ABSTRACT

Title of Dissertation: DETECTION OF PHYSICAL HAZARDS IN BONELESS POULTRY PRODUCT USING COMBINED X-RAY AND LASER RANGE IMAGING TECHNOLOGIES

Xin Chen, Doctor of Philosophy, 2003

Dissertation directed by: Professor Yang Tao
Department of Biological Recourses Engineering

Detection of bone fragments and other physical contaminations in deboned poultry meat has become increasingly important to ensure food quality and safety. Traditional X-ray imaging detection technologies have significant difficulties detecting contaminations because of the meat tissue thickness variation.

In order to address the thickness variation problem, in this study, a novel vision system with combined X-ray and laser 3D imaging technology has been developed for accurate physical contamination detection. The X-ray part of the combined system captures high resolution X-ray images in real-time, and the laser 3D part provides an accurate thickness profile for each piece of meat. In the combined system, the 3D

thickness information is used to cancel the thickness variation in the X-ray image, thus the process of physical contamination detection is significantly simplified.

The combined vision system is capable of detecting calcified bones (rib bones and pulley bones) at a 95% detection rate, and partially calcified bones (fan bones) at a 90% detection rate.

In order to handle the inspection tasks in real-time, a multithread architecture is used in this vision system. Various threads work simultaneously in the system, synchronized with each other, taking full advantage of system resources. It is shown that real-time capability is achieved due to the multithread framework.

The result of this study has the potential to promote food safety and quality by providing advanced and automated detection techniques to the poultry and food industries.

DETECTION OF PHYSICAL HAZARDS IN BONELESS POULTRY PRODUCT USING COMBINED X-RAY AND LASER RANGE IMAGING TECHNOLOGIES

By

Xin Chen

Dissertation 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
Doctor of Philosophy
2003

Advisory Committee:

Professor Yang Tao, Chair
Professor Lewis E. Carr
Professor Ramalingam Chellappa
Professor Yud-Ren Chen
Professor Arthur T. Johnson
Professor Fredrick W. Wheaton

ACKNOWLEDGEMENTS

First I would like to thank my advisor and committee chairman, Dr. Yang Tao, for his immeasurable guidance, patience, and support throughout the course of my Ph.D. study. I have learned much as a direct result of working with him. In addition, his confidence in my capability has enabled much of my creative work as well as diligent working habits.

I would also like to thank Dr. Lewis E. Carr, Dr. Ramalingam Chellappa, Dr. Yud-Ren Chen, Dr. Arthur T. Johnson, and Dr. Fredrick W. Wheaton for serving on my committee and for their helpful advice on this research.

I would like to thank all the members in our research group, especially Mr. Hansong Jing for his excellent efforts on laser imaging, Mr. Gary Seibel for his indispensable help to my research, and Ms. Abby Vogel for her valuable inputs on this dissertation.

Special thanks to my dear wife Xuemei Cheng, for her enduring love, encouragement, inspiration, patience, and support through all these years.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

BMD	=	bone mineral density
C	=	image contrast
CQI	=	combined quality index
DQE	=	detective quantum efficiency
E	=	X-ray energy
I	=	X-ray image intensity
I_{px}	=	pseudo X-ray image
I_{tc}	=	thickness compensated image
MTF	=	modulation transfer function
NPS	=	noise power spectrum
PLL	=	phase locked loop
PSM	=	pulse sampling module
PTM	=	pulse tracking module
SNR	=	signal to noise ratio
μ	=	X-ray attenuation coefficient
ρ	=	mass density
σ	=	standard deviation
Φ	=	X-ray quanta per unit area

CHAPTER 1

INTRODUCTION

The United States produces billions of pounds of poultry every year, a large portion of which is boneless poultry meat. Physical contamination is one of the major types of contamination compromising food safety. In boneless poultry meat, physical hazards include plastic, metal, glass, bone fragments, etc., which can lead to serious injury if ingested by the consumer. To ensure food quality and food safety, it is necessary for poultry processors to inspect each piece of boneless poultry product and make sure that bone fragments and any unwanted hazardous materials such as metals and plastics do not remain in the product.

In many poultry processing plants, the inspection of boneless meat is still performed manually. The accuracy of manual inspection depends on the sensitivity of workers' fingers. However, the meat products are so cold that the inspectors' fingers become quickly numb and insensitive to possible bone fragments. Additionally, cross-contamination and high labor costs are also problems of manual inspection.

