ABSTRACT

Title of Thesis:	Detection of Fecal Contamination on Cantaloupes and Strawberries Using Hyperspectral Fluorescence Imagery
	Angela M. Vargas, Master of Science, 2006
Thesis directed by:	Professor, Yang Tao Biological Resources Engineering Department

Fluorescence methods are widely used for investigation of biological materials, and in recent years have also been used to monitor food quality and safety. In this research, fluorescence imaging techniques for detecting fecal contamination on cantaloupes and strawberries were evaluated. Fluorescence images at emission peaks were examined for fecal classification. These images were subjected to further analysis utilizing band ratios and principal component analysis. Two-band ratio images and principal component images, compared to the single-band images, enhanced the contrast between the feces-contaminated spots and untreated sample surfaces. The images exhibited useful results for contamination detection, however, false positives resulting from natural color variation on strawberry surfaces present a problem throughout the methods. This study confirmed the capability of hyperspectral fluorescence imaging in detecting fecal matter on cantaloupes and strawberries and the potential for this method to be used for developing on-line applications.

DETECTION OF FECAL CONTAMINATION ON CANTALOUPES AND STRAWBERRIES USING HYPERSPECTRAL FLUORESCENCE IMAGERY

By

Angela M. Vargas

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
2006

Advisory Committee:
Professor Yang Tao, Chair
Dr. Moon Kim
Dr. Alan Lefcourt
Dr. Yud-Ren Chen
Professor Martin Lo

Foreword

The investigation for the work included in this Master's thesis was carried out at the Instrumentation and Sensing Laboratory (ISL), United States Department of Agriculture (USDA), Beltsville, Maryland, and at the Biological Resources Engineering (BRE) Department, University of Maryland (UMD), College Park, Maryland. The study is part of the project entitled "Safety Inspection of Fruits and Vegetables Using Optical Sensing and Imaging Techniques" granted in 2002 by the Joint Institute of Food Safety and Nutrition (JIFSAN). The supervisors were Professor Yang Tao (BRE, University of Maryland), and Senior Researcher Moon Kim (ISL, USDA). The advisors were Senior Researcher Alan Lefcourt (ISL, USDA), Research Leader Yud-Ren Chen (ISL, USDA), and Professor Martin Lo (Food Science and Nutrition Department, UMD). The thesis work includes a synopsis (Chapter 1), a set of publications (Chapter 2) of two journal articles, and a conclusion (Chapter 3).

Dedication

To my mother, Margarita, who has been a source of encouragement and support throughout my life. It is also dedicated to my uncle and aunt, Mario and Martha, who have been more than parents to me, and my siblings, Mauricio and Marcela, for being paragons of love and effort.

Acknowledgements

I would like to thank the Instrumentation and Sensing Laboratory (ISL), USDA, the Joint Institute of Food Safety and Nutrition (JIFSAN) and the Department of Biological Recourses Engineering (BRE) at the University of Maryland (UMD) for sponsoring this work. I greatly acknowledge Professor Yang Tao (BRE, UMD), who supported and supervised this research. I am very thankful to Dr. Moon Kim (ISL, USDA), for his scientific discussions, support and in educating me about the world of scientific research. Dr. Alan Lefcourt (ISL, USDA), who trained me about specific details on research and the Byzantine scientific English writing style. My deepest gratitude goes to the ISL research leader, Dr. Yud-Ren Chen, for incorporating me into the ISL staff. The internship at the ISL introduced me to spectroscopy sensing for food safety and enabled me to gain valuable knowledge for this thesis work. Special thanks to Dr. Martin Lo in the Department of Nutrition & Food Science at UMD for introducing me to the world of sensors for food safety and for supporting my academic growth on many occasions. I would also like to acknowledge other colleagues at the ISL, in particular Senior Researcher Dr. Stephen Delwiche, for discussing my data with me. I am also thankful to the research support staff Diane Chan, Frank Gwozdz, and John Kelly. I would like to express my appreciation to all members of the BRE Bio-Imaging and Machine Vision Laboratory at UMD.

Table of Contents

List of Figures

Chapter 1: Synopsis

The Instrumentation and Sensing Laboratory (ISL) of the United States Department of Agriculture (USDA), the Biological Resources Engineering Department (BRE) at the University of Maryland at College Park (UMCP), and the Food and Drug Administration (FDA) jointly worked on research using hyperspectral imaging to develop optical sensing techniques capable of instantly detecting fecal contamination and defects in fruits and vegetables. This endeavor was the basis for this master's research.

1. Importance of Imaging Systems to Ensure Safety and Quality of Fresh Produce

1.1. Contamination of Fresh Produce

Fruits and vegetables are popular among consumers. Consumers buy fruits and vegetables to maintain a healthier diet and because there is year-round availability of a large variety of domestic and imported produce. Reports indicate that over the past two decades there has been a significant increase in per capita consumption of raw fruits and vegetables in the United States. Fruit and vegetable consumption averaged 741 pounds per person annually during 1997-99, 25 percent above consumption levels during 1977-79 (USDA, 2001a). Simultaneously, outbreaks of foodborne illnesses in the U.S. have been frequently associated with the consumption of fresh produce (FDA, 1998). In an effort to address this problem, the FDA issued the

"Guide to Minimize Microbial Food Safety Hazards for Fresh Fruits and Vegetables" (FDA, 1998). The guide emphasized animal manure and human fecal matter as major harbors of pathogenic contamination. Additionally, *Salmonella, Escherichia coli 0157:H7* and *Shigella* were mentioned as major pathogens contributing to outbreaks of foodborne illness associated with fresh and minimally processed produce. Studies investigating factors contributing to these outbreaks have pointed out that a large number of variables influence contamination of fresh produce. Contamination points can be during production, harvest, processing, during transportation or retailing (USDA, 2001). In addition to faulty agricultural practices, fecal contamination can also be attributed to random natural events by wild animals. Examples of on-site sources of contamination from animal waste include animals pasturing near growing areas; manure storage adjacent to crop fields; leaking or overflowing manure lagoons; uncontrolled livestock access to surface waters, and high concentrations of wildlife (FDA, 1998). Additionally, produce can come into direct contact with human feces through poor sanitation practices of workers.

1.2. Cantaloupes and Strawberries

Muskmelons, commonly known in the U.S. as cantaloupes, are members of the gourd family *Curcurbitaceae*, which includes squash, pumpkins, cucumbers and watermelons. Muskmelons produce a sprawling vine that occupies a large ground area and thus exposes the fruit to the soil. Its fruit is characterized by sweet flesh within a netted rind (Porter, 1967). At maturity, the rind color changes from green to orange/yellow. Cantaloupes (*Cucumis melo* 'cantaloupensis') that are normally available are the Western Shipper. These cantaloupes have no sutures or ribs; they are

heavily netted and appear almost perfectly round. They are raised on the West Coast and are chosen more for their longer shelf life and perfect appearance than for their flavor.

The strawberry is a small plant of the *Rosaceae* (Rose) family. All varieties of the strawberry plant belong to the *Fragaria* genus (Sauer, 1993). Strawberries can be planted during spring or fall, when the temperatures are not excessively cool or warm. Initially, the fruit is greenish-white; but as it ripens, the fruit turns red. The calyx, a leafy star-shaped appendage, remains attached to the fruit even after harvest. The plant presents a growing pattern that, similar to cantaloupes, exposes the fruit to the soil. Its fruit consists of a fleshy receptacle whose surface is covered with achenes (commonly referred to as “seeds”) (Winton, 1902). For this research, strawberries belonging to the Earliglow variety were used. They are characterized by their wonderful flavor, glossy skin, firm flesh and medium size.

Throughout the centuries, cantaloupes and strawberries have been cultivated for their nutritional value. These produce have been found to be low in calories, fat and cholesterol. They are an excellent source of vitamins A and C, and other valuable minerals. They are also a good source of folate and soluble fiber (Wang, 2000, Wedge et al., 2001). Today, cantaloupes and strawberries continue to be food commodities valued for their natural health benefits and appealing physical appearance.

1.3. Health Issues Related to Imported Cantaloupes and Strawberries

International agreements have opened doors to year-round imports of tropical fruits from other countries where, unfortunately, strict standards for sanitation procedures while growing and handling are not always employed. Therefore, on numerous occasions, the USDA and the FDA have issued warnings and even sanctions on these imports. In October 2001, the FDA issued an important alert on cantaloupes from Mexico. The alert recommended that officials detain cantaloupes imported from Mexico (FDA, 2002). Investigations of *Salmonella* outbreaks in the U.S., responsible for two deaths and 18 hospitalizations between 2000 and 2002, revealed unsanitary conditions in the growing and packaging of cantaloupes in Mexico (CDC, 2002). This eminent risk led to a ban on imported Mexican cantaloupes. FDA investigations point out that contamination can be attributed to a broad source of factors such as irrigation of fields with water contaminated with sewage, poor hygienic practices of workers who harvest the produce, pests in packing facilities, and a lack of adequate cleaning and sanitizing of equipment (CFSAN, 2003). Because of the ability of pathogens to grow on the rind and flesh of cantaloupe, they are considered a potentially hazardous food.

Like cantaloupes, strawberries have traditionally been a popular fruit for fresh use, freezing, and processing. According to the USDA, strawberry consumption has doubled since the 1970s and strawberry consumption is second only to apples, among all fruits consumed in the U.S. (Cook, 2002). When growing, the strawberry plant exposes the fruit to the soil, increasing risk of fecal contamination. In addition,

strawberries are very perishable due to spoilage by fungus (i.e., *Botrytis cinerea*) (Browne et al., 1984), and have a very fragile physiology that makes them vulnerable to physical damage and consequently leads to rapid decomposition that could potentially harbor pathogens. In April 1997, the FDA announced the recall of about 1.7 million pounds of frozen strawberries purchased by the USDA for the school lunch program after school children in six states may have been exposed to the hepatitis A virus by eating the strawberries (GAO, 2000; Hutin et al., 1999). Like the cantaloupe case, general concern about the safety of fresh strawberries affected demand for berries from all sources, creating estimated financial losses of about $40 million (Richards and Patterson, 1999).

The harvesting and most of the packaging of strawberries and cantaloupes are still done manually. Strawberries and cantaloupes are hand harvested by field workers. Cantaloupes are placed in a conveyor system that feeds into a loading vehicle from which they can be unloaded by hand, or wet-dumped into tanks of chlorinated water to decrease physical damage. Alternatively, field debris can be removed with the use of spray washers and wet brushes (UGA, 2006). Finally, cantaloupes are packed into bins for store distribution. Strawberries are carefully hand picked from the field, directly arranged in boxes and rushed to cooling facilities for prompt delivery to stores. Rinsing strawberries before consumption is not recommended because their natural protective outer layer can be removed, affecting the quality of the fruit and inducing rapid deterioration. Quality control and safety grading operations of both produce are done by visual inspection. Although Good Agriculture Practices (GAP)

guidelines during the harvesting and packaging to ensure safety of the produce can be employed, this process is a difficult and labor-intensive job, in which workers can misjudge the appearance of the fruit, especially when dealing with very large amounts of produce. Therefore, the elimination of the potentially contaminated produce can be faulty. Mistakes at this level can result in the distribution of contaminated produce that may cause foodborne illness outbreaks.

The impacts of the aforementioned problem on the fresh produce industry depend on how quickly the contamination problem can be corrected. These outbreaks not only affect consumers but also suppliers, who can no longer export the produce to the U.S. In addition, the negative publicity not only affects the country in question but other domestic and foreign suppliers selling their product at the time. Automated imaging systems to ensure safety and defect inspection of the produce have been proposed as a solution to this problem. In the following sections this technology is introduced.

1.4. Nondestructive Methods to Ensure Safety and Quality

It has been established that it is essential to reduce microbial or disease-causing hazards before fresh produce reaches the consumer. 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 error (Tao, 2002). The ability to detect and classify

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.

1.4.1. Hyperspectral and Multispectral Imaging

Hyperspectral and multispectral sensors are relatively new classes of sensors that are based on the principle of imaging spectroscopy. Hyperspectral data includes measurements at many different wavelengths; in contrast, multispectral data is limited to a small number of selected wavelengths. Unlike other methods that collect a single spectrum at one point on a sample, hyperspectral imaging records a volume of data that contains a complete spectrum for every point in an image of the sample. The resulting characteristic spectrum depends on the composition of the testing material, serving as a fingerprint or signature that can be used to identify the material. As a result, hyperspectral data offers more detailed information about a sample object. These data can be used to generate complex models to discriminate, classify, identify and quantify materials present in the image.

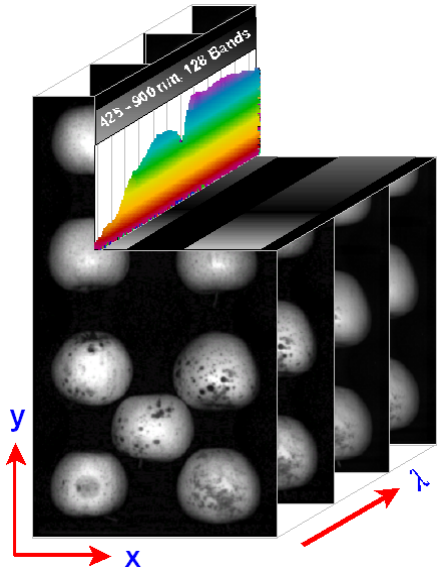


Figure 1. Hyperspectral Cube. This cube presents the data as a volume, composed of the spatial resolution and the number of contiguous spectral bands. In this manner, the spectral characteristics of the image can be visualized in a specific region of spectral space (Kim, 2000).

Hyperspectral Imaging Systems (HIS) generally include an illumination device, a spectrograph, a charge-coupled device (CCD) camera to acquire the images, and a computer to store the data. There are a number of common illumination sources. For agricultural produce, light in the visible, ultraviolet (UV), near infrared (NIR), infrared (IR), or a combination of these regions is typically used (Chen et al., 2002). The spectral region selected depends on the characteristics that the system must be able to identify. Subsequently, a few optimal wavelengths are selected to provide the most useful information about the characteristics of interest. These optimal wavelengths are then typically used in a multispectral imaging system, or in data analysis. A multispectral imaging system is much faster than a HIS because it acquires images only at the specific wavelengths of interest, rather than at many wavelengths over a broad spectral range.

There are many benefits to using these types of systems to evaluate agricultural products. The systems are non-destructive and, with adequate calibration for the product of interest, they can be very accurate. In addition, because light outside the visible spectrum can be used, features that are not visible to the naked eye can be detected. This technology has shown promise in detection for bio-terror agents, in water safety monitoring, and in food quality and safety control for fecal contamination detection on animal carcasses (News release: VerifEye^{TM} 2002 and patent 6,330,064). Spectral composition presents a wide range of major applications such as detecting contaminants in food, grading control for poultry and meat, and assorted processes in commercial settings for fresh produce and product quality control.

2. Hyperspectral Imaging Used For Food Safety

Detection systems based on hyperspectral imaging have been used to detect defects and diseases on the surfaces of meats, fruits, vegetables, grains and other agricultural products. The system can be used for either reflectance or fluorescence measurements. By observing fluorescence, fecal contamination has been detected on a variety of fruits and vegetables. This is accomplished by taking advantage of the fluorescent nature of feces at certain wavelengths. Similar systems can also identify bone fragments in meat (Marks et al., 1998). Nearly any attribute can be identified using HIS as long as a spectral region can be identified in which the feature of interest has reflective or fluorescent properties that differ from the surrounding regions.

Similar multispectral and machine vision systems have been used to inspect poultry to differentiate wholesome from unwholesome carcasses (Chen, 1993; Chen et al., 1998; Park et al., 1998). These systems have shown potential for use in on-line inspection systems and research in this area continues at the ISL.

Hyperspectral and multispectral imaging for food safety have been extensively studied at the ISL. In the laboratory, research has been conducted to develop imaging sensors and techniques to characterize the physical, chemical, biological and aesthetic properties of agricultural commodities. Kim et al. (2001) developed a rapid, noninvasive laboratory-based system to assess hazardous conditions in fresh produce. This Hyperspectral Reflectance and Fluorescence System (HRFIS) was used in this study.

The HRFIS utilizes a rapid, non-invasive method to assess hazardous conditions in fresh produce and is capable of both reflectance and fluorescence measurements. The imaging system covers spectral regions from 425 to 951 nm for reflectance measurements and from 425 to 760 nm for fluorescence emissions with 365 nm UV-A excitation at approximately 2 nm intervals. The spectral channels are separated at approximately 4.5 nm increments resulting in 117 and 79 channels for reflectance and fluorescence, respectively. The system is designed to be operated under dark conditions. Figure 2 shows the schematic diagram of the HRFIS. The system is composed of three modules: the sensor module, the optics module, and the lighting and sample module.

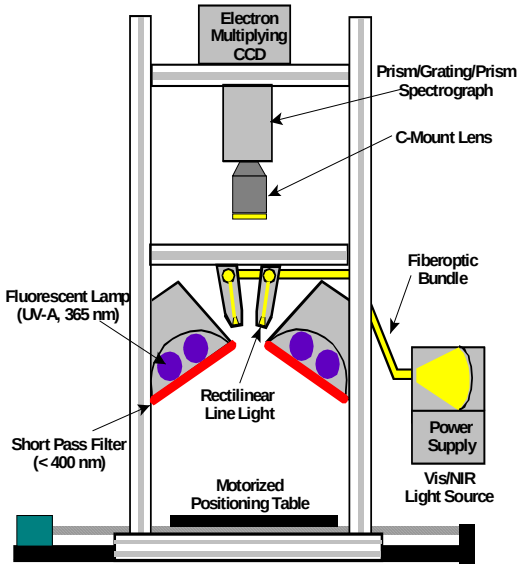


Figure 2. Schematic diagram of the USDA-ISL hyperspectral reflectance and fluorescence imaging system.

