples 56 to 60 (Color photo, Reflectance at 779nm, Detection results)



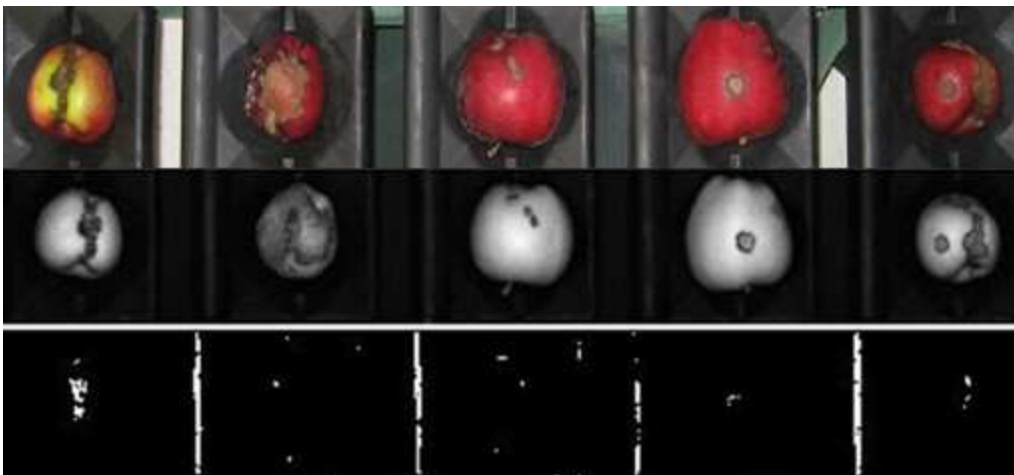
13) Defect apples 61 to 65 (Color photo, Reflectance at 779nm, Detection results)



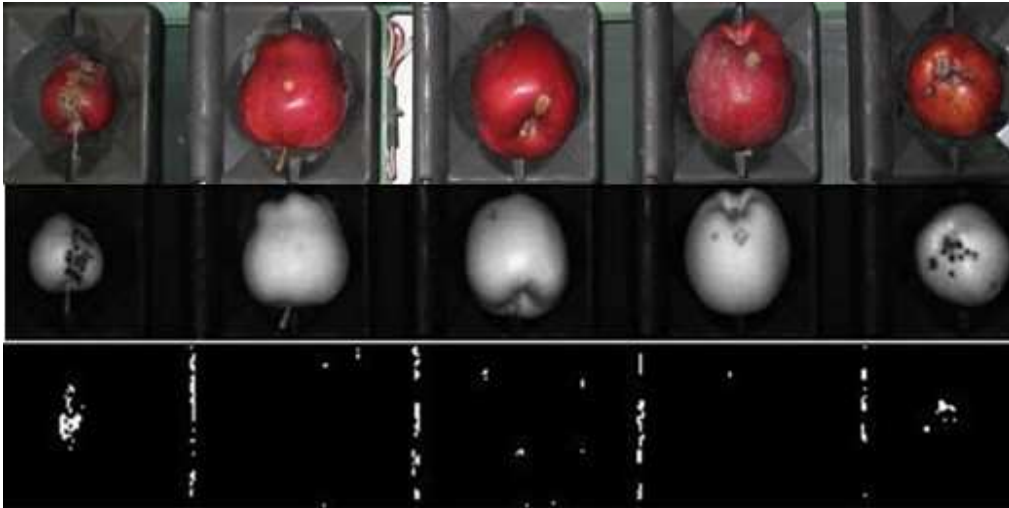
14) Defect apples 66to70(Color photo, Reflectance at 779nm, Detection results)



15) Defect apples 71to75(Color photo, Reflectance at 779nm, Detection results)



16) Defect apples 76 to 80(Color photo, Reflectance at 779nm, Detection results)