X-ray imaging techniques, a noninvasive inspection method, have been used for years to detect physical contamination in food products. However, the traditional X-ray inspection systems currently being used to detect bone fragments in meat fillets have a high rate of failure (over 30%). As suggested by numerous publications and industrial reports, X-ray inspection systems cannot succeed in detecting bone fragments in poultry unless the challenge of uneven meat thickness is addressed. Generally, the intensity (or grayscale) of an object in an X-ray image reflects the X-ray absorption, which is dictated

by the physical characteristics of the material and its thickness. The intensity of X-ray imagery ***I*** can be described as (Dowsett, et al., 1998):

$$I = I_0 \exp(-\mu d) \tag{1.1}$$

where

I_0 : the incident X-ray intensity,

μ : X-ray attenuation coefficient of the material of interest (assuming homogeneity), and

d: the length of the X-ray pathway through the object.

According to equation (1.1), the thickness variation leads to grayscale variation in the X-ray image, which has been confirmed by experimentation. In the case of bone fragment detection, the X-ray absorption differences are small between normal meat tissue and bone fragments, especially when the bone fragments are thin. This weak intensity difference in the X-ray image makes it very difficult to distinguish bone fragments from meat patterns even using the human eye. Furthermore, typical deboned poultry products, such as chicken breasts, have uneven thickness, which means that the X-ray intensity of a thicker area could be comparable to a bone fragment in a thinner area. The false pattern in the X-ray image caused by uneven thickness could lead to significant classification error.

In order to find and implement an effective solution to address the uneven thickness problem, researchers from academia and industry have tried numerous approaches (Graves and Batchelor, 2003). It is believed that the discovery of a viable solution should start from basic physics and finish by the insight gained through mathematics. Thus in this study, the basic approach is to estimate the thickness

interference by direct, nondestructive physics methods, and then identify the contaminant using computer vision methods.

The methods and results of the combined X-ray and laser range imaging system will be presented in this dissertation. Chapter 2 describes the overview of this study along with background information of the challenges. In chapter 3, X-ray image quality at different energy levels is evaluated, as well as the limitations of solutions based on single energy X-ray and dual energy X-ray technologies. Chapter 4 gives an extensive description of the combined X-ray and laser 3D imaging approach. Experimental results and discussions are provided in chapter 5, followed by chapter 6, the conclusions.

CHAPTER 2

REVIEW OF LITERATURE

2.1 Physical contaminations in deboned poultry

In the poultry processing industry, physical contamination, especially bone fragments in deboned poultry products, is a major concern to food safety. Bone fragments in presumed boneless poultry products, such as chicken fillets, nuggets, etc., could lead to consequences ranging from consumer complaints to health incidents. The USDA listed physical contamination as one of the three major types of food safety threatening contaminations (USDA, 1996) and it is required by USDA regulations that there be no bone fragments or other types of physical contaminations left in deboned products (USDA, 2002).

In practice, the acceptable quality standard for bone fragment occurrences in deboned poultry meat varies from plant to plant. According to a recent study on a typical poultry processing plant (Smith, 2001), there were approximately three bone fragments left in every 10,000 deboned fillets and one bone fragment left in every 10,000 deboned tenders. It was also found that the majority of bone fragments left in deboned chicken meat were clavicles or fan bones. Bone fragments originating from the clavicle tend to be highly calcified, and fan bones tend to be thin and sharp. Both types of bone fragments can cause serious medical problems to consumers, especially to children and the elderly.

A similar study was conducted in the UK, focusing on poultry meat purchased from supermarkets. The destructive method was used to inspect chicken meat purchased from several leading supermarket retailers in the UK. The results showed that the average

number of bone fragments left in the poultry meat was more than 87 per 100 kg (Graves and Batchelor, 2003), which is significantly higher than one would expect. The disparity between the US and UK studies was due to the difference in inspection methods, i.e., the destructive inspection method adopted by the UK study would result in a more accurate estimate of bone fragment occurrences. Despite the different numerical results of the two studies, both findings highlight the necessity of effective bone fragment detection.

According to a recent study commissioned by the Food Safety and Inspection Service, USDA (Morales, 2002), the majority of meat and poultry processing plants rely on manual visual inspection to detect and control physical hazards, including bones and metals, in deboned products. In a typical setup of a poultry deboning line, a crew of 24 to 32 workers sit shoulder to shoulder, working on each passing chicken carcass by cutting, pulling off muscles, and detaching skins. At the end of the line, another crew of several workers double checks each piece of deboned meat, making sure each is free of bone fragments. This re-inspection process is also done by hand, squeezing the chilled chicken meat in an effort to feel the embedded bone fragments.