The sensor module includes a 16-bit digital imager containing an electron multiplying charge-coupled device (EMCCD) with 288 (V) X 560 (H) pixels. The 288 pixels in the spectral dimension are binned by 2 for image processing and analysis. The spatial dimension is reduced from 560 pixels to 460 pixels to eliminate the area at the edges of the field of view where lighting is non-uniform. In order to reduce noise, the digital imager is thermo-electrically cooled. The optics module contains a Nikon C-mount lens that disperses incoming radiation along each spatial location on the scan line into the spectral dimension. The lens is attached to the spectrograph, which is composed of a Prism Grating Prism (PGP) construction with a spectral resolution of approximately 7 nm full width at half maximum (FWHM). Lastly, two halogen lamps that provide near-uniform illumination for reflectance measurements and two UV-A fluorescence lamps for near-uniform excitation for fluorescence measurements make up the lighting and sample module. The UV-A fluorescence lamps are arranged

toward the line of field of view (FOV) at 30º backward and forward. This lighting assembly has been equipped with short-pass filters in order to prevent transmittance of radiation at wavelengths greater than approximately 400 nm. The imaging system can longitudinally accommodate a range of sample sizes of up to 30 cm wide. In order to displace the samples transversely through the line of the field of view, the system is equipped with a programmable positioning table. A more detailed description of the system operation is given by Kim et al. (2004), and calibration is given by Kim et al. (2001).

In a two-part study, the reflectance and fluorescence applications of multispectral detection of fecal contamination on apples using the HRFIS was described (Kim et al., 2002). Apple images were acquired in the visible to NIR spectral regions. This spectral range has been previously identified as a powerful tool identifying and detecting defects and contamination on agricultural produce (Miller et al., 1998; Upchurch et al., 1990; Upchurch et al., 1994; Tao, 1997; Throop et al., 1995). In the first part of the study, a systematic method for using hyperspectral data to identify wavelengths to be used in multispectral detection systems is presented. Additionally, spatial and spectral responses captured in hyperspectral reflectance images of apples contaminated with feces were evaluated. Specifically, principal component analysis (PCA) was used to identify regions of interest in the images and to define several optimal wavebands for multispectral detection. Results indicate that fecal contamination can be detected using reflectance imaging; however, in the second article, results indicate that fecal contamination detection on apples is more sensitive

using fluorescence imaging techniques. Natural spectral emission evaluation and PCA identified the same multispectral bands (450, 530, 685, and 735 nm) for discrimination of fecal spots. Moreover, simple wavelength ratios were found to reduce the variation in apple surfaces and increase the contrast between treated and untreated areas.

Fluorescence excitation and emission characteristics of fecal material from cattle, deer, swine, chickens, and turkeys were studied (Kim et al., 2003). In the study, naturally occurring fluorescent markers in animal feces were evaluated in detecting fecal contamination of foodstuffs, including animal carcasses and cut meats. Results indicated that fluorescence emission from *chlorophyll a* and its metabolites are good markers for fecal contamination. In contrast to meat tissue, animal feces show a distinctive red fluorescence; therefore fluorescence techniques can be used for detecting fecal contamination on agricultural meat commodities.

2.1. Analysis Methods for Hyperspectral Data

Hyperspectral imaging data analysis is one of the central topics of interest in sensing research. Despite the aforementioned advantages of hyperspectral imaging, the ultimate goal of this technology for food safety applications is to develop online automated systems that use multispectral techniques. Different mathematical algorithms have been widely used to solve the problems and challenges inherent in hyperspectral imaging analysis. A common approach to selecting significant wavelength channels in hyperspectral data is to evaluate the spectral response. The

spectral plots typically display the relative response versus the wavelength. Inflection points are then examined to detect materials of interest.

Although single waveband images do allow materials of interest to be distinguished, the visualization process is made difficult by spectral features of unwanted regions (false positives) on the image. In order to identify targets, increase image interpretation reliability and improve classification, many data image processing techniques have been developed, including image fusion. A general definition of image fusion is given by Genderen and Pohn (1994) as the combination of two or more images to form a new image by using a certain algorithm. Image fusion can be applied to various types of data sets, including those from single sensor and multi-sensor systems (Genderen and Pohl 1994), and its main purposes as applied to digital imagery are to sharpen images, improve geometric correlation, and enhance certain features not visible in either of the single images alone. Image fusion techniques are categorized by C. Pohl (1998) into two groups: color related techniques and statistical/numerical methods. For this research, we are interested in the second group. In general, the mathematical combinations of image channels in addition, subtraction, multiplication, and division (ratio) operations, and techniques such as principal component analysis (PCA) form a part of the statistical/numerical methods group.

Different mathematical operations provide ground to further enhance detection in the images. Linear and non-linear combinations of images, such as ratios, can be used to increase the intensity of the object detected and the background. Specifically, the ratio

method is useful because of its capability to emphasize slight variations in the spectral signature (Pohl and Genderen, 1998). Kim et al. (2002b) effectively isolated fecal regions of interest (ROI) from apples of different varieties using the image ratio technique. These band ratios can be performed using images at wavebands showing inflections in the spectra.

More mathematically involved methods include PCA, also known as Karhunen-Loeve Transformation. PCA provides means to identify patterns in images, encoding, image data compression, image enhancement, digital change detection, and image fusion, among others applications (Pohl and Genderen, 1998). This mathematical algorithm computes the best orthogonal basis of the unit vectors of x points. In the process of creating the PCA images, a correlation matrix of the image is calculated. This correlation matrix is a diagonal matrix, which is then used to compute the eigenvalues. The eigenvalues are equivalent to the variance of each principal component (PC) image. Let S represent the scatter matrix (Duda et al., 2001):

$$S = \frac{1}{n-1}\sum_{k=1}^{n}(x_k - m)(x_k - m)^t$$

where S is $\frac{1}{n-1}$ times the sample covariance matrix, x represents the sample vector for class *k*, and $m = \frac{1}{n_i}\sum_{k=1}^{n} x_k$ where m is the estimated mean of class *k*. It should be noted that the eigenvectors of the scatter matrix minimize the quadratic error under the restraint that these vectors form an orthogonal basis for a vector space with the

same dimension of the original data. Thus, the resulting matrix A is diagonalized by an orthogonal eigenvectors matrix E, and D is a diagonal matrix (Lay, 2002).

$$A = EDE^t$$

As explained by Pohl and Genderen (1998), PCA transforms the original data set into a set of new un-correlated linear combinations of the original variables. In this manner, PCA restructures the data so that most of the variance in the original data is accounted for in a reduced number of variables. The images are then ordered in terms of its variance, where PC-1 accounts for the largest variance.

Iterative Self-Organizing Data Analysis Technique (ISODATA) is an unsupervised classification method that deals with clustering the pixels into a desired number of classes; the method does not require any *a priori* information about the classes (Ball and Hall, 1965). In this algorithm, pixels with similar spectral characteristics are grouped into distinct classes by calculating class means evenly distributed in the data space; the remaining pixels are then iteratively clustered using minimum distance techniques until the number of pixel in each class change by less than the selected pixels change threshold. The sum of the square error (SSE) is commonly used as the clustering principle (Richards, 1986):

$$SSE = \sum_{C_i} \sum_{x \in C_i} (x - m_i)^t (x - m_i) = \sum_{C_i} \sum_{x \in C_i} \left\| x - m_i \right\|^2$$

Where m_i is the mean of the ith cluster, $x \in C_i$ is a pattern assigned to that cluster and the outer sum, $\sum_{C_i}$, is over all the clusters. The number of classes is selected based on the number of possible unique features found in the object.

The methods described above have been used successfully in imaging algorithms for target detection for different applications. The main intention for the evaluation of these methods in this study is to assess their performance when used for fecal contamination detection on cantaloupes and strawberries. Primary questions during the utilization of these methods were: What wavelength(s) provides optimal contrast between the different features in the image? How are the different sources of spectral variation illustrated in the images? What is the best band combination to eliminate undesirable areas and isolate fecal matter? What additional information can be learned from plots of the fluorescence intensities of ROI?

*3. **Objective**s*

3.1. The Rationale of the Thesis Objective

Recent studies in spectral imaging have revealed the potential of this technology to assess a myriad of conditions in agricultural commodities. Studies of hyperspectral sensing data are important to the fresh produce industry to address safety and quality control issues. This technology has been previously utilized with apples and other produce, including some vegetables, but it has not been explored yet with cantaloupes and strawberries. In general terms, the present study contributes to research with

relation to spectral imaging used in the development of on-line machine vision and multispectal imaging systems to rapidly and non-destructively assess hazardous conditions on agricultural commodities.

Recently, fecal contamination detection on fresh produce has become a subject of attention for the industry. It has been well established that contamination of fresh produce with pathogenic microorganisms can happen through contact with animal feces. The spectral contribution of all the biochemical components in feces from different animals should be addressed. On the other hand, it is an arduous process beyond the scope of this investigation. This research was confined to bovine fecal matter, which is widely used in agriculture practices for soil fertilization purposes, and it has been previously analyzed in similar research.

3.2. Objective Statement

The objective of this thesis is to study the potential for using hyperspectral imaging techniques to detect fecal contamination on cantaloupes and strawberries. This research is focused on identifying effective wavebands from hyperspectral imaging, and subsequently using these wavebands for pattern recognition and detection of fecal contamination. The experimental methods used for both commodities are the same; however, due to the different physical characteristics of cantaloupes and strawberries, such as fruit texture and color, different observations were expected.

Chapter 2: Articles

1. Detection of Fecal Contamination on Cantaloupes Using Hyperspectral Fluorescence Imagery [1]

1.1. Abstract

To determine if detection of fecal contamination on cantaloupes is possible using fluorescence imaging, hyperspectral images of cantaloupes artificially contaminated with a range of diluted bovine feces were acquired from 425 to 774 nm in responses to ultraviolet-A (320 - 400 nm) excitation. Evaluation of images at emission peak wavelengths indicated that 675 nm exhibited the greatest contrast between feces contaminated and untreated surface areas. Two-band ratios compared to the single-band images enhanced the contrast between the feces contaminated spots and untreated cantaloupe surfaces. The 595/655, 655/520, and 555/655-nm ratio images provided relatively high detection rates ranging from 79 to 96 % across all feces dilutions. However, both single band and ratio methods showed a number of false positives caused by such features as scared tissues on cantaloupes. Principal component analysis (PCA) was performed using the entire hyperspectral images data. Second and fifth principal component (PC) images exhibited differential responses between feces spots and false positives. The combined use of the two PC images

[1] This article is published in the Journal of Food Engineering. Vol.70(8-2005):471-476. It can be found in the appendix section.

demonstrated the detection of feces spots (e.g., minimum level of 16-μg/ml dry fecal matter) with minimal false positives. Based on the PC weighing coefficients, the dominant wavelengths were 465, 487, 531, 607, 643, and 688 nm. This research demonstrated the potential of multispectral-based fluorescence imaging for on-line applications for detection of fecal contamination on cantaloupes.

1.2. Introduction

Spurred by multiple outbreaks of foodborne illness associated with fresh produce and the identification of fecal contamination as the major source of pathogens responsible for these outbreaks, the United States Department of Agriculture (USDA) has instituted a number of programs to attempt to reduce fecal contamination of food products. In general, fruits and vegetables can become contaminated with pathogens through contact with soil, animals, or humans during any stage of the food-handling chain (FDA, 2001), including growing and harvesting operations as well as while in the processing plants (Murdock and Brokaw, 1957). Since the Food and Drug Administration (FDA, 1998) issuance of the "*Guide to Minimize Microbial Food Safety Hazards for Fresh Fruits and Vegetables,*" the food processing industry has made concerted efforts to address microbial food safety hazards. Generally, foodborne pathogens originate from the intestinal tracts of animals and humans (FDA, 2001), thus making fecal matter a major source of contamination. For example, fecal contamination can result from use of manure-based fertilizers, unsanitary conditions in packaging plants, or through random natural events.

Outbreaks of salmonellosis have been epidemiologically linked to consumption of fresh cantaloupes (FDA, 2001); as a consequence, this produce has been targeted as a potentially hazardous food. Contaminated cantaloupes were found responsible for 2 deaths and 18 hospitalizations due to *Salmonella* between 2000 and 2002; subsequent investigations revealed unsanitary growing and packaging conditions (Anderson *et al*, 2002; FDA, 2002). FDA investigations noted a broad range of potential factors, such as field irrigation with sewage-contaminated water, poor hygienic practices of workers, pests in packing facilities, and inadequate cleaning and sanitizing of equipment (FDA, 2001).

The Instrumentation and Sensing Laboratory of the USDA in Beltsville, Maryland, has developed a hyperspectral imaging system (HIS). The system provides a non-invasive method to assess hazardous conditions for fresh produce, is capable of acquiring reflectance and fluorescence measurements, and has been successfully used to demonstrate the detection of fecal contamination on apples and other fresh produce (Kim *et al*, 2001). Hyperspectral imaging systems can simultaneously collect spectral data for hundreds of narrow contiguous wavebands over regions of the electromagnetic spectrum at every spatial pixel in an image. The resulting spectra can serve as fingerprints or signatures for target identification. For food safety applications, a central goal of using hyperspectral imaging data is to find several spectral bands that can be implemented to the development of multispectral inspection systems for on-line applications at processing plants (Chen *et al*, 2002).

Many image-processing techniques have been developed to identify targets, increase image interpretation reliability and improve classification. Image fusion is one such technique, defined by Van Genderen and Pohl (1994) as “the combination of two or more images to form a new image by using a certain algorithm”. Applicable to various types of data sets, its main purposes for digital imagery are to sharpen images, improve geometric correlation, and enhance certain features not readily apparent in single images. Image fusion includes mathematical combinations of spectral images in arithmetic operations, and techniques such as principal component analysis (PCA).

For this investigation, hyperspectral fluorescence images were collected from cantaloupes artificially contaminated with bovine feces at varying concentrations. The objective of the present study was to identify a few wavelength bands that could be used to detect fecal contamination on cantaloupes. Candidate wavebands were subjected to two-band ratio permutations, and to further enhance the detection process the band ratio images were subjected to unsupervised classification. In addition, the entire hyperspectral images were subjected to PCA, and principal component (PC) images were evaluated for fecal contamination detection and determination of multispectral bands.

1.3. Materials and Methods

1.3.1. Sample preparation

Forty Western Shipper cantaloupes were purchased from a local supermarket. To increase the number of samples and prevent sample movement during HIS scanning,

cantaloupes were cut in half and the 80 halves placed cut-side down in batches of 8 on black trays. Hyperspectral images were first collected for all 80 samples prior to any treatment. Fresh feces from the USDA dairy farm in Beltsville, Maryland were diluted by weight to 1:10, 1:50, 1:100, 1:300 and 1:500 with deionized water. Dry fecal matter concentration for 1:100 dilution was 16 μg/ml as determined by drying samples to constant weight in a 90˚C oven. Using a variable pipette, a matrix of dilutions was applied to each sample at 40, 30, 20, and 10 μl (Figure 3).

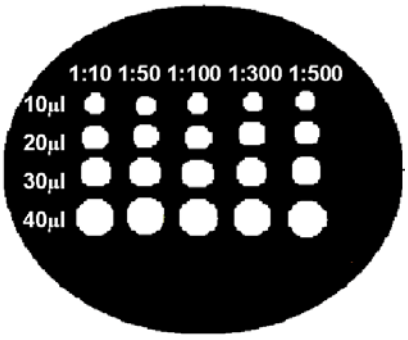


Figure 3. Schematic illustration of sample treatments. Using a variable pipette, dilutions of bovine feces at 1:500, 1:300, 1:100, 1:50 and 1:10 dilutions were applied to the cantaloupe halves in volumes of 10, 20, 30 and 40 μl. Fecal matter content for 1:100 dilution was 16 μg/ml as determined by drying sample to constant weight in a 90˚C oven.

Samples were allowed to air dry with the aid of a fan for several hours, until fecal spots were completely dried, and then scanned again. It was observed that areas contaminated with the 1:10 dilution were usually visible to the human eye. As dilutions increased, spots became transparent and less visible. At 1:100 dilutions, contaminated spots became difficult to visually identify. Digital color photos of samples were acquired prior to taking hyperspectral images.

1.3.2. Hyperspectral Fluorescence Images Acquisition

A hyperspectral imaging system incorporating a line scan spectrograph with a spectral resolution of approximately 7-nm full width at half maximum (FWHM) was used to acquire images (Kim *et al*, 2001). A motorized table was used to move samples at 1-mm intervals and the zoom lens was adjusted to yield a pixel resolution of 1 mm^2. The effective spectral range for fluorescence imaging was from 425 nm to 774 nm, with about 4.5-nm intervals, which was captured using 79 spectral channels. Two continuous wave UV-A (360 nm) fluorescent lamp assemblies provided the samples with a near-uniform illumination. The system was operated under dark conditions. A more detailed description of the system operation is given by Kim *et al* (2004), and calibration is given by Kim *et al* (2001). Each batch of eight cantaloupes required approximately 5 minutes to scan. The line scan data was saved in 16-bit binary files and processed later to create hyperspectral image cubes containing spatial and spectral data. Software developed in Visual Basic version 6 (Microsoft, Seattle, WA) was used to acquire data, and to perform further processing. Additional data analyses were done using ENVI version 3.2 (Research System Inc., Boulder, CO).

1.3.3. Data Processing and Analyses

The spectral dimension was smoothened using a 3-point running average prior to image visualization or analyses. Masks to eliminate background were created manually using threshold values determined by evaluating background pixel intensities. The largest possible rectangular region of interest (ROI) within each fecal spot was used to extract representative spectral data. Images taken before treating the

samples were used to create spectra of uncontaminated surfaces. ROI for untreated cantaloupes encompassed the largest rectangular area within the sample. The averaged ROI fluorescence spectra were used to characterize fluorescence emission of feces and untreated cantaloupe surfaces. Images at individual wavebands where spectra showed emission peaks were visually evaluated for efficacy of fecal contamination detection. Fecal spots for all dilutions and volumes were visually tallied from the selected single-band images.