## Reference

- Ariana, D., Guyer, D. E. and Shrestha, B., “Integrating Multispectral Reflectance and Fluorescence Imaging for Defect Detection on Apples”, in *Computers and Electronics in Agriculture*, 2006, pp.148-161.
- Ariana, Diwan P., Lu, R.F. and Guyer, Daniel E., “Near-Infrared Hyperspectral Reflectance Imaging for Detection of Bruises on Pickling Cucumbers”, in *Computers and Electronics in Agriculture* 53(1), 2006, pp. 60-70.
- Campbell, J. B., “Introduction to Remote Sensing,” 3rd ed. Oxford, U.K.: Taylor and Francis, 2002.
- Cheng, X., Chen, Y.R., Tao, Y., Wang, C.Y., Kim, M.S. and Lefcourt, A.M., “A Novel Integrated PCA and FLD Method on Hyperspectral Image Feature Extraction for Cucumber Chilling Damage Inspection,” in *Transactions of the ASAE* 47 (4), 2002, pp.1313–1320.
- Cheng, Xuemei, “Hyperspectral Imaging and Pattern Recognition Technologies for Real Time Fruit Safety and Quality Inspection,” Doctor Dissertation, 2002.
- Knopf, David, “National Agricultural Statistic Service,” USDA, 2010.
- ElMasry, G., Wang, Ning, Elsayed, Adel and Ngad, Michael, “Hyperspectral Imaging for Nondestructive Determination of Some Quality Attributed for Strawberry”, in *Journal of Food Engineering*, 2007, pp.98–107.

- ElMasry, G., Wang, Ning, Vigneault, Clement, Qiao, Jun and ElSayed, Adel, “Early Detection of Apple Bruises on Different Background Colors using Hyperspectral Imaging,” In Swiss Society of Food Science and Technology, 2007, Doi:10.1016/J.Lwt.2007.02.022.
- Fukunaga, K., “Introduction to Statistical Pattern Recognition”. 2nd Ed, New York, N.Y.: Academic Press, 1990.
- Giraldo, V.A.M, “Detection of Fecal Contamination on Cantaloupes and Strawberries using Hyperspectral Fluorescence Imagery,” Master Thesis, University of Maryland, 2006.
- Hahn, F., Lopez, I. and Hernandez, C., “Spectral Detection and Neural Network Discrimination of Rhizopus Stolonifer Spores on Red Tomatoes,” in Biosystems Engineering 89 (1), 2004, pp.93-99.
- Jayas, D.S., Paliwal, J. and Visen, N.S., “Multi-layer Neural Networks for Image Analysis of Agricultural Products,” In Journal of Agriculture Engineering Research, 2000, pp.119-128.
- Jiang, L., Zhu, B., Jing, H., Chen, X., Rao, X. and Tao, Y., “Gaussian Mixture Model Based Walnut Shell and Meat Classification in Hyperspectral Fluorescence Imagery,” in Transactions of the ASABE, 50(1), 2007, pp.153-160.
- Jiang, L., Zhu, B., Rao, X., Berney, G. and Tao, Y. , “Discrimination of Black Walnut Shell and Pulp in Hyperspectral Fluorescence Imagery using Gaussian Kernel Function Approach,” in Journal of Food Engineering, 81(1), 2007, pp.108-117.

- Kavdir, I. and Guyer, D.E., "Apple Sorting using Artificial Neural Networks and Spectral Imaging," in Transactions of the ASAE 45(6), 2002, pp.1995-2005.
- Kim, M.S., Chen, Y. and Mehl, P., "Hyperspectral Reflectance and Fluorescence Imaging System for Food Quality and Safety," in Transactions of the ASAE 44(3), 2001, pp. 721-729.
- Kim, M.S., Chen, Y.R., Cho, B., Lefcourt, A.M, Chao, K., and Yang, C.C., "Hyperspectral Reflectance and Fluorescence Line-Scan Imaging for Online Quality and Safety Inspection of Apples," in Sensing and Instrumentation for Food Quality and Safety 1(3), 2007, pp.151-159.
- Kim, M.S., Lefcourt, A.M., Chao, K., Chen, Y.R., Kim, I. and Chan, D.E., "Multispectral Detection of Fecal Contamination on Apples Based on Hyperspectral Imagery: Part I Application of Visible and Near-Infrared Reflectance Imaging," in Transactions of the ASAE 45 (6), 2002, pp.2027–2037.
- Kim, M.S., Chen, Y. R. and Mehl, P.M., "Hyperspectral Reflectance and Fluorescence Imaging System for Food Quality and Safety," in Transactions of the ASABE 44(3), 2001, pp.721-729.
- Kim, Moon, Chen, Y-Ren, Cho, Byoung-Kwan, Chao, Kuanglin, Yang, Chun-Chieh, Lefcourt, Alan M. and Chan, Diane, "Hyperspectral Reflectance and Fluorescence Line-Scan Imaging for Online Defect and Fecal Contamination Inspection of Apples," in Sensing and Instrumentation for Food Quality and Safety, 2007, pp.151-159.