Certainly, the hand checking approach has several disadvantages. First, lengthy manual inspection is a demanding task for the workers, and inevitable muscle and mental fatigue could lead to health issues. Secondly, it is labor intensive, and the throughput is low compared to automatic approaches. Third, it raises hygiene problems, because in the wet and greasy environment, contamination could spread from one worker to another or from one piece of chicken meat to another. Finally and most notably, the accuracy of bone detection is not ensured, because there is no guarantee that a bone fragment embedded in meat tissue can be felt by hand, especially when workers' fingers have

become numb after touching the chilled poultry meat for hours. It is certainly desirable for a poultry processing plant to design an automatic technique for bone fragment detection, and X-ray imaging seems to be the most viable option.

2.2 X-ray technology and food inspection

2.2.1 Overview of X-ray imaging

X-ray imaging is a well-established technology and has been widely used for numerous non-destructive testing applications, since Dr. Roentgen's discovery of *a new kind of ray* on November 8, 1895 (Selman, 1993).

X-rays are a form of electromagnetic radiation with a wavelength ranging from 10^{-3} nm to 10 nm, with photon energy ranging from 120 eV to 1.2 MeV (Selman, 1993). The penetrating nature of X-rays makes them ideal tools to detect the internal structure of objects based on the interaction between the X-rays and matter (Hale, 1975). Usually the energy range of 40 keV to 120 keV is used for radiography (Selman, 1993). Two processes that lead to X-ray attenuation have been studied, namely photo electronic absorption and Compton scattering (Dowsett, et al., 1998). The attenuation processes are both energy dependent and material dependent. For most materials with small atomic numbers (<20), photo electronic absorption dominates at low energies while Compton scattering dominates at high energies (Graham and Cloke, 2003). Here the threshold between low and high energies could be anywhere from 50 keV to 500 keV, depending on the atomic number of the material of interest.

Typically, X-rays are generated in an X-ray tube by shooting a beam of high-energy electrons from a heated filament to an anode. X-ray radiation is generally classified into two types: Bremsstrahlung radiation and characteristic radiation (Dowsett,

et al., 1998). Bremsstrahlung is also called braking radiation, which is emitted when electrons lose energy in the form of electromagnetic radiation due to a change in velocity. Characteristic radiation is caused when vacancies created in inner electron orbitals are filled by the electrons from outer orbitals, which give off a photon of radiation and make a quantum jump. The spectrum of Bremsstrahlung radiation is continuous, while characteristic radiation forms line spectra, which leads to an overall radiation spectrum shown in Figure 2.1 (Hamamatsu, 2000).

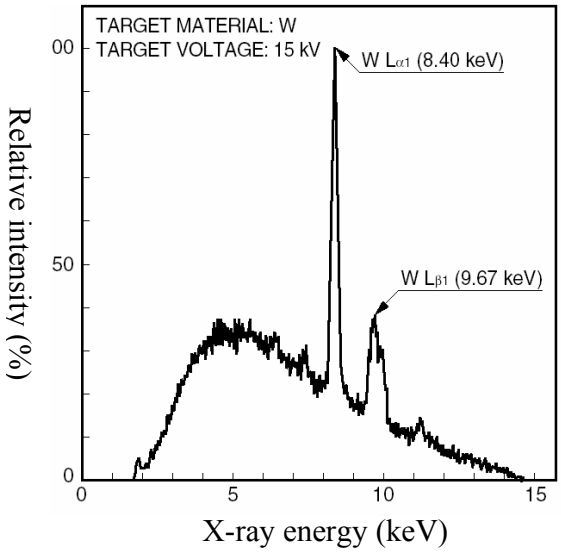


Figure 2.1 Radiation spectrum of an X-ray tube

A typical X-ray imaging system is depicted in Figure 2.2, where an X-ray beam is generated by the X-ray tube, attenuated by the target object, and then intercepted by an X-ray imager. The imager can be a radiographic film (Selman, 1993; Thomas, et al., 1995), a CCD video camera coupled with an image intensifier (Molloi, et al., 1993), a linear digital X-ray detector array (Haff and Schatzki, 1997), or an amorphous silicon flat

panel for X-ray imaging (Fann, et al., 2003). Radiographic films are specially designed to be highly sensitive to X-ray radiation.