All possible ratio permutations of emission peak and valley wavelengths were generated. These ratio images were visually assessed for fecal contamination detection and the best ratio images were selected. Subsequently, selected ratio images were subjected to an unsupervised classification method, ISODATA (Iterative Self-Organizing Data Analysis Technique), which clusters pixels into a desired number of classes without any *a priory* information about the classes (Ball and Hall, 1965). Pixels with similar intensities are grouped into distinct classes by calculating class means evenly distributed in the data space; the remaining pixels are then iteratively clustered using minimum distance techniques until the number of pixel in each class change by less than the selected pixels change threshold. The sum of the square error (SSE), commonly used as clustering principle, is expressed as:

$$SSE = \sum_{C_i} \sum_{x \in C_i} (x - m_i)^t (x - m_i) = \sum_{C_i} \sum_{x \in C_i} \left\| x - m_i \right\|^2$$

where m_i is the mean of the ith cluster, $x \in C_i$ is a pattern assigned to that cluster and the outer sum, $\sum_{C_i}$ is sum over all the clusters (Richards, 1986). Only one-iteration

was used to recalculate means, and the threshold parameter was set at 5%. The number of classes was selected based on the number of possible features found on the surface of cantaloupes, such as rind and vein tracts, bruises, cuts, variations in coloration, and fecal spots. For this study, classes ranged between 5 and 10, with no less than 3 pixels per class. Subsequently, several classes in the resulting images were merged to produce a binary image of feces contaminated regions and background.

In addition, the entire image data set (79 channels) was subjected to PCA. In the process of creating the PCA images, a correlation matrix of the image is calculated. This correlation matrix is a diagonal matrix, which is then used to compute the eigenvalues. The eigenvalues are equivalent to the variance of each PC image. The resulting matrix A is diagonalized by an orthogonal eigenvectors matrix (E), and D is a diagonal matrix (Lay, 2002).

$$A = EDE^t$$

As explained by Pohl and Van Genderen (1998), PCA transforms the original data set into a set of new un-correlated linear combinations of the original variables. In this manner, PCA restructures the data so that most of the variance in the original data is accounted for in a reduced number of variables. The images are then ordered in terms of variance sizes, where first PC accounts for the largest variance.

1.4. Results and Discussion

1.4.1. Spectral Responses

Figure 4 shows the averaged fluorescence spectra for areas treated with 40 μl of fecal contamination at 1:10, 1:50, 1:100, and 1:300 dilutions, and for untreated sample areas. For the feces treated spots, the average size of the rectangular ROI was 52 pixels (with a minimum of 20 and maximum of 99). For untreated spots, the average size of the rectangular ROI was 13,845 pixels (with a minimum of 11,880 and maximum of 15,720). Fecal spots for 1:500 dilutions were difficult to identify, suggesting that they were similar to uncontaminated areas. For this reason, the averaged spectrum for dilutions of 1:500 was not shown in (Figure 4).

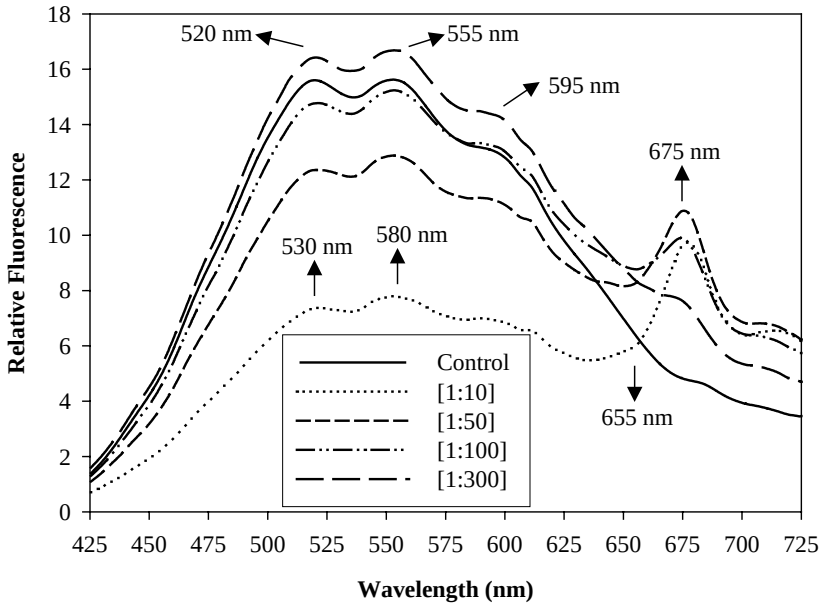


Figure 4. Representative fluorescence spectra obtained from the region of interests (ROI). Note that spectrum for 1:500 dilution spot was not included since the spectral characteristics were similar to those of the cantaloupe surfaces.

The spectra show fluorescence emission peaks in the green region at 520, 555, and 595 nm, and an additional peak for the feces treated spots in the red at 675 nm. At

high feces concentrations, 1:10 and 1:50 dilution spots, there was a distinct response in the green region due to the presence of feces in that relative fluorescence intensities were lower compared to the 1:100, 1:300 and 1:500 dilution spots. This observation was the results of reabsorption characteristics of animal fecal matter existing in relatively high concentrations (Kim *et al*, 2004). The application of the 1:100 and 1:300 dilutions resulted transparent feces spots, and the fluorescence responses in the green bands resembled those of untreated areas.

Fluorescence emissions in the 650 to 750 nm region from intact green plant materials are due to membrane-bound chlorophyll *a* with emission maxima at 685 nm and 730 nm (Papageorgiou, 1975). The averaged spectrum for untreated cantaloupe surfaces exhibited very low chlorophyll *a* fluorescence emission, indicating a well-ripen state of the samples. However, animal fecal matter showed blue-shifted emission peak at 675 nm, which emanated from chlorophyll *a* and its by-products such as pheophobide *a* (Kim *et al*, 2003). Multiple plant constituents, including phenolic compounds and riboflavin, are responsible for the fluorescence emission in the blue-green region of the spectrum (Chappelle *et al*, 1991; Kim *et al*, 2001). Other factors that may affect fluorescence emissions include changes in accumulation of the anthocyanin pigments in association with fruit ripeness (Abbott *et al*, 1997).

Figure 5 shows representative gray-scale fluorescence images of cantaloupes contaminated with feces at 520, 555, 595, and 675 nm. These wavelengths correspond to fluorescence emission maxima observed in the representative fluorescence spectra

(Figure 2). In general, the images illustrated progressive decreases in intensities from the center portions toward the edges due to the hemispherical shapes of the cantaloupe halves. In addition, intensity variations in localized regions (asides from the feces treated spots) were noted, suggesting heterogeneous nature of the cantaloupe surfaces in fluorescence responses. In the green region bands at 520, 555, and 595 nm, 1:10 and 1:50 dilution spots were shown darker than surrounding cantaloupe surfaces. For 1:300 and 1:500 dilution spots, the visual identification was difficult; as these dilutions created transparent feces contaminated spots, the fluorescence responses started to blend with those of cantaloupe surfaces. Fluorescence responses at 675 nm were markedly different than those of the green bands in that feces contaminated spots were brighter than cantaloupe surfaces. Similar responses were observed on apples contaminated with bovine feces and this observation was attributed to additive effects of fluorescent animal feces and apples (Kim *et al*, 2003).

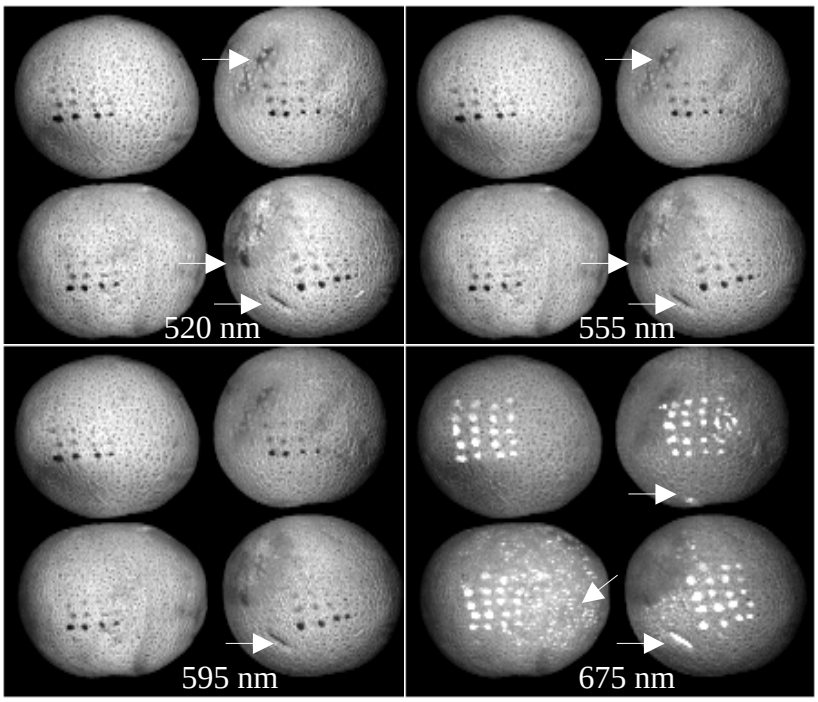


Figure 5. Fluorescence images of bovine feces treated cantaloupes acquired using the ISL hyperspectral imaging system. The wavelengths correspond to fluorescence emission maxima observed in the representative spectra. Arrows indicate some potential false positives in the images. Cantaloupe in the bottom right corner shows a physical damage (scared tissues).

Detection results based on the visual assessment of the 520, 555, 595, and 675 nm bands for 40 μl spots are shown in Figure 6. Results for other feces volumes were similar and omitted for brevity. All the wavelength images showed evidence (95-100 % detection rates) of the 1:10 feces dilution spots. The 675 nm band demonstrated the best potential for detection of feces across the range of volumes and concentrations tested (95 to 73 % detection rates for 1:10 to 1:500 dilution spots, respectively); although 100 % of the 1:10 dilution spots could be detected at the green bands, detection rates for 1:300 and 1:500 dilutions were zero. Based on the single-band imaging, most of the cantaloupes exhibited a number of false positives (Figure 5) which could be attributed to physical damage such as bruises or cuts (scared

tissues). False positives in the red (650 - 700 nm) region of the spectrum may also emanate from netted rind and vein tracts that contained chlorophyll *a*. For the green band images, the relatively lower fluorescence intensities of the edge portions of cantaloupes were also potential false positives.

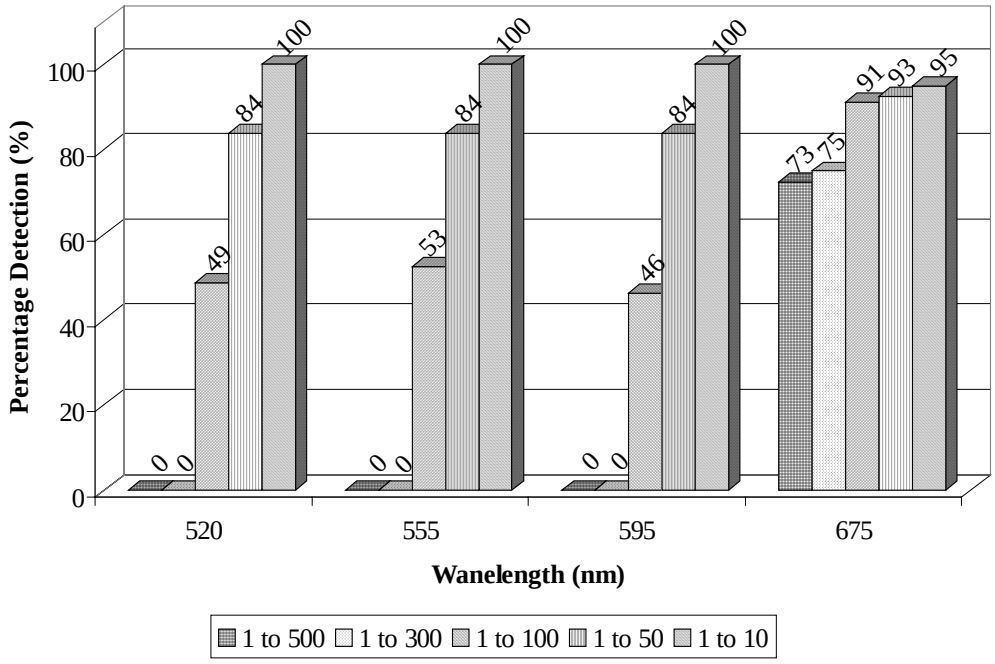


Figure 6. Fecal contamination detection rates for 40-µl treatment spots based on the single wavelength images at 520 nm, 555 nm, 595, nm, and 675 nm (n = 80 per dilution).

1.4.2. Ratio Images

Although single waveband images showed potential for detection of fecal matters on cantaloupes, mathematical combinations of images acquired at different wavelengths can enhance fecal contamination detection and reduce false positives. Addition and multiplication of images have been shown to enhance contrast, while difference or ratio images are suitable for detecting changes (Pohl and Van Genderen, 1998). Kim *et al* (2002) effectively demonstrated the use of ratio methods to isolated fecal

contaminated spots from apples of different varieties. For this investigation, the most promising results were also obtained with two-band ratio images. Thus, discussion is limited to those ratios that resulted in high detection rates for brevity.

Representative gray-scale, two-band ratio images (595/655, 655/520, and 555/655 nm), and ratio images subjected to the ISODATA method are shown in Figure 7a and 5b, respectively. These ratios compared to the single-band images enhanced the contrast between the feces treated spots and cantaloupe surfaces and produced more uniform responses across the cantaloupe surfaces. Detection results for 40-μl spots (n = 80 per dilution) for 595/655, 655/520, 555/655, and 675/555-nm ratio images are shown in Figure 8. The 595/655-nm ratio image followed by 655/520, 555/655, and 675/555 achieved relatively high detection rates across all feces concentrations. Note that ratio in Figure 7 images shown excluded the use of the best single band at 675 nm in the ratio combinations.

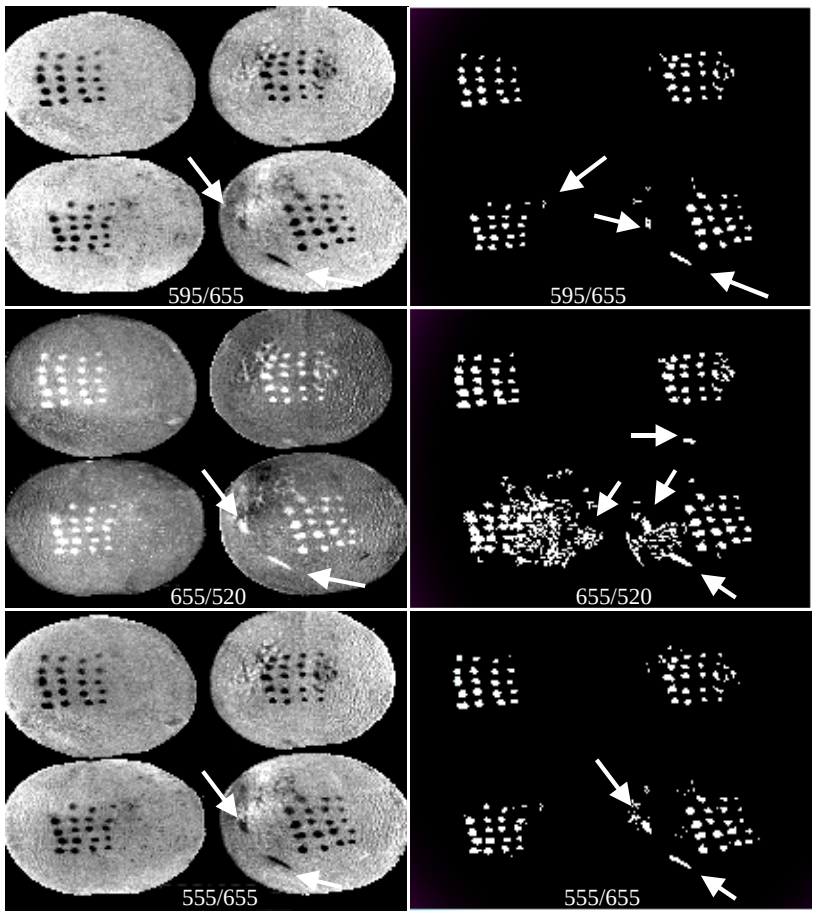


Figure 7. a) Representative two-band ratio images of 595/655, 655/520, and 555/655 nm. Out of all possible two-band ratio permutations of peak and valley wavelengths observed in the fluorescence spectra, these ratios provided the best contrast between the feces treated spots and cantaloupe surfaces and produced more uniform responses across the cantaloupe surfaces. b) Binary images for feces contamination spots obtained by subjecting ratio images to the ISODATA method. Arrows indicate some false positives. 595/655 and 555/655 nm ratios were the most effective in reducing false positives.

For reducing false positives, the 595/655 and 555/655-nm ratio images provided the most satisfactory results (Figure 7b) with detection rates ranging 80-95 % and 79-91 %, respectively for 1:500 to 1:10 feces dilution spots. With the use of ISODATA method, removal of clusters of three pixels or less resulted in the exclusions some false positives. However, false positives

were still apparent; the most prominent false positives were due to the scared tissues as indicated on the lower-right cantaloupe in Figure 7.

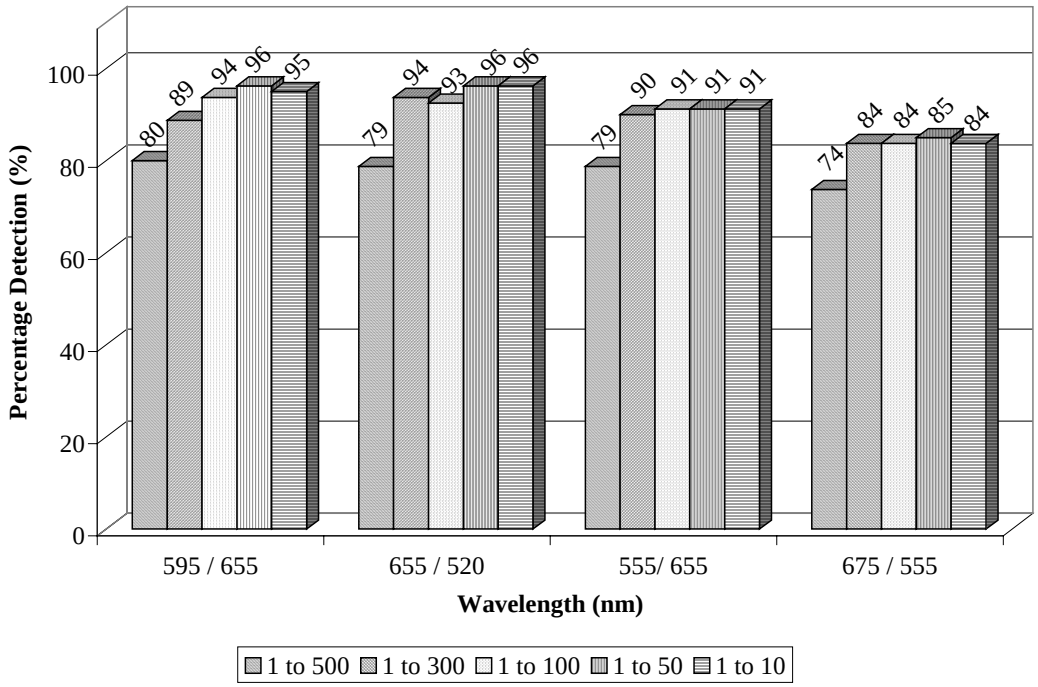


Figure 8. Fecal contamination detection rates for 40-μl treatment spots for ratio images of 595/655, 655/50, 555/655, and 675/555 nm subjected to the ISODATA classification method (n = 80 per dilution). Note that these ratios represent the two-band ratio combinations with the highest detection rates.