- Kim, M.S., Chao, K., Chane, D. and Lefcourt, A., “Line-Scan Imaging for High-Speed Food Safety Inspection,” SPIE Newsroom, 2009, DOI: 10.1117/2.1200903.1564.
- Liu, Y., Chen, Y. R., Kim, M. S., Chan, D. E., & Lefcourt, A. M., “Development of Simple Algorithms for the Detection of Fecal Contaminants on Apples from Visible/Near Infrared Hyperspectral Reflectance Imaging,” in *Journal of Food Engineering* 81(2), 2007, pp.412-418.
- Liu, Y., Chen, Y. R., Wang, C. Y., Chan, D. E. and Kim, M. S., “Development of Simple Algorithm for the Detection of Chilling Injury in Cucumbers from Visible/Near-Infrared Hyperspectral Imaging,” in *Applied Spectroscopy* 59(1), 2005, pp.78--85.
- Liu, Y., Chen, Y. R., Wang, C. Y., Chan, D. E. and Kim, M. S., “Development of Hyperspectral Imaging Techniques for the Detection of Chilling Injury in Cucumbers; Spectral and Image Analysis,” in *Applied Engineering in Agriculture* 22(1), 2006, pp.101-111.
- Lu, R. and Peng, Y., “Hyperspectral Scattering for Assessing Peach Fruit Firmness,” in *Biosystems Engineering* 93(2), 2006, pp.161–171.
- Noh, H. and Lu, R., “Hyperspectral Laser-Induced Fluorescence Imaging for Assessing Apple Fruit Quality,” in *Postharvest Biology and Technology* 43, 2007, pp.193-201.



- Park, B., Lawrence, K. C. , Windham, W. R. and Smith, D. P., “Detection of Fecal Contaminants in Visceral Cavity of Broiler Carcasses using Hyperspectral Imaging,” in *Applied Engineering in Agriculture* 21(4), 2005, pp. 627-635.
- Park, B., Windham, W.R., Lawrence, K.C., and Smith, D.P. , “Contaminant Classification of Poultry Hyperspectral Imagery using a Spectral Angle Mapper Algorithm,” in *Biosystems Engineering* 96 (3), 2007, pp.323–333
- Park, B., Windham, W.R., Lawrence, K.C., and Smith, D.P., “Hyperspectral image Image Classification for Fecal and Ingesta Identification by Spectral Angle Mapper”, in *ASAE/CSAE Meeting presentation*, 2007.
- Perry West, “High Speed, Real time Machine Vision,” *Automated Vision Systems, Inc.*
- Roggo, Y., Edmond, A., Chalus, P. and Ulmschneider, M., “Infrared Hyperspectral Imaging for Qualitative Analysis of Pharmaceutical Solid Forms,” in *Sensing and Instrumentation for Food Quality and Safety*, 2007, pp. 151-159.
- Tao, Yang, Buchanan, Robert, Song, Yoonseok, Luo, Yanguang, Chen, Yud-Ren and Kim, Moon, “Safety Inspection of Fruit and Vegetables using Optical Sensing and Imaging Techniques,” *Proposal for the USDA*, 2002.
- USDA ERS., “Economics, Statistics and Market Information System,” Table 3--U.S. Apple Production and Utilization, by State, 1980-2009.
- USDA, “United States Standard for Grades of Apples,” *USDA*, 2002.

Vargas, A.M., Kim, M.S., Tao, Y., Lefcourt, A.M., Chen, Y.R., Luo, Y., and Song, Y., “Detection of Fecal Contamination on Cantaloupes using Hyperspectral Fluorescence Imagery,” In *Journal of Food Science*. 2005, 70:8.

Wen, Z., and Tao, Y., “Dual-Camera NIR/MIR Imaging for Stem-end/Calyx Identification in Apple Defect Sorting,” in *Transactions of the ASAE* 43(2), 2000, pp.449-452.

Wikipedia

Xing, J. and, Baerdemaeker, J.D., “Bruise Detection on ‘Jonagold’ Apples using Hyperspectral Imaging,” in *Postharvest Biology and Technology* 37 (2), 2005, pp.152-162.

Xing, J., Bravo, C., Jancsok, P., Ramon, H., Baerdemaeker, J.D., “Detecting Bruises on ‘Golden Delicious’ Apples using Hyperspectral Imaging with Multiple Wavebands,” in *Biosystems Engineering*, 2005, pp.27-36.

Xing, J., Jancsok, P. and Baerdemaeker, J. D., “Stem-end/Calyx Identification on Apples using Contour Analysis in Multispectral Images,” in *Biosystems Engineering* 96 (2), 2007, pp.231–237.

Xing, J., Saeys, W. and Baerdemaeker, J. D., “Combination of Chemometric Tools and Image Processing for Bruise Detection on Apples,” in *Computers and Electronics in Agriculture* 56(1), 2007, pp.1-13.