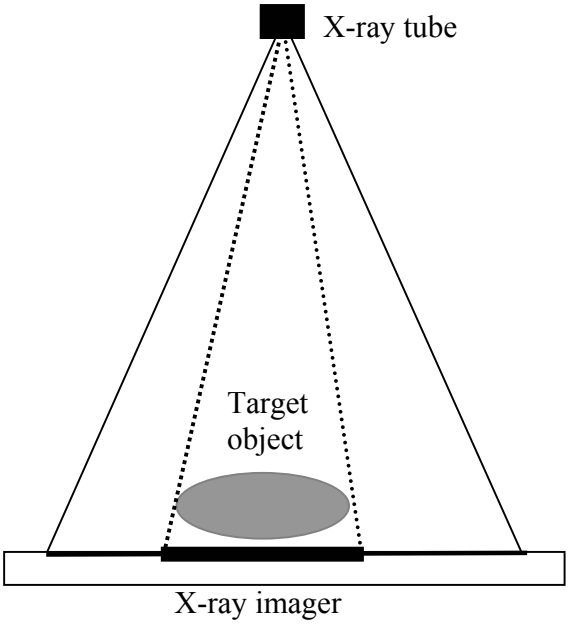


Figure 2.2 Typical setup of an X-ray imaging system

In general, digital imagers are preferred over analogue imagers for many reasons including high spatial resolution, rapid image acquisition, and convenient image storage and manipulation (Dowsett, et al., 1998). With a digital X-ray imager and sophisticated electronic control mechanism, it is possible to calibrate the imaging parameters in real-time, and to take sequential exposures at different X-ray energies quickly. It is reported that a GE digital radiographic system can take two calibrated exposures only 200 ms apart. The major benefit of this approach is improved image quality because both exposures are calibrated and there is reduced displacement between the snapshots. This also makes it more convenient to infer meaningful information by comparing the two images (Dobbins, et al., 2003).

Image quality of the resultant X-ray image can be assessed from different aspects, such as the noise level, spatial resolution, and image contrast. A successful X-ray imaging system should deliver images with low noise level, high spatial resolution, and high contrast of the object. It is believed that X-ray image noise is caused by insufficient photon flux from the X-ray source (Graham and Cloke, 2003) or thermal noise from the X-ray imager (Speller, et al., 2001). Spatial resolution is largely limited by the pixel pitch of the X-ray imager, while a large focal spot of the X-ray source could also affect the image resolution. Image contrast can often be improved by reducing X-ray photon energy, but the noise level could also get worse because fewer photons are able to penetrate the object at a lower X-ray energy.

It is well know that X-ray radiation could pose health threat to human body, thus it is extremely important to carefully enforce certain protection procedures for X-ray imaging personals (Selman, 1993).

2.2.2 Typical applications of X-ray imaging

X-ray imaging has been successfully used in many areas, including medical diagnosis, homeland security, material research, and industrial inspections. In medical imaging applications, X-ray technology has made enormous progress over the decades. In 1890s, X-ray films were used to capture skeletal structures but were limited by low contrast and slow X-ray exposure. Nowadays, X-ray imaging technologies are serving as indispensable diagnosis tools, as shown in Table 2.1.

Table 2.1 Medical applications of X-ray imaging (Imaginis, 2000).

Medical applications	Function of X-ray imaging
Angiography	Imaging of the blood vessels
Arthrography	Imaging of the joints
Barium X-ray	Imaging of the gastro-intestinal (GI) tract
Chest X-ray	Imaging of the thoracic cavity and heart
Cholangiography	Imaging of the bile ducts
Cholecystography	Imaging of the gallbladder
Dental X-rays	Imaging of the teeth and jaw
Lymphangiography	Imaging of the lymphatic system
Mammography	Imaging of the breasts
Myelography	Imaging of the spinal cord
Pyelography	Imaging of the urinary tract
Skeletal X-rays	Imaging of bones
Urography	Imaging of the kidneys and bladder

In terms of technology, computed tomography (CT) is one of the most significant achievements in medical X-ray imaging, and has dramatically improved the imaging of various body parts, especially soft tissue structures which are challenging to traditional radiographic techniques (Selman, 1993). In a typical CT scanner, the patient is scanned by a circular array of X-ray detectors and an X-ray tube moving through a circular field around the patient. The image data collected by the detector array are recorded by a computer synchronized with the moving X-ray tube, and the image of the region of interest is then obtained using the fan-beam reconstruction method (Jain, 1989). The

spiral scanning technique is used in many modern CT machines, where the patient moves longitudinally while the X-ray tube rotates, thus 3-D image reconstruction of body parts is obtained (Bovik, 2000).