1.4.3. Principal Component Analysis

Figure 9 illustrates representative first PC (PC-1) to sixth PC (PC-6) images obtained from the PCA of the entire hyperspectral fluorescence image data. PC-1 through PC-6 accounted for 99.94 % of the data variability and images beyond PC-6 contained no useful attributes for detection of feces treated spots. PC-1 image reflects a weighted sum of all the spectral bands, and showed features causing the largest variations of the data; the intensity decreases from the center portions to the edges of the

cantaloupe surfaces. Subsequent PC images depicted other features affecting variations in spectral responses. For instance, PC-3, PC-4 and PC-6 images showed responses that may be attributes of color variations or sides in contact with the grounds (ground spots).

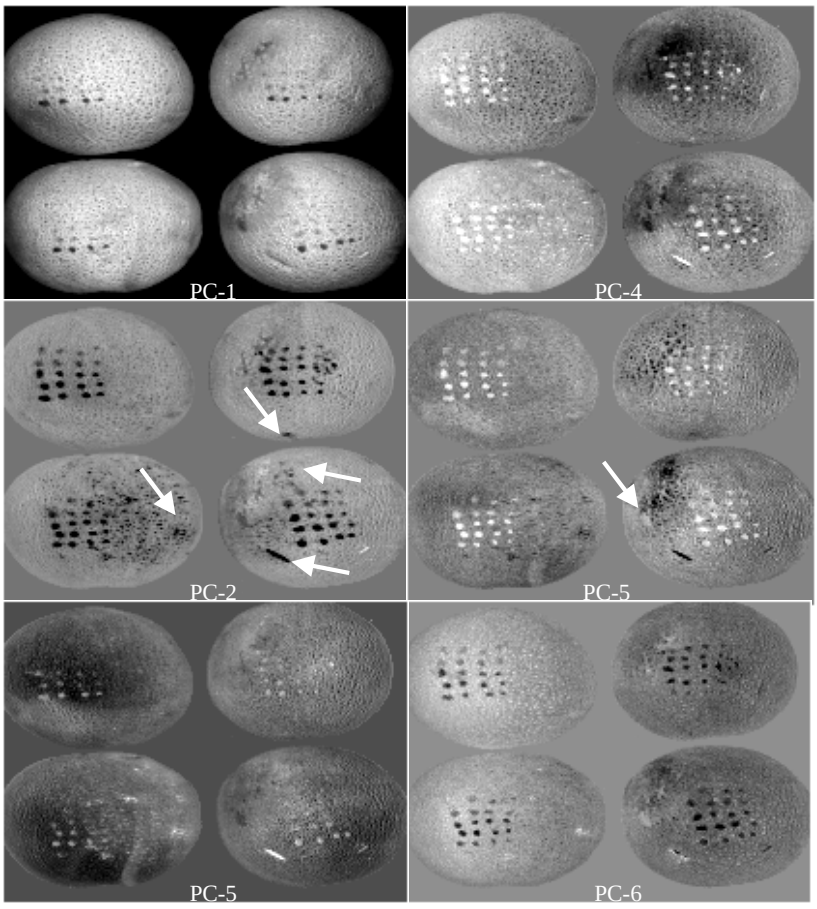


Figure 9. First to sixth principal component (PC) images of bovine feces treated cantaloupes obtained from principal component analysis (PCA) of the entire hyperspectral fluorescence image data (79 spectral channels). Arrows indicate some of false positives.

PC-2 and PC-5 images exhibited the evidence of the feces treated spots as darker and lighter spots, respectively, in contrast to the cantaloupe surfaces. In addition, PC-5 showed the scared tissues (on the lower-right cantaloupe) that were false positives in

the single-band and ratio images as non-false positives. PC-2 and PC-5 images were subjected to a simple thresholding method (Kim *et al*, 2004) to create binary images for fecal contamination spots (Figure 10a and 8b). These images exhibited false positives. However, the PC-2 false positives did not coincide with those of PC-5. Figure 10c shows the spots (pixels) where the PC-2 and PC-5 binary images overlapped. Although the combined use of PC-2 and PC-5 compromised the detection of some 1:300 and 1:500 spots, the resultant image in Figure 8c illustrated the 1:10, 1:50 and 1:100 spots (a minimum of 16-μg/ml dry fecal matter) with minimal false positives.

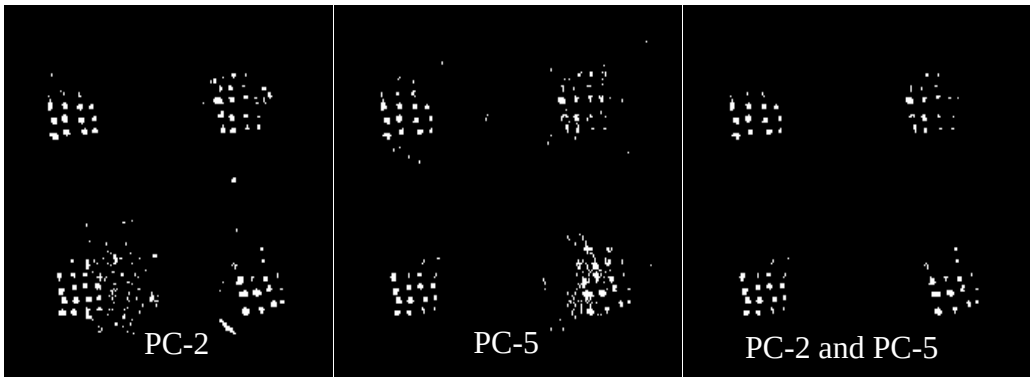


Figure 10. Binary images for feces contamination spots obtained by subjecting a) PC-2 and b) PC-5 images to a simple thresholding method (Kim *et al*, 2004). c) Binary image shows where both PC-2 and PC-5 binary results overlapped. As the resulting binary image illustrates, pixel locations of the PC-2 false positives did not coincide with those of PC-5.

The PCA was performed using the data in entire spectral regions. However, individual PC images could be approximated by the use of few significant wavelengths (Kim *et al*, 2002).

Figure 11 shows weighing coefficients (eigenvetors) for the PC-2, and PC-5 images. The peaks and valleys indicated the dominant wavelengths, 487 and 607 nm for PC-2, and 465, 531, 643 and 688 nm for PC-5. Weighted sums of the original images at the dominant wavelengths can create near identical images comparable to the PC-2 and PC-5 images, respectively. Above wavelengths can be implemented to a multispectral imaging system for on-line applications. Readily available common aperture-based, multispectral adaptors can be used to simultaneously capture images up to 8 spectral bands (Chen and other, 2002). Further research is needed to incorporate samples with various stages of maturity and to develop automated processing and detection methods for on-line applications.

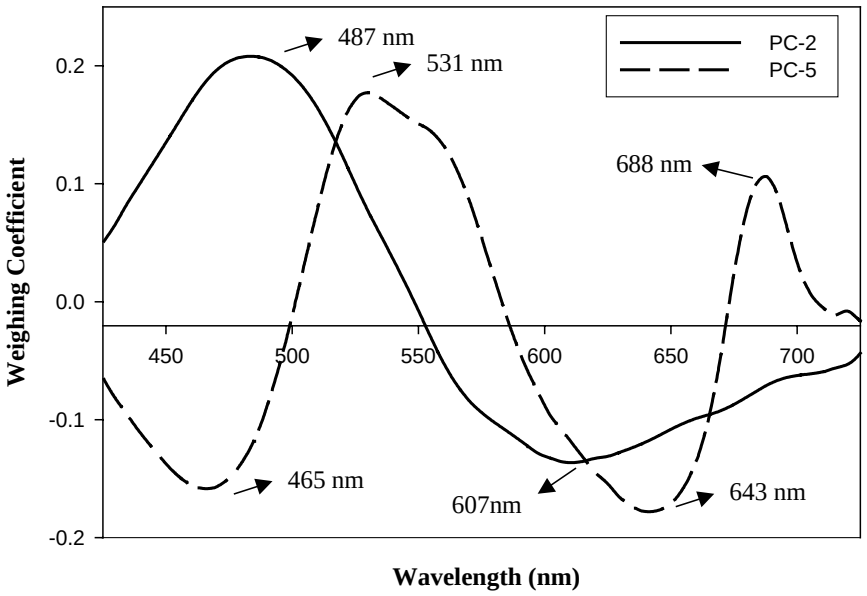


Figure 11. Spectral weighing coefficients (eigenvectors) for PC-2, and PC-5. The dominant wavelengths are indicated on the graph.

1.5. Conclusion

In this investigation, hyperspectral fluorescence images were evaluated for detection of fecal contamination on cantaloupes. Single waveband images showed natural variations in fluorescence responses of the cantaloupe surfaces, and were not suitable for detection of fecal contamination on cantaloupes due to the presence of false positives. Images constructed using two-band ratios enhanced the contrast between the fecal contaminated spots and cantaloupe surfaces. However, ratio images also exhibited false positives emanating from features found on cantaloupes such as scared tissues. The PCA of hyperspectral images demonstrated the potential detection of feces contaminated spots (a minimum of 16-μg/ml dry fecal matter) on cantaloupes with minimal false positives. Examination of PC weighing coefficients defined several dominant wavelengths that can be implemented to a multispectral imaging system for on-line applications.

2. Hyperspectral Analysis of Fecal Contamination Detection on Strawberries[2]

2.1. Abstract

Spectral imaging has shown potential for assessing quality and safety conditions of fresh produce in a rapid, non-invasive manner. In this study, hyperspectral fluorescence images between 425 and 775 nm of strawberries artificially contaminated with bovine fecal matter and placed under Ultraviolet-A excitation were analyzed. Evaluation of the fluorescence spectra using Principal Component Analysis (PCA) identified 553, 597 679 and 736 nm as optimal wavelengths to classify small areas of diluted feces on the strawberries. Visual assessment and fluorescence intensity plots indicated that the image at 679 nm exhibited greatest contrast for all diluted feces spots on strawberries. A major impediment in the fecal spot classification process was similar intensity between the fecal spots and the leafy calyx of the fruit. The two-band ratio image at 679/730 nm improved the detection rates of fecal-contaminated areas by reducing the number of false positives and increasing the contrast between the surface, the fecal-contaminated areas and the calyx. Using PCA over the entire spectrum (79 channels), PC images 2 and 3 showed the best responses for detecting the fecal-contaminated areas and calyx. The combined use of these two PC images increased the detection rates for all concentrations and decreased the presence of false positives. This study confirmed the capability of hyperspectral imaging in detecting fecal matter on strawberries and the potential for this method to be used for developing on-line applications.

[2] This article is submitted to the Journal of Food Engineering on 2006 (in peer review)

2.2. Introduction

Fecal matter contamination of produce has become a major concern for the produce industry. According to the Food and Drug Administration (FDA), fecal contamination is a major source of human pathogens associated with fresh produce (FDA, 1998). Strawberries have traditionally been a popular fruit for fresh use, freezing, and processing. According to the United States Department of Agriculture (USDA), strawberries are the fourth most valuable fruit produced in the United States (Bertelsen, 1995). In addition to its appealing flavor, a major stimulus to strawberry consumption is its year-round production and its health benefits (Törrönen, 2002). The growing pattern of the strawberry plant exposes the fruit to the soil, increasing fecal contamination risk. In April 1997, the FDA announced the recall of about 1.7 million pounds of frozen strawberries purchased by the USDA for the school lunch program after school children in six states were exposed to the hepatitis A virus by eating the strawberries (GAO, 2000; Yvan *et al*, 1999). General concern about the safety of fresh strawberries affected the demand for berries from all sources, creating estimated financial losses of about $40 million (Richards and Patterson, 1999).

Current methods of strawberry inspection involve visual and manual examination by human inspectors. Research studies addressing quality and safety issues in the fresh produce industry have considered the implementation of machine vision based on multispectral imaging in processing plants as part of a Hazard-Analysis-and-Critical-Control-Point (HACCP) inspection process. Online machine vision inspection systems to screen and discard unsafe produce in the processing line could improve the

effectiveness of the inspection, increase consumers' confidence about produce safety, and increase processing productivity. In particular, machine vision potentially allows the handling of large quantities of fresh produce to assess quality issues in a nondestructive and rapid manner. Current research at the USDA Instrumentation and Sensing Laboratory (ISL) is evaluating hyperspectral and multispectral fluorescence imaging systems and techniques for online inspection of apples (Kim *et al.*, 2002), cantaloupes (Vargas *et al.*, 2005) and vegetables for fecal and soil contamination, diseased surfaces, and open skin cuts and bruising. Hyperspectral fluorescence imaging analysis has been used to select a few candidate wavelengths for the development of multispectral imaging systems for inspecting contaminated produce. The development of methods for identification of fecal matter using hyperspectral and multispectral technology has been based on analyses of responses from single band intensities, simple two-band ratios and Principal Component Analysis (PCA) (Kim *et al*, 2002). With the use of hyperspectral fluorescence imaging techniques, the objectives of this study are to determine a few suitable wavelength bands that can be used to detect bovine fecal contamination on strawberries.

2.3. Materials and Methods

2.3.1. Strawberry sample preparation

A total of 162 randomly selected Earliglow strawberries purchased from a local supermarket were used for this investigation. To prevent sample movement during the image acquisitions, each batch of 27 samples was placed on a custom-made holder covered with non-fluorescent black cloth (in a 3 x 9 sample arrangement). Samples of fresh cow feces obtained from the USDA dairy farm in Beltsville, MD were diluted 1:2, 1:25, and 1:100 by weight with deionized water (concentration for 1:25 dilution was 15.78 μg/ml as determined by drying sample to a constant weight in a 90˚C oven. The feces dilutions were applied to the strawberries in 10-μL spots (Figure 12), and the samples were allowed to air dry for 30 minutes. Treated spots at 1:25 and 1:100 dilutions became transparent and difficult to visually identify. An additional 27 untreated strawberries were used as control samples.

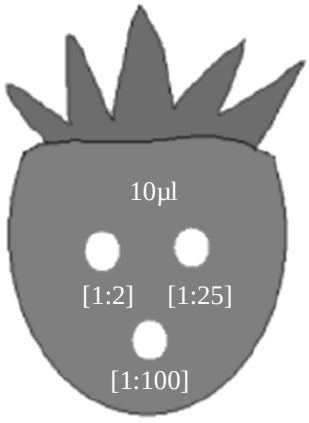


Figure 12. Schematic illustration of sample treatment.

2.3.2. Hyperspectral Fluorescence Imaging

Hyperspectral imaging data of strawberries were collected using the ISL hyperspectral imaging system. The hyperspectral imaging system incorporates a line scan spectrograph with a spectral resolution of approximately 7-nm full width at half maximum (FWHM). A motorized table was used to move samples at 1-mm intervals and the zoom lens was adjusted to yield a pixel resolution of 1 mm^2. The effective spectral range for fluorescence imaging was from 425 nm to 774 nm, with about 4.5-nm intervals, which was captured using 79 spectral channels. Two continuous wave UV-A (360 nm) fluorescent lamp assemblies provided the samples with a near-uniform illumination. The system was operated under dark conditions. Additional description of the spectral and spatial calibration can be found in Kim *et al*, 2001. Each batch of 27 strawberries took approximately 3 minutes to scan. Individual line-scan data were saved in 16-bit binary files and were subsequently processed to create hyperspectral image cubes containing spatial and spectral data. Data acquisition, processing, and analyses were performed using Software developed in Visual Basic version 6 (Microsoft, Seattle, WA). Additional analyses were done using ENVI version 3.2 (Research Systems, Inc., Boulder, CO).

2.3.3. Data Processing and Analysis

All image processing and data analysis was performed on a merged hyperspectral image data set containing data for all 135 treated strawberries. Prior to image visualization and analyses, the spectral dimension was subjected to a 3-point running average for all treated and control batches. Next, in order to eliminate the

background and only visualize the strawberry leaves and the treated and untreated regions of the strawberries, a masking image was created by applying a threshold based on relative fluorescence intensity at the 679 nm waveband of the control sample data set. Investigators have observed that vegetation, when exposed to UV radiation, show maxima fluorescence emission at 440, 525, 685 and 740 (Chappelle *et al.*,., 1984). This information along with visual assessment of the images indicated that the image at 679 nm exhibited significant contrast for the leaves and strawberries surface. Therefore, this waveband was utilized to generate a background-masking image. Because the histogram of pixel intensities showed a bimodal distribution that corresponded well to the two regions to be separated (dark image background vs. strawberry surfaces and leaves), a threshold intensity value suitable for separating the two regions could be approximated through visual examination of the histogram.

The representative spectra for non-treated strawberries were acquired from three random rectangular regions of interest (ROI) of 50 pixels each in the surface of every control sample. For the treated samples, spectra were retrieved from ROIs of 30 pixels per concentration spot. For the leaf areas, ROIs contained 20 pixels each. The dimensions of the rectangular ROIs varied from sample to sample, but their areas (number of pixels) remained constant. This approach was used to encompass the larger rectangular area with identical number of pixels throughout the samples. To construct the representative emission spectra, the averaged response from the ROI for the leaves, control, and treated spots (1:2, 1:25 and 1:100) were plotted against the

wavelength. Images at maximum and minimum fluorescence emission wavelengths were visually evaluated for fecal spot detection.