There are various forms of X-ray imagers used for different medical applications, including traditional X-ray film, film/screen combination, X-ray intensifier-TV system, and digital solid state detector. Recently, solid state X-ray imagers based on amorphous semiconductors, especially amorphous silicon (A-Si) are of significant interest (Speller, et al., 2001). There are several benefits to using an A-Si imager over other types of medical X-ray detectors. First, an A-Si imager can be manufactured as a uniform flat panel, large enough to cover certain body parts. Secondly, the digital imaging interface enables rapid processing of the X-ray image. Thirdly, unlike CCD based imagers, the A-Si material does not degrade if exposed to X-rays. Finally and most importantly, an A-Si imager can provide much better spatial resolution and dynamic range for medical applications. A typical A-Si X-ray imaging flat panel of 40 by 30 cm can deliver a spatial resolution of 127 μm (Varian, 2003), and medical imaging systems based on A-Si technology are commercially available (Dobbins, et al., 2003).

Besides medical imaging, security applications are another major field for X-ray imaging technology. In airports and other locations where security is critical, X-ray scanning has become a routine procedure, with guards looking for firearms, knives, and other types of illegal objects. By being safe, non-destructive, accurate and efficient, X-ray imaging has been the *de facto* standard scanning technology for security applications (Hallowell, 2001). In many X-ray machines, images at two X-ray energies are acquired at the same time, and by combining the images at different energies, it is possible to detect

certain materials which are challenging or undetectable by traditional single energy X-ray systems (Fainberg, 1992). Another promising X-ray technology for security applications is backscatter imaging, which can form an image from the X-rays backscattered from a subject. The backscatter X-ray originates from Compton scattering, and the intensity is determined by the atomic number and density. Backscatter imaging can provide an image of the human surface without showing internal body structures, which makes it an ideal tool to detect illicit objects concealed under clothes (Hallowell, 2001). Recently, a mobile backscatter imaging system was introduced for rapid X-ray inspection of suspected objects without contact, which could dramatically improve the process of cargo screening in ports and on ships (AS&E, 2003).

X-ray imaging is also widely used in industrial applications, and only a few examples are listed as follows. A high resolution X-ray diffraction imaging system was applied to detect semiconductor wafer defects, and the detectable resolution was 1 μm^2 (Lubbert, et al., 2000). A portable digital X-ray imaging system was successfully used to scan valves and pipes in civil engineering (Sawicka, et al., 1999). Another X-ray system was used for printed circuit board (PCB) solder joint inspection, where the final cross-section image was reconstructed from multiple images projected from different views (Roh, et al., 2003).

2.2.3 Food inspection using X-ray imaging

In the field of food inspection, X-ray technology also provides an invaluable means to ensure food safety and to assess food quality. There have been numerous studies conducted on X-ray based food inspection, striving to make X-ray imaging

systems more efficient, effective, accurate and dependable. Graves, et al (1994) proposed a method of analyzing the X-ray imaging system in order to evaluate the contrast obtained when viewing small defects, which linked X-ray imaging contrast to the total system sharpness. Tollner (1993) developed and investigated a model to interpret X-ray imaging pixel standard deviation. Zwiggelaar, et al., (1996) assessed simulation techniques with respect to X-ray imaging applications in food inspection, and discussed the trade-off between irradiation levels and imaging signal to noise ratio (SNR). Zwiggelaar, et al., (1997) also investigated a selective X-ray energy method to improve image contrast and spatial resolution for soft material imaging. Thomas, et al., (1995) used an X-ray system for detection of seed weevil-infested mango. Schatzki, et al., (1997) proposed to use X-ray imaging for defect detection in apples. Statistical (Shahin, et al., 1999) and neural network (Kim and Schatzki, 2000) approaches were applied to X-ray images to detect waterholes in red delicious apples. In both of the apple waterhole studies, texture features were extracted from the apple X-ray images and then classified by either a Bayes classifier or a neural network. It was reported (Shahin, et al., 2002B) that surface bruises of apples could be detected using line-scan X-ray imaging. In this study, spatial edge features and select discrete cosine transform (DCT) coefficients comprised the basis for bruise detection, and it was found that a neural network classifier performed better than the Bayesian classifier.

It has been shown that X-ray imaging can also be applied to the quality assessment of sweet onions (Shahin, et al., 2002A). In this study, features were selected using Bayes methods and a Bayes classifier was used to sort the onions into two classes: good or defective.

A noteworthy study on pistachio nut inspection was conducted using real-time X-ray imaging and digital image processing methods (Casasent, et al., 2001). In the study, line-scan digital X-ray images were acquired by a custom made X-ray machine. The digital images and another set of X-ray film images of the same set of nuts were analyzed using efficient image processing operations, including histogram adjustment, morphological processing, image segmentation, blob analysis and watershed transform. Furthermore, for the same application, a radial basis function (RBF) neural network classifier along with new training procedures were developed to achieve improved classification performance (Casasent and Chen, 2003).

Significant research efforts have been made to improve meat inspection. In 1996, a study was conducted by Schatzki, et al., to test the efficacy of detecting particulate contaminants in processed meat samples by using visual observations of X-ray images. Visual recognition of contaminants in meat samples were studied and analyzed as a function of the thickness, size and shape of the meat as well as the X-ray image texture. It was found that inclusions were more difficult to recognize in textured X-ray images, and the errors varied with the size, shape, and thickness of the inclusions and samples. The processing time required for image acquisition and analysis in this work was seven minutes per sample. As a new approach, Morita et al., (1997) proposed an X-ray system for detecting foreign materials in food and agricultural products that used soft X-ray radiation and a high-resolution image intensifier. A noteworthy study of X-ray imaging for the accurate detection of bone and cartilage fragments in poultry meat was presented by Papanicolopoulos, et al., (1992, 1995). The research was based on Rayleigh X-ray scattering and Compton scattering. The method of detecting bones used the angles of

Rayleigh scattering and the ratio of Rayleigh and Compton (R/C) scattering. While the theory was plausible, the implementation demanded a very precise angle (0.1 degree angle) for the detection, because at slightly different angles, the R/C value indicated materials with an atomic number different from bone or meat. The method had inevitable signal noises, which made it difficult to achieve the needed resolution and accuracy for the application and implementation. More importantly, as in any other method, the uneven thickness of each piece of meat (and from one piece to the next) caused vital problems such as incident energy variation, transmission field intensity variation, detection shifting, and scattering changes as realized by the investigators. These variations limited the accuracy, resolution, and application of a real system. The suggested X-ray system unit costs nearly $900,000, making it not feasible for the poultry processing industry.

2.2.4 Challenge of bone detection

The difficulty of X-ray imaging in the poultry inspection application is that the thickness of the poultry meat is not uniform, which makes it difficult to differentiate bone fragments and thicker meat portions. This challenge makes the X-ray machines used for packed food inspection inadequate in this application.

Various approaches have been used to address this thickness issue. Research shows that by immersing the poultry meat in water, the thickness variation was largely smoothed out. It is believed that this success was due to the significant water content in poultry meat. Although successful in thickness cancellation, the method is unlikely to be acceptable for on-line inspection applications. Sanitation would be a major problem.

During the inspection operation, it would be extremely important to prevent microbial cross-contamination through water flow and grease deposit. Efficient disposal of dirty water should be addressed, and frequent cleaning of the food contact surface. Moreover, because the meat products are immersed in water, it would be challenging to implement an efficient meat handling and rejection apparatus.

Another thickness cancellation method adopted by the industry is to press the meat mechanically. A pump is used to press the meat into a pipe, compressing it into a rectangular block of uniform thickness before it is scanned by an X-ray imager. It was reported that the method worked particularly well for ground meat inspection (Hartman, 2001). However, for products such as poultry fillets, where the preservation of the natural shape is desired, this compression method seems unattractive due to inevitable meat damage. Furthermore, once the X-ray machine reports a contamination in the pipe, it is often difficult for human inspectors to determine the exact location of the defect in the chunk of rejected meat (Graves and Batchelor, 2003).

2.2.5 Dual energy X-ray Imaging

Note that all of the above methods for thickness cancellation are based on X-ray imaging with a single energy X-ray. It is arguable that by combining X-ray images from two energies, more information could be inferred, which could help address the thickness variation problem. An experiment using dual energy X-ray was conducted in this study, which is described in section 3.2. A brief overview of the background and applications of the dual energy X-ray imaging system are presented as follows.

It is known that the attenuation characteristic of any material changes with X-ray energy, and for different materials, the change is different. This difference in changes

could make it possible to differentiate two materials based on the images at different X-ray energies. As mentioned above, X-ray attenuation is the combined processes of photo electronic absorption and Compton scattering. It was found that the photo electronic absorption process is dependent on the atomic number and dominant at lower energies, while Compton scattering is dependent on electron density of the object and dominant at higher energies. It has been shown (Lehmann, et al., 1981) that by combining the X-ray images at two energies (low and high), two new transformed images can be generated to represent the integrated photo electronic and Compton scattering components. The two components are energy independent, and can be linearly combined to cancel unwanted material or highlight targeted material in the final image.

The dual energy imaging technology has been widely employed in the medical and security fields. Medical applications such as Bone absorptiometry, digital mammography, and computed radiography (CR) are discussed below, as well as a brief description of explosive detection with dual energy X-ray.