The spatial profiles crossing features of interest (leaves and treated spots) of characteristic wavelengths used for classification were graphed. This approach was used to explore the spatial profile differences between the strawberries surface, treated spots and leaves. The fluorescence intensity measures of quantum yields and are expressed in terms of relative fluorescence intensities (RFI) (Kim *et al.*, 2001).

With the objective to further facilitate and enhance contrast for fecal contamination detection, all ratio combinations of the wavebands corresponding to the emission maxima and minima of the fluorescence spectra were generated. Resulting images were visually examined, and representative images were selected for further analysis. Unwanted regions were removed by evaluating the RFI and determining minimum and maximum values to be masked out. This method renders a binary image were the masked pixels appear black.

In addition, the entire unmasked data set (79 channels) was subjected to PCA. Individual PC images were evaluated to identify the fecal spots at all concentrations, and to identify an image that could provide opposite contrast between the strawberry surfaces, the leaves, and the fecal spots. Similarly to the ratio image, PC images with opposite contrast characteristics could serve as means to mask out undesirable features and positively classify the fecal spots. In PCA, the orthogonal projections that maximize the amount of data variance is calculated. This transform is based on

the eigen-decomposition of the covariance matrix of the data set (Weeks, 1998). The information can be presented in the form of principal component (PC) images, which are ordered in terms of decreasing variance, where first PC (PC-1) accounts for the largest variance. The maximum and minimum weighing coefficients (eigenvectors) inflections are indicative of dominant spectral regions (Kim *et al*, 2004).

2.3. Results and Discussion

Figure 13 shows the average fluorescence spectra of control (untreated) ROIs, leaf ROIs, and treated ROIs at 1:2, 1:25 and 1:100 concentrations. All these ROIs exhibited a broad fluorescence emission in the green region at 553nm, and three distinct red fluorescence emission peaks at 679, 688, and 730 nm. Lawrence *et al* (1997) reported that fluorescence emission near 685 (red) and 740 nm (near-infrared) are governed by chlorophyll *a* concentrations and photosynthetic activity. Note that the treated spots compared to the leaf spectra exhibited a blue- shift from 688 to 679 nm. Fluorescence emission spectra for treated regions is similar in shape for 1:2 and 1:25 concentrations with relatively high emission, while emission for 1:100 resembled emission for control ROI. The leaf ROI exhibited higher fluorescence emission in the red band (between 675 and 735 nm). The fluorescence emission in the blue-green band with emission maximum at 553 nm is in the vicinity emission peak observed for riboflavin in plants at 525 nm (Chappelle *et al.*,, 1991). However, the blue-green band fluorescence emission remained relatively low in comparison to the red band fluorescence emission. The point at which the treated, control and leaves change direction was observed at 597 nm.

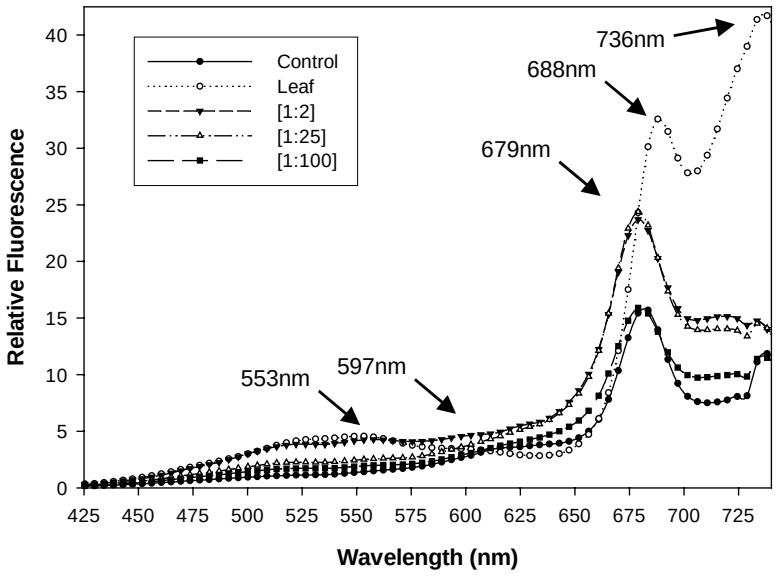


Figure 13. Representative fluorescence emission spectra for control ROIs, leaf ROIs and treated ROIs

Figure 14 shows the histogram for the pixel intensities of the merged florescence image at 679 nm containing all the treated samples. At this wavelength, the treated and control ROIs show a fluorescence emission maximum. This histogram was used to determine a cut off value to create a background mask for images. In the plot, two distinctive distribution regions can be observed. The first region corresponds to the relative fluorescence intensities (RFI) occurring for the background areas. In this region, there are very high frequencies of lower intensity pixels, within a relatively narrow intensity range. In the second region, there are relatively low frequencies of higher-intensity pixels over a much wider range of intensities. The highly variable responses in the second region correspond to pixels in the strawberry surfaces, leaves, and fecal contaminated spots. The cut off value for the background mask was

determined to be 180 by approximating the point at which the slopes of the two regions change.

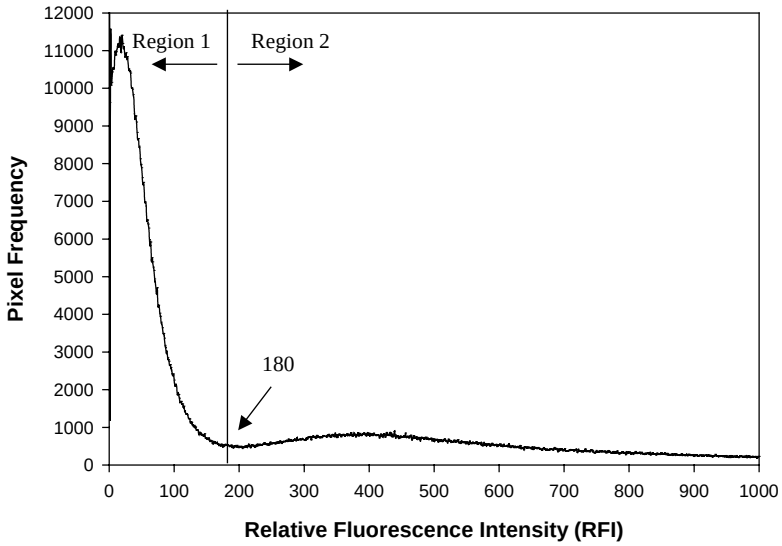


Figure 14. Histogram of control samples at 679 nm. The cut off value for the two regions at 180 (ROI) was calculated by approximating the point at which the slopes of the two regions change. Region 1 is representative of the background of the image and Region 2 is representative of the strawberries' surface and leaves

Figure 15 shows representative gray-scale images of feces-treated strawberries at 553, 597, 679 and 730 nm, where these wavelengths were selected based on the evaluation of the average fluorescence spectra (Figure 13). Note that the background was masked out using the 679 nm mask image obtained by the histogram threshold. At shorter wavelengths (553 and 597 nm), the images appear grainy and lower-concentration feces spots (1:25 and 1:100) were not clearly distinguishable from the strawberry surface. At 597 nm, the second bottom strawberry showed higher fluorescence emission from the leaf and upper part of the fruit compared to other strawberries and feces spots. Using these waveband images individually for detection

results in the potential for false positives: all fecal spots can be identified, but some non-fecal spots will also be misidentified as contaminated spots. In the percentage detection graph (Figure 21) high-concentration fecal spots (1:2) present high detection rates between 79% and 90% for 553 and 597 nm. Lower-concentration fecal spots (1:25 and 1: 100) performed rather poorly with detection rates as low as 21%, and not exceeding 56%. The images in the red band and near-infrared bands (679 and 730 nm) appeared smoother compared to the images at 553 and 597 nm and to all feces treated spots regardless of the concentrations that were visible. The contrast between the strawberry surfaces and the treated areas is the greatest in the 679 nm image compared to the other wavelength images. A major impediment in the classification process is similar RFI of the leaves, strawberry surfaces, and feces-treated spots. Variations in RFI for the control may be the result of the differences in fruit ripeness and constituent concentration. The arrows in the image in point at potential false positives. Overall, the 1:2 and 1:25 fecal concentrations exhibited brighter fluorescence emission in contrast to the 1:100 fecal spots. These plots demonstrate how the fluorescence emission is dependent upon spectral regions.

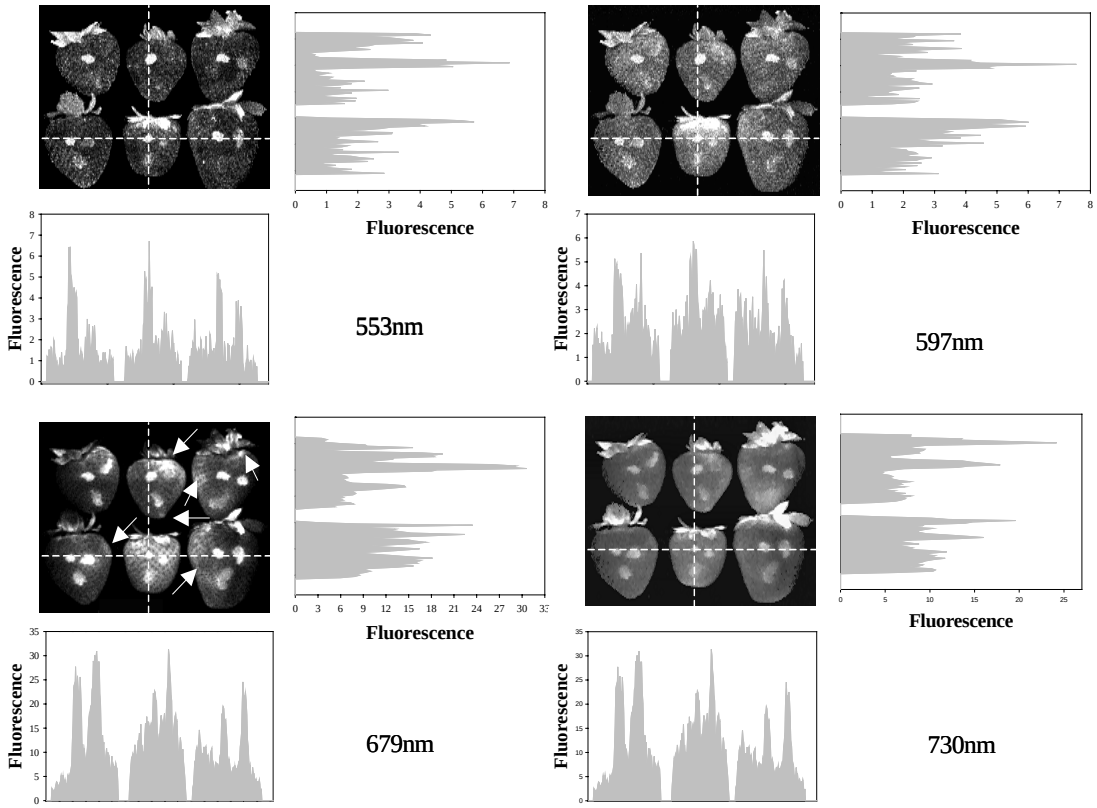


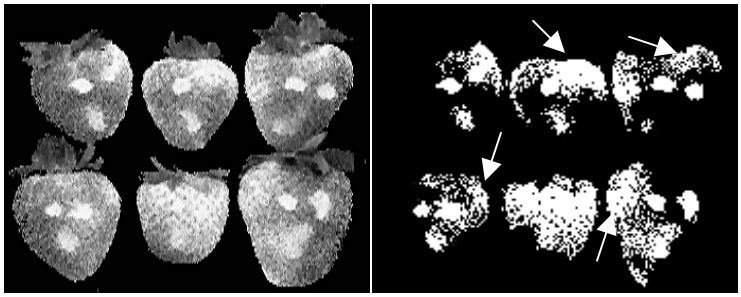
Figure 15. Gray-scale images of the strawberries at 553, 597, 679, and 730 nm. For each image a horizontal and vertical fluorescence intensity variation across the dash line is shown. The plot to the right of each image corresponds to the vertical profile, and the plot below each image corresponds to the horizontal profile. The horizontal line crosses the 1:2 and 1:25 fecal concentration spots, and the vertical line crosses the leaves, 1:2 and 1:100 fecal concentration spots. Arrows point at potential false positives.

The spatial intensity profiles that intersect the surface of the fruit, leaves, and fecal spots are shown in the plots adjacent to the four images in Figure 15. The plots illustrate the spatial responses along the vertical and horizontal dated lines. It was observed that on the fruit surface in non-treated regions, not including leaves, fluorescence emission was non-uniform indicating the presence of heterogeneous constituents concentrations. These heterogeneous constituents maybe dependent on the characteristics of fresh strawberries such as thin shiny fruit surface with abundant

moisture underneath, and ripening stage at harvest time. At their initial stage strawberries are greenish-white, but as they ripen, the fruit turns to an intense red color. The ripening process in strawberries is blocked if the fruit is harvested before it has ripened and the fruit may have not fully developed its uniform red color. When comparing the spatial intensity profiles that intersect the contamination spots for each of the four images, the 679 nm image more clearly contrasts the high-intensity fecal spots against the strawberry surfaces and leaves, with less variation in the surface/leaf areas. Higher fluorescence emissions were registered in the 679 and 730 nm wavebands. In all the plots, the fecal spots and leaf regions generally show higher emissions than the strawberry surfaces. Some other emission variations may be due to achenes, sun exposure side, and natural intensity variations previously discussed. In terms of fecal matter concentration, a gradual reduction of fluorescence emission was observed as fecal concentration decreased. Detection rates (Figure 21) for 679 and 730 nm at high fecal concentration (1:2) ranged between 75% and 78%. In contrast to emission on the red band (553 and 597 nm) these two wavelengths present higher detection rates (between 73% and 74%) for 1:25 fecal concentration spots, almost matching the detection rates for the 1:2 concentration. However, 679 nm presents a higher detection rate for low concentrations (1:100) than 730 nm. 679 nm detects 32% more than 730 nm. At 730 nm, the contrast between the fecal spots and the fruit surface is not sufficient to be able to detect low fecal concentration. Another reason for this observation can be the presence of false positives that limit the ability to discriminate fecal spots.

Ratio images of two wavelengths have been previously used to create images that improve fecal contamination detection on produce. Ratio images of red band fluorescence images to blue or green band fluorescence images have been shown to reduce the variation due to colorations and enhance the contrast between fecal spots and fruit surface (Kim *et al*, 2001, 2002; 2003, Vargas *et al*, 2005). Two-band ratio images using combinations of 553 nm, 597 nm, 679 nm, and 730 nm were tested. Of these combinations, the ratio image of 679/730 nm shown in Figure 16a was found to improve the detection of fecal-contaminated spots compared to single-waveband detection rates. This ratio image produced significant intensity difference among the leaves, fruit surface and the treated spots; therefore, this ratio can be used to create a mask to remove the leaves. Figure 16b was constructed by masking out the leaves and as much of the strawberries surface with out affecting the fecal spots in band ratio 679/730 nm image. The range values to construct the mask were defined by inspecting the leaves intensity response and setting mask range with minimum and maximum values between 0 and 200 RFI respectively. The image effectively removes the leaves from the strawberries; nonetheless, natural intensity variations within the fruit become more apparent and hindered the classification process. These false positives are indicated with arrows in the image. In some cases, the 1:100 concentrations can be difficult to identify. In comparison to the single wavelength images, the percentage detection in Figure 21 shows a higher detection rate at low concentration (1:100). This is not the case for high fecal matter concentrations (1:2 and 1:25), the percentage detection is only 69% for both of these concentrations. In general, because of the heterogeneous fluorescence responses associated with the

various components present in the fruit, the two wavelength ratio approach did not provide a mean to emphasize differences between the strawberry surface, false positive and the treated spots. This method may be appropriate to mask the leaves from the strawberries surface in the image. In previous studies PC images have exhibited unique features that provide basis for fecal spot and color variation of the fruit detection (see below).



(a) (b)

Figure 16. Fluorescence ratio images of treated strawberries. (a) 679/730 nm, (b) Binary image of masked image of 679/730 nm using the range values.

Figure 17*a* shows the first four Principal Component (PC-1 to PC-4) images which accounted for 98.88% variability; higher PC images did not illustrate useful information for fecal spots detection. The PC-2 image provided opposite contrast between the leaves and fecal spots; leaves appear dark and treated spots bright. This observation suggests that the image may provide the means for discrimination of the two classes and classify the fecal spots. Both PC-2 and PC-3 showed, near the leaves in the second strawberry in the upper row, a section with brightness intensity similar to the fecal spots.

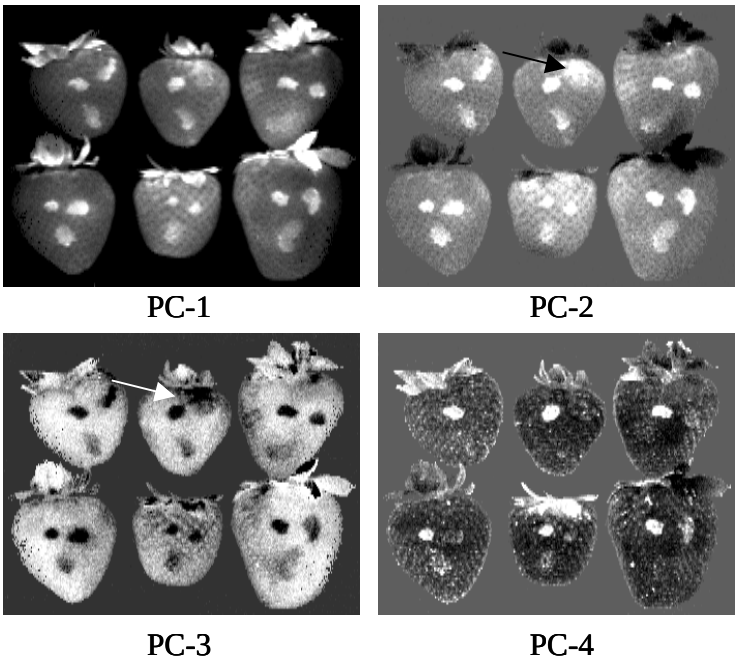


Figure 17. First through fourth Principal Component (PC) images for treated strawberries. The arrow in PC-2 points at a potential false positive.