Dual energy X-ray absorptiometry (DXA) is seen as the single most useful method for bone mineral analysis, and the technique of choice to diagnose osteoporosis (Genant, et al., 1996, Blake and Fogelman, 2002). In a typical DXA machine, X-ray images at two distinct X-ray energies are captured of part of a human body. In both images, pixels representing bones are separated visually from the non-bone regions. The grayscales of "bone pixels" in the two images are then used to calculate the bone mineral density (BMD). The BMD of a specific patient is compared with a population average to generate a score, indicating risk of osteoporosis. A DXA scan typically takes several minutes, needs only negligible radiation to provide an accurate prediction for

osteoporotic fracture, and can be used as an excellent measure of treatment response (Moses, 2000).

Dual energy X-ray imaging can be applied in mammography to enhance the contrast of calcified tissues against the background structure caused by soft adipose and glandular tissues (Johns, et al., 1985). A recent study showed that digital subtraction technology could be applied to dual energy mammograms to detect micro calcifications, as well as provide an estimation of noise levels and factors such as X-ray energies, tissue composition, and breast thickness (Lemacks, et al., 2002). X-ray sources used in traditional mammography are believed to be no longer suitable for dual energy detection, and dichromatic X-ray sources were developed to improve the detection performance of dual energy mammography (Tuffanelli, et al., 2002). The most significant benefits of dual energy mammography are improved sensitivity and reduced radiation dosage.

For computed radiography, dual energy X-ray imaging serves as a powerful tool to enhance images. In a typical application, X-ray posterior-anterior (PA) images of a patient's chest are captured at two different energies. Typically the high energy is 110-150 keV and the low energy is usually 60-80 keV. After combining the low and high-energy images, two new images are generated. One of the new images shows only the skeletal and calcified structures, while the other represents only the soft tissue information without any bone occultation (General Electric Company, 2003). This technology provides radiologists a much better opportunity to diagnose lung cancer in its early stage, because it was found that lung cancers that are missed on traditional chest radiographs were at least partly obscured by bone (MacMahon, 2001).

For aviation security, the dual energy X-ray imaging technology can be used to detect explosives and other illicit materials which cannot be identified by signal energy X-ray systems (Fainberg, 1992). In a typical dual energy baggage scanner (Tyson, 2003), two X-ray detectors are used for the high and low energy X-rays. The two images of different energies are then used to categorize the pixels into organic (i.e. explosives, drugs, and food), inorganic (i.e. glass and plastic), and metallic materials. Usually a color-indexing scheme is used to visually discriminate the different types of materials (so called color X-ray), and the human operator can make further decisions based on the conveniently displayed information.

Based on the numerous applications that were successful using dual energy imaging, it is presumable that the dual energy X-ray approach would also be a promising solution for the thickness variation problem in deboned poultry detection (Graves and Batchelor, 2003). Commercial inspection machines using dual energy methods are available for bone detection (Jamieson, 2002). The dual energy system can detect ossified bone fragment in chicken fillets, thighs and nuggets. However, soft bones are still not detectable by the X-rays (Graves and Batchelor, 2003).

Because in both the high and low energy X-ray images, different materials or even the same material with different thickness cause great variations to image intensity levels, a neural network is usually used as a classifier to analyze the images based on historical training. Human inspectors are required to recalibrate the neural network during operation (Graves and Batchelor, 2003).

Limitations of this dual energy method come from the underlying physics. Compared to the human body, poultry fillets are much thinner and softer, thus in order to

have high quality X-ray images, it is necessary to use lower energies than those for medical radiography. This requirement implies that the difference between the high and low energies is small, which leads to insufficient contrast between the images taken at the two energies. Certainly this limited contrast can be improved by using a neural network in the post-processing stage to extract some information of bone existence. However, it is not surprising that the less obvious bones, those that are thinner or softer, would be missed due to the imperfect input information.

In order to address the thickness variation problem, a more plausible method is to acquire the thickness information directly, and then compensate for the thickness effects in the X-ray images. A brief review of optical methods for range finding is given in the next section.

2.3 Optical range imaging technologies

Optical range imaging refers to the technology of determining three-dimensional (3D) information of an object via optical sensing. A range imaging system can provide surface geometry measurements in the form of a range image, which is in fact an array of 3D coordinates for the object surface.

The optical range imaging methods can be categorized into passive and active approaches. Typical passive approaches include structure from shading (Horn and Brooks, 1986; Zheng and Chellappa, 1991), and passive stereo imaging, also known as stereo vision (Grimson, 1981). The passive methods are well established in computer vision fields, and can recover high accuracy 3D information about the surface (Godin, et al., 2002). The passive methods are based on the visible feature points under ambient illumination. However, in many cases, the available feature points are often coarsely

dispersed over the surface, which prevents the passive methods from achieving sufficient spatial resolution. Fortunately, active methods address this issue by introducing active illumination (or other excitation methods) to "manufacture" dense features on the object surface, thus making it possible to recover high-resolution 3D information. The active methods can be divided into two broad categories: time of flight methods and active triangulation methods. Due to the advantages of the active approaches, only active range imaging methods are described below.