Figure 18 illustrates the masked PC-2 image. This mask was built by inspecting the leaves and strawberries surface score values of PC-2 and setting mask range with observed minimum and maximum values. The score values for the leaves range were between 0 and –6000. In this image the leaves and strawberry surface share the same dark intensity. However, the method failed to completely eliminate the false positives.

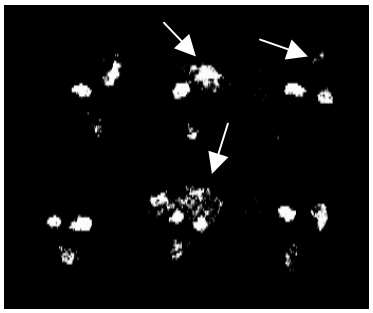
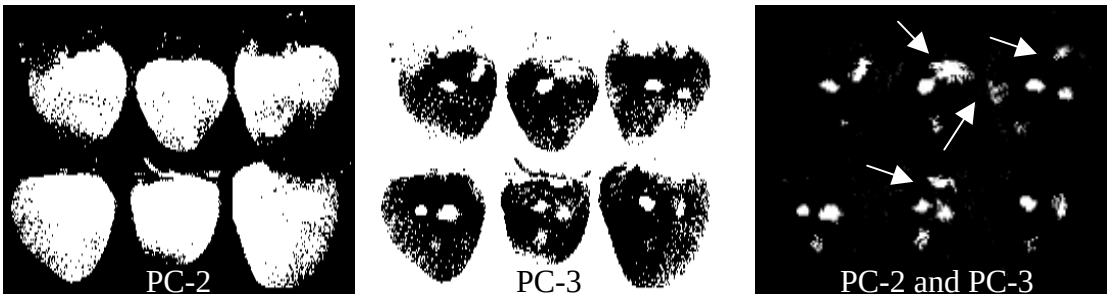


Figure 18. Simple threshold image of PC-2. Arrows indicate false positive.

Figure 19*a*, and Figure 19*b* illustrate the masks created from PC-2 and PC-3 to use for removing the leaves and strawberry surface, respectively. The mask score values ranges to remove the leaves from PC-2 were 5545-8610 and for PC-3 were 6780-3006. The combined result of the masking is shown in Figure 19c. In this image more false positives were removed in comparison to the simple threshold image of PC-2 (Figure 19); nevertheless, false positives are still present and the detection of low fecal concentrations (1:100) was compromised. The concentration detection rates (Figure 21) for this method are more promising than for single band and ratio images. All three concentrations present high detection rates, ranging from 70% for 1:100 to 82% for 1:2. The graph shows an increase of 24% detection for 1:100 concentrations and the highest detection rate previously recorded for the 679/736 nm ratio image at 46%.



(a) (b) (c)

Figure 19. a) Masks created from PC-2 and b) from PC-3. c) Combined effect of PC-2 and PC-3 masks. Arrows point at false positives.

Figure 20 shows the weighing coefficients (eigenvectors) for PC-2 and PC-3 of the treated strawberries. It has been previously observed that the dominant spectral regions (positive or negative peaks) observed in the weighing coefficients plot

approximately coincide with the fluorescence emission maxima in Figure 13 (Kim *et al*, 2003).

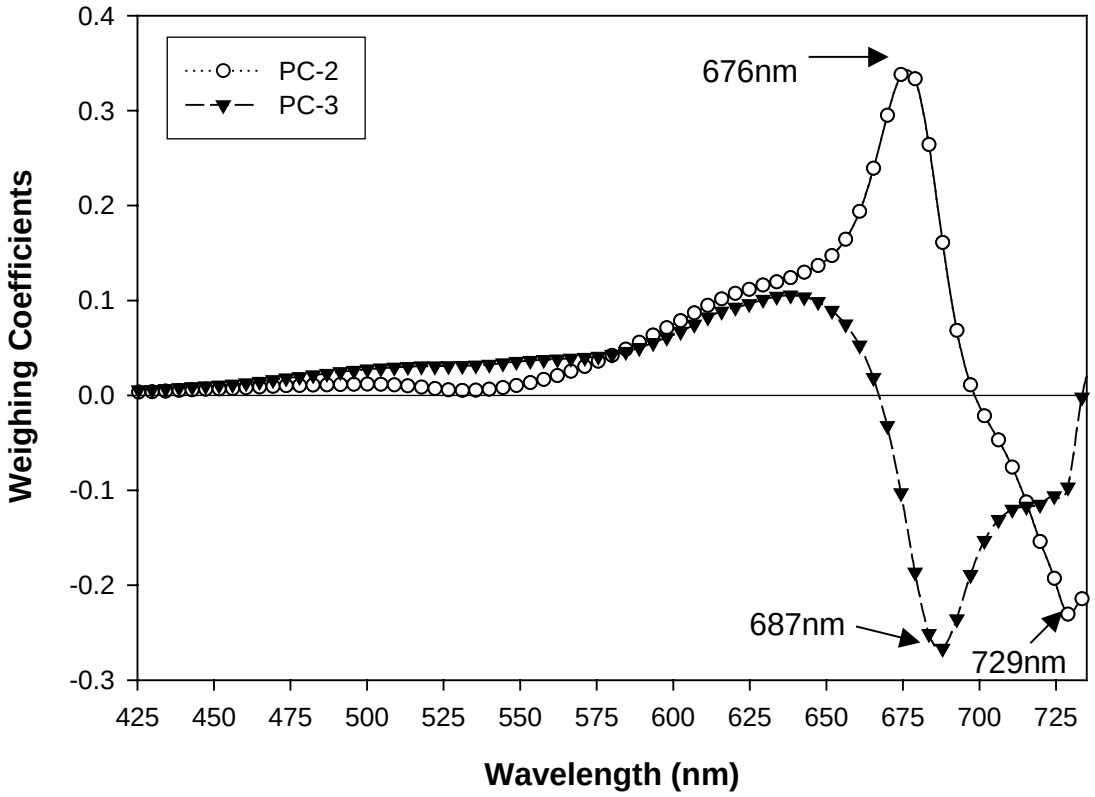


Figure 20. Spectral weighing coefficients of PC-2 and PC-3 obtained by subjecting the entire data set (79 channels) of the fecal contaminated strawberries to PCA. Dominant spectral bands are indicated in the image.

Two major inflection points at 676 nm and 730 nm closely coincided with the emission maxima observed in the fluorescence spectra (Figure 13). Similarly, a negative inflection in PC-3 at 687 nm coincided with one of the leaves emission maxima (Figure 13) at 688 nm. In contrast with previous hyperspectral imaging studies on apples (Kim *et al*, 2002) and cantaloupes (Vargas *et al*, 2005), the detection rates of the feces spots on strawberries appeared to have relatively lower

detection rates throughout all the methods (e.g., single band, band ratio, PC images) under investigation. Color variation of strawberries with abundant moisture underneath of the fruit surfaces may have adversely contributed to the fluorescence responses and thus hindered fecal spot classification. Additionally, the excitation illumination may not have been the best for imaging this produce. Optimal excitation could potentially improve contrast for fecal matter detection on strawberries.

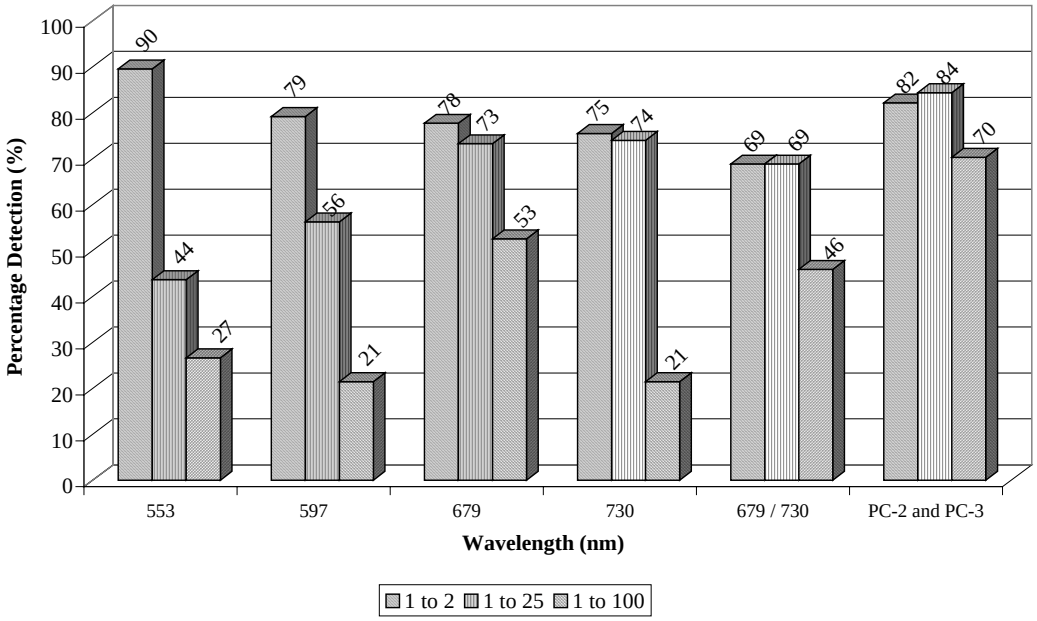


Figure 21. Fecal contamination detection rates for 10-μL treated spots on single wavelength, ratio wavelength and combined PC-2 and PC-3 image (n=135).

2.4. Conclusion

In this study, hyperspectral images of strawberries artificially contaminated with bovine fecal matter to different concentrations were analyzed. Although single band images show high detection rates for high fecal concentration spots, only a fraction of the low fecal concentration spots could be detected. Major impediments in

the classification of the fecal spots were the presence of false positives and similar contrast between the fecal spots and the leafy calyx of the fruit. Two-band ratio images increased the detection rate for low fecal concentration spots, but the results were not as satisfactory as expected, from similar studies previously conducted with cantaloupes, indicating the lack of sensitivity of the method for strawberries. However, the PCA images increased the detection rates for all concentrations and decreased the presence of false positives; these results indicated the potential of the method for fecal spot classification. The analysis of both the representative fluorescence spectra and the weighing coefficients plot identified the same optimal wavebands to use to classify features of interest in the image. The methodology developed in this study can serve as a preamble for the use of spectral analysis of strawberries; nevertheless, further research is required.

Chapter 3: Conclusion

This thesis contributes to research in the area of spectral sensing and its application to food safety and quality control using imaging systems. The investigation results can be considered in the development of multispectral fluorescence imaging for on-line use in cantaloupe and strawberry packing plants.

The main question of this study was whether it was possible to classify fecal matter at different concentrations on the surface of cantaloupes and strawberries using hyperspectral imaging, given that the other chemical and physical aspects of the fruit surface may present similar spectral characteristics. Research results reported in two different articles for both types of produce indicate that it is feasible to classify fecal matter on the surface of cantaloupes and strawberries. The most important conclusions of this study are as follows:

This study showed that the greatest contrast between fecal contaminated and uncontaminated surface areas on cantaloupes was attained at 675 nm. Two-band ratio images at 595/655, 655/520, and 555/655 nm enhanced the contrast between the fecal contaminated spots and untreated cantaloupe surfaces. These images provided relatively high detection rates across all feces dilutions concentrations. PCA minimized false positives in the images by exhibiting differential responses between fecal spots and false positives. This observation was noted in the PC-2 and PC-5

images. The combined use of these two images further enhanced the detection of feces spots with minimal false positives. Based on the PC weighing coefficients, the dominant wavelengths were 465, 487, 531, 607, 643, and 688 nm. These wavelengths can be implemented in a multispectral imaging system for on-line applications.

Evaluation of the fluorescence spectra and principal component weighing coefficients of artificially contaminated strawberries identified 553, 597 679 and 736 nm as optimal wavelengths to classify fecal spots. Visual assessment and fluorescence intensity plots indicated that the image at 679 nm exhibited greatest contrast for all fecal spots concentrations detection. The two-band ratio image at 679/730 nm improved the detection rates by reducing the appearance false positive spots and providing opposite contrast between the surface, the fecal spots and the calyx. PCA provided differential responses for the fecal spots and calyx in the PC-2 and PC-3 images. The combined use these two PC images increased the detection rates for all concentrations and decreased the presence of false positives.

Evaluation of the strawberry data shows that the greatest contrast between the strawberry surfaces and the treated areas is in the 679 nm image. The ratio image of 679/730 nm was found to improve the detection of fecal-contaminated spots compared to single-waveband images. However, because of the heterogeneous fluorescence responses associated with the various components present in the fruit, the single wavelength and the two-wavelength ratio approaches did not provide

means to emphasize differences between the strawberry surface, false positives and the treated spots. PCA images improved the detection of the fecal spots by providing opposite contrast between the leaves and fecal spots. The detection rates for this method were more promising than for single band and ratio images. Weighing coefficients (eigenvectors) plot approximately coincide with the fluorescence emission maxima.

In contrast with hyperspectral imaging studies on cantaloupes, the detection rates of the feces spots on strawberries appeared to have relatively lower detection rates throughout all the methods under investigation. Color variation of strawberries with abundant moisture underneath of the fruit surfaces may have adversely contributed to the fluorescence responses and thus hindered fecal spot classification.

To improve detection results, it is recommended that appropriate excitation lamps be used for each type of produce. It is believed that by the using the correct excitation lamps for each commodity, contrast of fecal matter detection can be improved.

Appendices

JFS E: Food Engineering and Physical Properties

Detection of Fecal Contamination on Cantaloupes Using Hyperspectral Fluorescence Imagery

ANGELA M. VARGAS, MOON S. KIM, YANG TAO, ALAN M. LEFCOURT, YUD-REN CHEN, YAGUANG LUO, AND YOONSEOK SONG, AND ROBERT BUCHANAN

ABSTRACT: To determine whether detection of fecal contamination on cantaloupes is possible using fluorescence imaging, hyperspectral images of cantaloupes artificially contaminated with a range of diluted bovine feces were acquired from 425 to 774 nm in responses to ultraviolet-A (320 to 400 nm) excitation. Evaluation of images at emission peak wavelengths indicated that 675 nm exhibited the greatest contrast between feces contaminated and untreated surface areas. Two-band ratios compared with the single-band images enhanced the contrast between the feces contaminated spots and untreated cantaloupe surfaces. The 595/655-nm, 655/520-nm, and 555/655-nm ratio images provided relatively high detection rates ranging from 79% to 96% across all feces dilutions. However, both single band and ratio methods showed a number of false positives caused by such features as scarred tissues on cantaloupes. Principal component analysis (PCA) was performed using the entire hyperspectral images data; 2nd and 5th principal component (PC) image exhibited differential responses between feces spots and false positives. The combined use of the 2 PC images demonstrated the detection of feces spots (for example, minimum level of 16-μg/mL dry fecal matter) with minimal false positives. Based on the PC weighing coefficients, the dominant wavelengths were 465, 487, 531, 607, 643, and 688 nm. This research demonstrated the potential of multispectral-based fluorescence imaging for online applications for detection of fecal contamination on cantaloupes.

Keywords: hyperspectral imaging, fluorescence, cantaloupe, fecal contamination, ratio image

Introduction

Spurred by multiple outbreaks of foodborne illness associated with fresh produce and the identification of fecal contamination as the major source of pathogens responsible for these outbreaks, the United States Dept. of Agriculture (USDA) has instituted a number of programs to attempt to reduce fecal contamination of food products. In addition, the Food and Drug Administration issued the *Guide to Minimize Microbial Food Safety Hazards for Fresh Fruits and Vegetables* (USFDA 1998). Fruits and vegetables can become contaminated with pathogens through contact with soil, animals, or humans during any stage of the food-handling chain, including growing and harvesting operations as well as while in the processing plants (Murdock and Brokaw 1957). In general, foodborne pathogens originate from the intestinal tracts of animals and humans (USFDA 2001), thus making fecal matter a major source of contamination. For example, fecal contamination can result from use of manure-based fertilizers, and unsanitary conditions in processing and packaging plants.

Outbreaks of salmonellosis have been epidemiologically linked to consumption of fresh cantaloupes (USFDA 2001); as a consequence, this produce has been targeted as a potentially hazardous food. Contaminated cantaloupes were found to be responsible for 2 deaths and 18 hospitalizations due to *Salmonella* between 2000 and 2002; subsequent investigations revealed unsanitary growing and packaging conditions (Anderson and others 2002; USFDA 2002). USFDA investigations noted a broad range of potential factors, such as field irrigation with sewage-contaminated water, poor hygienic practices of workers, pests in packing facilities, and inadequate cleaning and sanitizing of equipment (USFDA 2001).

The Instrumentation and Sensing Laboratory of the USDA in Beltsville, Md., has developed a hyperspectral imaging system (HIS). The system provides a noninvasive method to assess hazardous conditions for fresh produce, is capable of acquiring reflectance and fluorescence measurements, and has been successfully used to demonstrate the detection of fecal contamination on apples and other fresh produce (Kim and others 2001). Hyperspectral imaging systems can simultaneously collect spectral data for hundreds of narrow contiguous wavebands over regions of the electromagnetic spectrum at every spatial pixel in an image. The resulting spectra can serve as fingerprints or signatures for target identification. For food safety applications, a central goal of using hyperspectral imaging data is to find several spectral bands that can be implemented to the development of multispectral inspection systems for online applications at processing plants (Chen and others 2002).

Many image-processing techniques have been developed to identify targets, increase image interpretation reliability, and improve classification. Image fusion is one such technique, defined by Van Genderen and Pohl (1994) as "the combination of 2 or more images to form a new image by using a certain algorithm." Applicable to various types of data sets, its main purposes for digital imagery are to sharpen images, improve geometric correlation, and enhance certain features not readily apparent in single images. Image fusion includes mathematical combinations of spectral images in arithmetic operations and techniques such as principal component analysis (PCA).

MS 20050015 Submitted 1/7/05, Revised 3/2/05, Accepted 6/15/05. Authors Vargas and Tao are with Biological Resources Engineering, Univ. of Maryland, College Park, Md. Authors Kim, Lefcourt, and Chen are with USDA-ARS, Instrumentation and Sensing Laboratory, Beltsville, MD 20705. Author Luo is with USDA-ARS, Food Quality and Safety Laboratory, Beltsville, Md. Author Song is with FDA/NCFST, Div. of Food Processing and Packaging, Summit-Argo, Ill. Author Buchanan is with FDA/CFSAN, College Park, Md. Direct inquiries to author Kim (E-mail: kimm@ba.ars.usda.gov).

For this investigation, hyperspectral fluorescence images were collected from cantaloupes artificially contaminated with bovine feces at varying concentrations. The objective of the present study was to identify a few wavelength bands that could be used to detect fecal contamination on cantaloupes. To enhance the detection process, the selected wavelength bands were subjected to 2-band ratio permutations and an unsupervised classification method was used on the resultant images. In addition, the entire hyperspectral images were subjected to PCA, and principal component (PC) images were evaluated for fecal contamination detection and determination of multispectral bands.

Materials and Methods

Sample preparation

Forty Western Shipper cantaloupes were purchased from a local supermarket. To increase the number of samples and prevent sample movement during HIS scanning, cantaloupes were cut in half, and the 80 halves placed cut-side down in batches of 8 on black trays. Hyperspectral images were 1st collected for all 80 samples before any treatment. Fresh feces from the USDA dairy farm in Beltsville, Md. were diluted by weight to 1:10, 1:50, 1:100, 1:300, and 1:500 with deionized water. Dry fecal matter concentration for 1:100 dilution was 16 μg/mL as determined by drying samples to constant weight in a 90 °C oven. Using a variable pipette, a matrix of dilutions was applied to each sample at 40, 30, 20, and 10 μL (Figure 1). Samples were allowed to air-dry with the aid of a fan for several hours until fecal spots were completely dried, and then scanned again. It was observed that areas contaminated with the 1:10 dilution were usually visible to the human eye. As dilutions increased, spots became transparent and less visible. At 1:100 dilutions, contaminated spots became difficult to visually identify. Digital color photos of samples were acquired prior to taking hyperspectral images.

Hyperspectral fluorescence images acquisition

A hyperspectral imaging system incorporating a line scan spectrograph with a spectral resolution of approximately 7-nm full width at half maximum (FWHM) was used to acquire images (Kim and others 2001). A motorized table was used to move samples at 1-mm intervals, and the zoom lens was adjusted to yield a pixel resolution of 1 mm². The effective spectral range for fluorescence imaging was from 425 nm to 774 nm, with about 4.5-nm intervals, which was captured using 79 spectral channels. Two continuous wave UV-A (360 nm) fluorescent lamp assemblies provided the samples with a near-uniform illumination. The system was operated under dark conditions. A more detailed description of the system operation is given by Kim and others (2004), and calibration is given by Kim and others (2001). Each batch of 8 cantaloupes required approximately 5 min to scan. The line scan data was saved in 16-bit binary files and processed later to create hyperspectral image cubes containing spatial and spectral data. Software developed in Visual Basic version 6 (Microsoft, Seattle, Wash., U.S.A.) was used to acquire data and to perform further processing. Additional data analyses were done using ENVI version 3.2 (Research System Inc., Boulder, Colo., U.S.A.).

Data processing and analyses

The spectral dimension was smoothened using a 3-point running average before image visualization or analyses. Masks to eliminate background were created manually using threshold values determined by evaluating background pixel intensities. The largest possible rectangular region of interest (ROI) within each fecal spot was used to extract representative spectral data. Images taken before treating the samples were used to create spectra of uncontaminated surfaces. ROI for untreated cantaloupes encompassed the largest rectangular area within the sample. The averaged ROI fluorescence spectra were used to characterize fluorescence emission of feces and untreated cantaloupe surfaces. Images at individual wavebands where spectra showed emission peaks were visually evaluated for efficacy of fecal contamination detection. Fecal spots for all dilutions and volumes were visually tallied from the selected single-band images.

All possible ratio permutations of emission peak and valley wavelengths were generated. These ratio images were visually assessed for fecal contamination detection and the best ratio images were selected. Subsequently, selected ratio images were subjected to an unsupervised classification method, ISODATA (Iterative Self-Organizing Data Analysis Technique), which clusters pixels into a desired number of classes without any a priori information about the classes (Ball and Hall 1965). Pixels with similar intensities are grouped into distinct classes by calculating class means evenly distributed in the data space; the remaining pixels are then iteratively clustered using minimum distance techniques until the number of pixels in each class change by less than the selected pixels change threshold. The sum of the square error (SSE), commonly used as clustering principle, is expressed as:

$$SSE = \sum_{C_i} \sum_{x \in C_i} (x - m_i)^t (x - m_i) = \sum_{C_i} \sum_{x \in C_i} \|x - m_i\|^2$$

where m_i is the mean of the ith cluster, $\chi \epsilon C_i$ is a pattern assigned to that cluster and the outer sum, and $\sum_{C_i}$ is the sum over all the clusters (Richards 1986). Only 1 iteration was used to recalculate means, and the threshold parameter was set at 5%. The number of classes was selected based on the number of possible features found on the surface of cantaloupes, such as rind and vein tracts, bruises, cuts, variations in coloration, and fecal spots. For this study, classes ranged between 5 and 10, with no less than 3 pixels per class. Subsequently, several classes in the resulting images were merged to produce a binary image of feces contaminated regions and background.

In addition, the entire image data set (79 channels) was subjected to PCA. In the process of creating the PCA images, a correlation matrix of the image is calculated. This correlation matrix is a diagonal matrix, which is then used to compute the eigenvalues. The eigenvalues are equivalent to the variance of each PC image. The

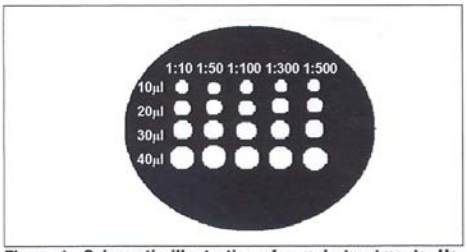


Figure 1—Schematic illustration of sample treatments. Using a variable pipette, bovine feces at 1:500, 1:300, 1:100, 1:50, and 1:10 dilutions were applied to the cantaloupe halves in volumes of 10, 20, 30, and 40 μL. Fecal matter content for 1:100 dilution was 16 μg/mL as determined by drying sample to constant weight in a 90 °C oven.

resulting matrix A is diagonalized by an orthogonal eigenvectors matrix (E), and D is a diagonal matrix (Lay 2002).

$$A = EDE^t$$

As explained by Pohl and Van Genderen (1998), PCA transforms the original data set into a set of new uncorrelated linear combinations of the original variables. In this manner, PCA restructures the data so that most of the variance in the original data is accounted for in a reduced number of variables. The images are then ordered in terms of variance sizes, where the 1st PC accounts for the largest variance.

Results and Discussion

Spectral responses

Figure 2 shows the averaged fluorescence spectra for areas treated with 40 μL of fecal contamination at 1:10, 1:50, 1:100, and 1:300 dilutions, and for untreated sample areas. For the feces treated spots, the average size of the rectangular ROI was 52 pixels (with a minimum of 20 and maximum of 99). For untreated spots, the average size of the rectangular ROI was 13845 pixels (with a minimum of 11880 and maximum of 15720). Fecal spots for 1:500 dilutions were difficult to identify, suggesting that they were similar to uncontaminated areas. For this reason, the averaged spectrum for dilutions of 1:500 was not shown in (Figure 2).

The spectra show fluorescence emission peaks in the green region at 520, 555, and 595 nm, and an additional peak for the feces treated spots in the red at 675 nm. At high feces concentrations, 1:10 and 1:50 dilution spots, there was a distinct response in the green region due to the presence of feces in that relative fluorescence intensities were lower compared with the 1:100, 1:300, and 1:500 dilution spots. This observation was the results of reabsorption characteristics of animal fecal matter existing in relatively high concentrations (Kim and others 2004). The application of the 1:100 and 1:300 dilutions resulted in transparent feces spots, and the fluorescence responses in the green bands resembled those of untreated areas.

Fluorescence emissions in the 650 to 750 nm region from intact green plant materials are due to membrane-bound chlorophyll *a* with emission maxima at 685 nm and 730 nm (Papageorgiou 1975). The averaged spectrum for untreated cantaloupe surfaces exhibited very low chlorophyll *a* fluorescence emission, indicating a well-ripened state of the samples. However, animal fecal matter showed blue-shifted emission peak at 675 nm, which emanated from chlorophyll *a* and its by-products such as pheophobide *a* (Kim and others 2003). Multiple plant constituents, including phenolic compounds and riboflavin, are responsible for the fluorescence emission in the blue-green region of the spectrum (Chappelle and others 1991; Kim and others 2001). Other factors that may affect fluorescence emissions include changes in accumulation of the anthocyanin pigments in association with fruit ripeness (Abbott and others 1997).

Figure 3 shows representative gray-scale fluorescence images of cantaloupes contaminated with feces at 520, 555, 595, and 675 nm. These wavelengths correspond to fluorescence emission maxima observed in the representative fluorescence spectra (Figure 2). In general, the images illustrated progressive decreases in intensities from the center portions toward the edges due to the hemispherical shapes of the cantaloupe halves. In addition, intensity variations in localized regions (asides from the feces treated spots) were noted, suggesting heterogeneous nature of the cantaloupe surfaces in fluorescence responses. In the green region bands at 520, 555, and 595 nm, 1:10 and 1:50 dilution spots were shown darker than surrounding cantaloupe surfaces. For 1:300 and 1:500 dilution spots, the visual identification was difficult; as these dilutions created transparent feces contaminated spots, the fluorescence responses started to blend with those of cantaloupe surfaces. Fluorescence responses at 675 nm were markedly different than those of the green bands in that feces contaminated spots were brighter than cantaloupe surfaces. Similar responses were observed on apples contaminated with bovine feces and this observation was attributed to additive effects of fluorescent animal feces and apples (Kim and others 2003).

Detection results based on the visual assessment of the 520-nm, 555-nm, 595-nm, and 675-nm bands for 40-μL spots are shown in Figure 4. Results for other feces volumes were similar and omitted for brevity. All the wavelength images showed evidence (95% to 100% detection rates) of the 1:10 feces dilution spots. The 675-nm band demonstrated the best potential for detection of feces across the range of volumes and concentrations tested (95% to 73% detec-

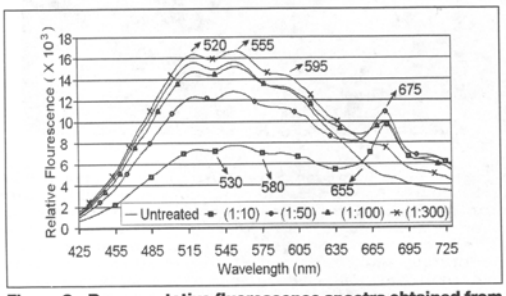


Figure 2—Representative fluorescence spectra obtained from the region of interests (ROI). Note that spectrum for 1:500 dilution spot was not included because the spectral characteristics were similar to those of the cantaloupe surfaces.

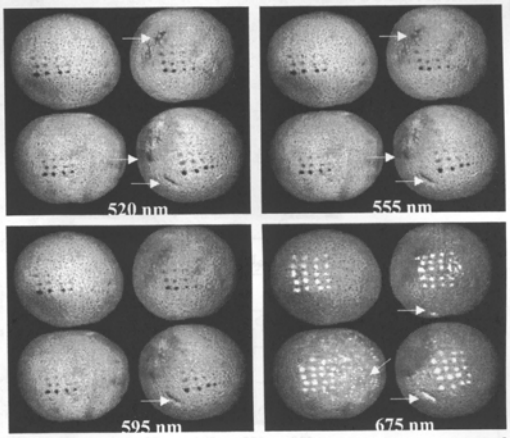


Figure 3—Fluorescence images of bovine-feces treated cantaloupes acquired using the Instrumentation and Sensing Laboratory (ISL) hyperspectral imaging system. The wavelengths correspond to fluorescence emission maxima observed in the representative spectra. Arrows indicate some potential false positives in the images. Cantaloupe in the bottom right corner shows a physical damage (scarred tissues).

URLs and E-mail addresses are active links at www.ift.org

tion rates for 1:10 to 1:500 dilution spots, respectively); although 100% of the 1:10 dilution spots could be detected at the green bands, detection rates for 1:300 and 1:500 dilutions were zero. Based on the single-band imaging, most of the cantaloupes exhibited a number of false positives (Figure 3), which could be attributed to physical damage such as bruises or cuts (scarred tissues). False positives in the red (650 to 700 nm) region of the spectrum may also emanate from netted rind and vein tracts that contained chlorophyll *a*. For the green band images, the relatively lower fluorescence intensities of the edge portions of cantaloupes were also potential false positives.

Ratio images

Although single-waveband images showed potential for detection of fecal matters on cantaloupes, mathematical combinations of images acquired at different wavelengths can enhance fecal contamination detection and reduce false positives. Addition and multiplication of images have been shown to enhance contrast, whereas difference or ratio images are suitable for detecting changes (Pohl and Van Genderen 1998). Kim and others (2002) effectively demonstrated the use of ratio methods to isolated fecal contaminated spots from apples of different varieties. For this investigation, the most promising results were also obtained with 2-band ratio images. Thus, discussion is limited to those ratios that resulted in high detection rates for brevity.

Representative gray-scale, 2-band ratio images (595/655, 655/520, and 555/655 nm), and ratio images subjected to the ISODATA method are shown in Figure 5a and 5b, respectively. These ratios compared with the single-band images enhanced the contrast between the feces treated spots and cantaloupe surfaces and produced more uniform responses across the cantaloupe surfaces. Detection results for 40-μL spots (n = 80 per dilution) for 595/655-nm, 655/520-nm, 555/655-nm, and 675/555-nm ratio images are shown in Figure 6. The 595/655-nm ratio image followed by 655/520 nm, 555/655 nm, and 675/555 nm achieved relatively high detection rates across all feces concentrations. Note that ratio images shown in Figure 5 excluded the use of the best single band at 675 nm in the ratio combinations.

For reducing false positives, the 595/655-nm and 555/655-nm ratio images provided the most satisfactory results (Figure 5b) with detection rates ranging 80% to 95% and 79% to 91%, respectively, for 1:500 to 1:10 feces dilution spots. With the use of the ISODATA method, removal of clusters of 3 pixels or less resulted in the exclusions of some false positives. However, false positives were still apparent; the most prominent false positives were due to the scarred tissues as indicated on the lower-right cantaloupe in Figure 5.

Principal component analysis

Figure 7 illustrates representative 1st PC (PC-1) to 6th PC (PC-6) images obtained from the PCA of the entire hyperspectral fluorescence image data. PC-1 through PC-6 accounted for 99.94% of the data variability and images beyond PC-6 contained no useful attributes for detection of feces treated spots. The PC-1 image reflects a weighted sum of all the spectral bands and showed features causing the largest variations of the data; the intensity decreases from the center portions to the edges of the cantaloupe surfaces. Subsequent PC images depicted other features affecting variations in spectral responses. For instance, PC-3, PC-4, and PC-6 images showed responses that may be attributes of color variations or sides in contact with the grounds (ground spots).

PC-2 and PC-5 images exhibited the evidence of the feces-treated spots as darker and lighter spots, respectively, in contrast to the cantaloupe surfaces. In addition, PC-5 showed the scarred tissues (on the lower-right cantaloupe) that were false positives in the single-band and ratio images as non–false positives. PC-2 and PC-5 images were subjected to a simple thresholding method (Kim and others 2004) to create binary images for fecal contamination spots (Figure 8a and 8b). These images exhibited false positives. However, the PC-

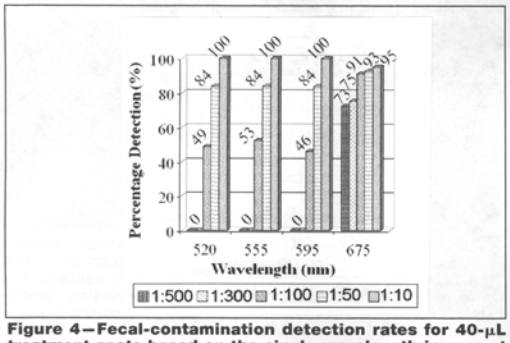


Figure 4—Fecal-contamination detection rates for 40-μL treatment spots based on the single-wavelength images at 520 nm, 555 nm, 595, nm, and 675 nm (n = 80 per dilution)

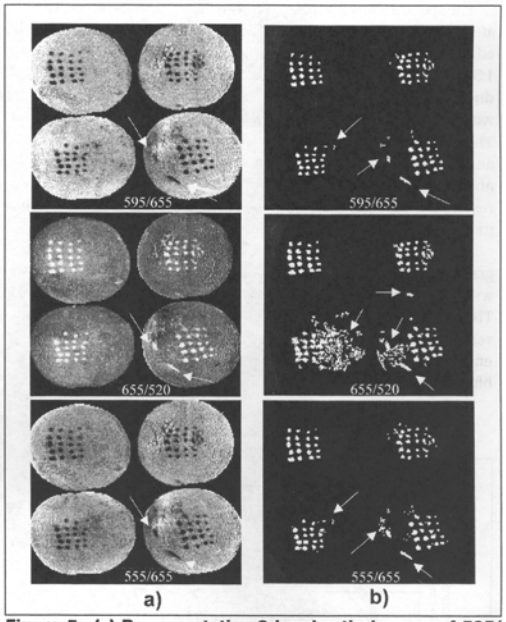


Figure 5—(a) Representative 2-band ratio images of 595/655, 655/520, and 555/655 nm. Out of all possible 2-band ratio permutations of peak and valley wavelengths observed in the fluorescence spectra, these ratios provided the best contrast between the feces-treated spots and cantaloupe surfaces and produced more uniform responses across the cantaloupe surfaces. (b) Binary images for feces-contamination spots obtained by subjecting ratio images to the ISODATA method. Arrows indicate some false positives. 595/655-nm and 555/655-nm ratios were the most effective in reducing false positives.

2 false positives did not coincide with those of PC-5. Figure 8c shows the spots (pixels) where the PC-2 and PC-5 binary images overlapped. Although the combined use of PC-2 and PC-5 compromised the detection of some 1:300 and 1:500 spots, the resultant image in Figure 8c illustrated the 1:10, 1:50, and 1:100 spots (a minimum of 16-μg/mL dry fecal matter) with minimal false positives.

The PCA was performed using the data in entire spectral regions. However, individual PC images could be approximated by the use of few significant wavelengths (Kim and others 2002). Figure 9 shows weighing coefficients (eigenvetors) for the PC-2 and PC-5 images. The peaks and valleys indicated the dominant wavelengths, 487 and 607 nm for PC-2, and 465, 531, 643, and 688 nm for PC-5. Weighted sums of the original images at the dominant wavelengths can create near-identical images comparable to the PC-2 and PC-5 images, respectively. Above wavelengths can be implemented to a multispectral imaging system for online applications. Readily available common aperture-based, multispectral adaptors can be used to simultaneously capture images up to 8 spectral bands (Chen and other 2002). Further research is needed to incorporate samples with various stages of maturity and to develop automated processing and detection methods for online applications.

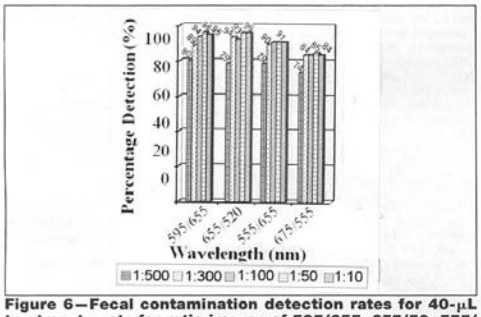


Figure 6—Fecal contamination detection rates for 40-μL treatment spots for ratio images of 595/655, 655/50, 555/655, and 675/555 nm subjected to the ISODATA (Iterative Self-Organizing Data Analysis Technique) classification method (n = 80 per dilution). Note that these ratios represent the 2-band ratio combinations with the highest detection rates.

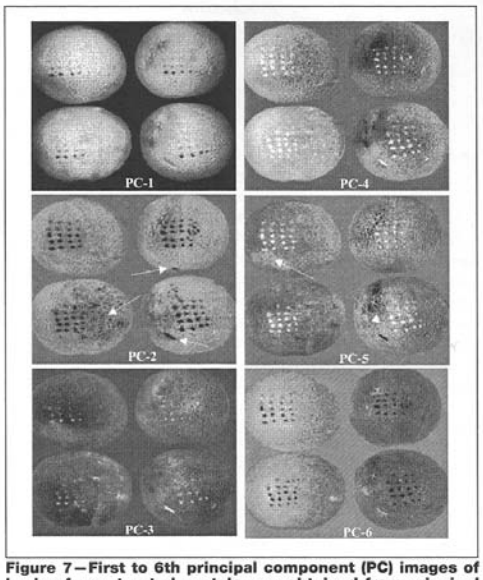


Figure 7—First to 6th principal component (PC) images of bovine feces-treated cantaloupes obtained from principal component analysis (PCA) of the entire hyperspectral fluorescence image data (79 spectral channels). Arrows indicate some of false positives.

Conclusions

In this investigation, hyperspectral fluorescence images were evaluated for detection of fecal contamination on cantaloupes. Single waveband images showed natural variations in fluorescence responses of the cantaloupe surfaces, and were not suitable for detection of fecal contamination on cantaloupes due to the presence of false positives. Images constructed using 2-band ratios enhanced the contrast between the fecal-contaminated spots and cantaloupe surfaces. However, ratio images also exhibited false positives emanating from features found on cantaloupes such as scarred tissues. The PCA of hyperspectral images demonstrated the potential detection of feces-contaminated spots (a minimum of 16-μg/mL dry fecal matter) on cantaloupes with minimal false positives. Examination of PC weighing coefficients defined several dominant

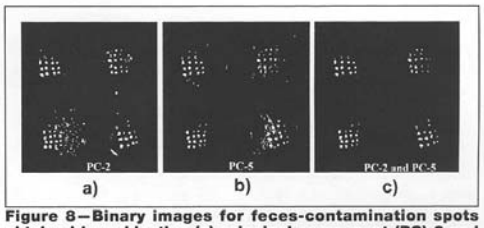


Figure 8—Binary images for feces-contamination spots obtained by subjecting (a) principal component (PC)-2 and (b) PC-5 images to a simple thresholding method (Kim and others 2004). (c) Binary image shows where both PC-2 and PC-5 binary results overlapped. As the resulting binary image illustrates, pixel locations of the PC-2 false positives did not coincide with those of PC-5.

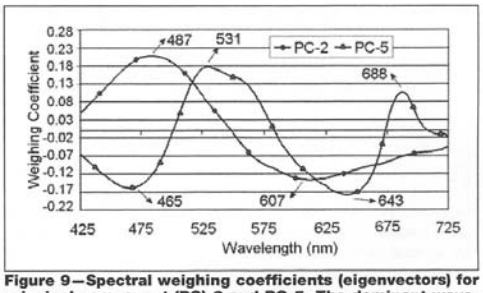


Figure 9—Spectral weighing coefficients (eigenvectors) for principal component (PC)-2 and PC-5. The dominant wavelengths are indicated on the graph.

wavelengths that can be implemented to a multispectral imaging system for online applications.

Acknowledgments

The authors would like to thank the Instrumentation and Sensing Laboratory (ISL), United States Dept. of Agriculture (USDA), the Food and Drug Administration (USFDA), The Joint Inst. of Food Safety and Nutrition (JIFSAN), and the Dept. of Biological Resources Engineering (BRE) at the Univ. of Maryland at College Park (UMCP) for supporting this work. The authors would like to thank Ms. Diane Chan and Mr. John Kelly of ISL for their assistance in sample collection.

References

Abbott JA, Lu R, Upchurch BL, Stroshine RL. 1997. Technologist for non-destructive quality evaluation of fruit and vegetables. Hortic Rev 20:1–120.

Anderson J, Stenzel S, Smith K, Labus B, Rowley P, Shoenfeld S, Gaul L, Ellis A, Fyfe M, Bangura H, Varma J, Painter J. 2002. Multistate outbreaks of *salmonella* serotype poona infections associated with eating cantaloupe from Mexico—United States and Canada, 2000-2002. Morb Mortal Wkly Rep 51(46):1044–7.

Ball GH, Hall DJ. 1965. ISODATA, a novel method of data analysis and pattern classification. Stanford: Stanford Research Inst. Technical Report.

Chappelle EW, McMurtrey JE, Kim MS. 1991. Identification of the pigment responsible for the blue fluorescence band in laser induced fluorescence spectra of green plants, and the potential use of this band in remotely estimating rates of photosynthesis. Remote Sens Environ 36:213–8.

Chen YR, Chao K, Kim MS. 2002. Machine vision technology for agricultural applications. Comp Electr Agric 36:173–91.

Duda RO, Hart PE, Stork DG. 2001. Pattern classification. 2nd ed. New York: Wiley-Interscience. p 114–7.

Kim MS, Chen YR, Mehl PM. 2001. Hyperspectral reflectance and fluorescence imaging system and safety. Trans ASAE 44(3):721–9.

Kim MS, Lefcourt AM, Chen YR. 2003. Optical fluorescence excitation and emission bands for detection of fecal contamination. J Food Protect 66(7):1198–207.

Kim MS, Lefcourt AM, Chen YR, Kang S. 2004. Hyperspectral and multispectral laser induced techniques for food safety inspection. Key Eng Mater 270:1055–63.

Kim MS, Lefcourt AM, Chen YR, Kim I, Chan DE, Chao K. 2002. Multispectral detection of fecal contamination on apples based on hyperspectral imagery. Part II. Application of hyperspectral fluorescence imaging. Trans ASAE 45(6):2039–47.

Lay D. 2002. Linear algebra and its applications. New York: Addison-Wesley. p 441–86.

Murdock DI, Brokaw CH. 1957. Some specific source of contamination in processing frozen concentrated orange juice. 1. handling and preparing fruit for extraction. Proc Fla State Hort Soc 70:231–7.

Papageorgiou G. 1975. Chlorophyll fluorescence: an intrinsic probe of photosynthesis. In: Govindjee JA and Ford DC. Bioenergetic of photosynthesis. 1st ed. New York: Academic Press. p 320-71.

Pohl C, Van Genderen JL. 1998. Multisensor image fusion in remote sensing: concepts, methods, and application. J Remote Sens 19(5):823–54.

Richards JA. 1986. Remote sensing digital image analysis: an introduction. BerliN: Springer-Verlag.

[USFDA] U.S. Food and Drug Admin. 1998. Guide to minimize microbial food safety hazards for fresh fruit and vegetables. Washington D.C.: Center for Food Safety and Applied Nutrition. Available from: http://www.cfsan.fda.gov. Accessed Oct 3, 2005.

[USFDA] U.S. Food and Drug Admin. 2001. Analysis and evaluation of prevention control measures for the control and reduction/elimination of microbial hazards on fresh and fresh-cut produce. Washington D.C.: Center for Food Safety and Applied Nutrition. Available from: http://www.cfsan.fda.gov. Accessed Oct 3, 2005.

[USFDA] U.S. Food and Drug Admin. 2002. Import alert IA2201: Detection without physical examination of cantaloupes from Mexico. College Park, Md.: Center for Food Safety and Applied Nutrition. Available from: www.fda.gov. Accessed Oct 3, 2005.

Van Genderen JL, Pohl C. 1994. Image fusion: issues, techniques and applications. Intelligent image fusion. Sept 11, 1994. Proceedings EARSeL Workshop; Strasbourg, France. p 18-26.

Bibliography

Abbott, J. A., R. Lu, B. L. Upchurch, and R. L. Stroshine. 1997. Technologist for non-destructive quality evaluation of fruit and vegetables. *Hort. Reviews* 20:1-120.

AccessScience. 2003. The Online Encyclopedia of Science & Technology. [online available] WWW:http://www.accessscience.com/server-java/Arknoid/science/AS.

Ball, G.H., and D. J. Hall. 1965. ISODATA, a novel method of data analysis and pattern classification. Stanford Research Institute Technical Report. Stanford.

Bertelsen, D. 1995. The U.S. Strawberry Industry. Commercial Agriculture Division, Economic Research Service, U.S. Department of Agriculture. Statistical Bulletin No. 914.

Browne, K. M., J. D. Greeson, and C. Dennis. 1984. The effects of harvest date and CO_2 enriched storage atmospheres on the storage and shelf-life of strawberries. *J. Hortic. Sci.* 59:197-204.

CDC. 2002. Multistate outbreaks of salmonella serotype poona infections associated with eating cantaloupe from Mexico - United States and Canada, 2000-2002. Morbidity and mortality weekly report, Center for Disease Control and Prevention 51(46): 1044-1047.

CDC. 2002. Multistate Outbreaks of Salmonella Serotype Poona Infections Associated with Eating Cantaloupe from Mexico --- United States and Canada, 2000—2002. *Morbidity and Mortality Weekly Report* 51(46):1044-1047.

Chappelle, E. W., F. M. Wood, Jr., J. E. McMurtrey III, and W. W. Newcomb. 1984. Laser-induced fluorescence of green plants. 1: A technique for the remote detection of plant stress and species differentiation. *Applied Optics* 23:134-138.

Chappelle, E. W., M. S. Kim, and J. E. McMurtrey. 1991. Identification of the pigment responsible for the blue fluorescence band in the laser induced fluorescence (LIF) spectra of green plants, and the potential use of this band in

remotely estimating rates of photosynthesis. *Remote Sens. Environ.* 36:213-218.

Chen, P., M. J. McCarthy, and R. Kauten. 1989. NMR for internal quality evaluation of fruits and vegetables. *Transactions of ASAE* 32:1747-1753.

Chen, Y. R., K. Chao, and M. S. Kim. 2002. Machine vision technology for agricultural applications. *Computer and Electronics in Agriculture* 36:173-191.

Corp, L. A., J. E. McMurtrey, E. W. Chappelle, C. S. T. Daughtry, and M. S. Kim. 1997. UV band fluorescence (in vivo) and its implications for the remote assessment of nitrogen supply in vegetation. *Remote Sens. Environ.* 61:110-117.

Duda, R.O, P. E. Hart, and D. G. Stork. 2001. Pattern classification. 2nd ed. New York, NY: Wiley-Interscience. 114-117.

FDA. 1998. Guide to minimize microbial food safety hazards for fresh fruit and vegetables. [online available] WWW:http://www.cfsan.fda.gov.

FDA. 2001. Analysis and evaluation of prevention control measures for the control and reduction/elimination of microbial hazards on fresh and fresh-cut produce. Center for Food Safety and Applied Nutrition. [online available] WWW:http://www.cfsan.fda.gov.

FDA. 2002. Import alert IA2201: Detection without physical examination of cantaloupes from Mexico. [online available] WWW:http://www.fda.gov/.

GAO, 2000. School Meal Program: Few Outbreaks of Foodborne Illness Report. GAP Reference No. GAO/RCED-00-53 School Meal Program. Washington, D.C.: United States General Accounting Office.

Hancock, J. 1999. Strawberries. Wallingford; New York : CABI Pub.

Howard, A. 1987. Elementary Linear Algebra 5e. New York, NY: John Wiley & Sons, Inc.

Hutin, Y. J. F., V. Pool, E. H. Cramer, O. M. Nainan, J. Weth, I. T. Williams, S. T. Goldstein, K. F. Gernsheimer, B. P. Bell, C. N. Shapiro, M. J. Alter, and H. S. Margolis, and the National Hepatitis A Investigation Team. 1999. A

multistate, foodborne outbreak of hepatitis A. *New England Journal of Medicine*. 340:595-602.

Jones, J. K. 1976. Strawberry, Fragaria ananassa (Rosaceae). In: N.W. Simmonds (ed), Evolution of crop plants. London: Longman. pp. 237-242.

Kim, M. S., A. M. Lefcourt, and Y. R. Chen. 2003. Optical fluorescence excitation and emission bands for detection of fecal contamination. *J Food Protection* 66(7): 1198-1207

Kim, M. S., A. M. Lefcourt, K. Chao, Y. R. Chen, and I. Kim. 2002. Multispectral detection of fecal contamination on apples based on hyperspectral imagery—part I: application of visible-near infrared reflectance imaging. *Transactions of the ASAE* 45:2027-2037.

Kim, M. S., A. M. Lefcourt, Y. R. Chen, and S. Kang. 2004. Hyperspectral and multispectral laser induced techniques for food safety inspection. *Key Engineering Materials* 270:1055-1063.

Kim, M. S., A. M. Lefcourt, Y. R. Chen, I. Kim, D. E. Chan, and K. Chao. 2002. Multispectral detection of fecal contamination on apples based on hyperspectral imagery—part II: application of hyperspectral fluorescence imaging. *Transactions of the ASAE* 45(6):2039-2047.

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

Lay, D. 2002. Linear algebra and its applications. New York: Addison-Wesley. 441-486p.

Marks, J. S., J. Schmidt, M. T. Morgan, J. A. Nyenhuis, R. L. Stroshine. 1998. Nuclear magnetic resonance for poultry meat fat analysis and bone chip detection. Industry summary for US Poultry and Egg Association.

Milne, L. and M. Milne. 1975. Living plants of the world. New York, NY: Random House.

Murduck, D. I, and C.H. Brokaw. 1957. Some specific source of contamination in processing frozen concentrated orange juice-1 handeling and preparing fruit for extraction. *Proc. Fla. State Hortic. Soc.* 70:231-237.

Papageorgiou, G. 1975. Chlorophyll fluorescence: an intrinsic probe of photosynthesis. In bioenergetic of photosynthesis. Govindjee, 1st ed. New York, N.Y.: Academic Press.

Pohl, C. and J. L. Van Genderen. 1998. Multisensor image fusion in remote sensing: concepts, methods, and application. *J. Remote Sensing* 19(5): 823-854.

Porter, C. L. 1967. Taxonomy of Flowering Plants. San Francisco, CA: W. H. Freeman and Company.

Richards, J.A, 1986. Remote sensing digital image analysis: an introduction. Berlin, Germany: Springer-Verlag.

Richards, T. and P. Patterson, P. 1999. The economic value of public relations expenditures: food safety and the strawberry case. *Journal of Agricultural and Resource Economics* 24:440-462.

Schatzki, T.F., R. P. Haff, R. Young, I. Can, L. C. Le, and N. Toyofuku. 1997. Defect detection in apples by means of X-ray imaging. *Trans. ASAE* 40:1407-1415.

Stein, W. J., S. G. Beaven, L. E. Hoff, E. M. Winter, A. P. Schaum, and A. D. Stocker. 2002. Anomaly detection from hyperspectral imagry. *IEEE Signal Processing Magazine* 19(1):58-69.

Tao, Y., R. Buchanan, Y. Song, Y. Luo, Y. R. Chen, M. S. Kim. 2002. Safety inspection of fruit and vegetables using optical sensing and imaging techniques. Proposal for the USDA.

Törrönen, R. and K. Määttä. 2002. Bioactive substances and health benefits of strawberries. *Acta Hort.* (ISHS) 567:797-803.

UGA. 2006. Cantaloupe and Specialty Melons. [online available] WWW:http://pubs.caes.uga.edu/caespubs/pubcd/b1179.htm.

Van Genderen, J. L, and C. Pohl. 1994. Image fusion: issues, techniques and applications. Intelligent image fusion, Proceedings EARSeL Workshop, Strasbourg, France. pp 18-26.

Vargas, A. M., M. S. Kim, Y. Tao, A. M. Lefcourt, Y. R. Chen, Y. Luo, and Y. Song. 2005. Detection of fecal contamination on cantaloupes using hyperspectral fluorescence imagery. *Journal of Food Science* 70(8): 471-476.

Wang, L. 2000. Antioxidant activity in fruits and leaves of blackberry, raspberry and strawberry. *Journal of Agriculture Food Chemistry* 48:140-146.

Wedge, D. and M. Kumundi. 2001. Anticarcinogenic activity of strawberry, blueberry and raspberry extracts to breast and cervical cancer cells. *Journal of Medicinal Food* 4:49-51.

Weeks, A. 1996. Fundamentals of electronic image processing. Bellingham, WA: SPIE Optical Engineering Press.

White, P.R. 1927. Studies of the physiological anatomy of the strawberry. *J. Agr. Res.* 35:481-492.

Willoghby, C. T, M. A. Folkman, and M. A. Figueroa. 1996. Application of hyperspectral imaging spectrometer systems to industrial inspection. *Proceedings of SPIE* 2599:264-272.

Winton, A.L. 1902. The anatomy of edible berries. *Conn. Exp. Sta. Rept.* pp. 288-325.