The time of flight (TOF) method determines the distance from a light source to an object by directly measuring the time of light traveling back and forth between the light transmitter/receiver and the reflecting object. Due to the similarities with the working principles of Radar, laser range imaging that uses the TOF method is often referred as Light Detection and Ranging (LIDAR) (Pace, et al., 2003). The TOF method is widely used in fields such as military applications, remote sensing, and atmospheric research. In applications where high accuracy of distance measurements is desired, strict precision requirements are imposed on time measurements. For example, in order to achieve a distance accuracy of 0.1 mm, it is necessary to measure time with an accuracy of 0.67 picoseconds.

A similar range imaging technique called light in flight holography uses a very short light pulse to scan the object. In a typical light in flight system, two short coherent laser beams are used, a reference beam and an object beam. The reference beam reaches an image screen directly, while the object beam is reflected by the object surface. The interference of the two beams on the image sensor is then analyzed to give a full reconstruction of the object surface (Abramson, 1991, Carlsson, 1993). Because the

method is based on interferometry, superior distance measurement resolution can be achieved. For example, coupled with digital reconstruction, the depth resolution reaches 7 μm (Carlsson, et al., 2001).

The range imaging method that takes advantage of Moiré patterns is called the Moiré method. A Moiré pattern is the interference pattern created when two uniform linear gratings superimpose, and there is an orientation mismatch and/or pitch mismatch between the two gratings. The dark and bright bands in a Moiré pattern are called Moiré fringes, which can be correlated to the 3D geometry of the object surface (Xu, et al., 2001). The advantages of the Moiré ranging method are its high accuracy, low cost and simple optical setup (Ratnam, et al., 2001).

The active triangulation method, also referred to as the structured light method, is also widely used for 3D reconstruction (Jalkio, et al., 1985; Toyooka and Iwasa, 1986; Sorgel and Schalkoff, 1997; Sjodahl and Synnergren, 1999). In a typical structured light system, a specially designed lighting pattern is projected on the object, and the deformed fringe pattern on the surface is recorded by a camera. If the spatial relation of the light source and camera is calibrated beforehand, it is possible to reconstruct the 3D information of the points in the fringe pattern using triangulation. This active triangulation approach is preferred in many applications over other ranging methods. The advantages include its low cost, high 3D resolution, convenient customization, and robustness against environment or object variation (Chen, et al., 2000). In this study, a structured light system using a laser imaging sheet and two high-speed cameras has been developed.

2.4 Digital image processing and machine vision

Digital image processing refers to manipulation and analysis of a two dimensional picture by a digital computer, or digital processing of any two dimensional data (Rosenfeld and Kak, 1982; Jain, 1989). The discipline of digital image processing covers many topics including representation, enhancement, compression, analysis, and recognition. In this study of deboned poultry product inspection, which can be seen as a typical machine vision application, various digital image processing techniques are used to ensure reliable and efficient defect detection. In the following parts of this section, selected aspects of digital image processing in machine vision applications are briefly discussed.

2.4.1 Machine vision

As machines that can "see", modern machine vision systems have a broad spectrum of applications, such as industry, security, and medical diagnosis. For industrial applications, machine vision technology often includes the engineering of an integrated inspection system with optical, electronics, mechanical and software components. Machine vision systems are used to detect defects, improve efficiency and ensure safety (Graves and Batchelor, 2003).

A typical machine vision system consists of one or more image sensors, illumination sources appropriate for the specific application, appropriate mechanism for presenting the objects, one or more image digitizers, and one or more digital image processors or computers. In a typical system, the image of an object is acquired by the image sensor, converted to digital format by the digitizer (also known as the video

grabber), and processed by the image processor. Then the image processor analyzes the image, extracts necessary information, and makes a decision about the object, for instance, to pass or reject the product.

Machine vision and computer vision are closely related in many aspects. However, it is necessary to recognize the significant distinctions between these two well-established disciplines. Computer vision approaches typically emphasize the analysis of given image data, attempting to maximize the insight and make sensible decisions, while machine vision approaches tend to control and optimize the image sensing process, which can maximize available useful information and ease the subsequent image processing tasks. On the tradeoff of performance and speed, computer vision researchers usually pursue optimal performance and regard processing time a secondary issue, while machine vision engineers push for maximum throughput as long as the system performance is satisfactory according to certain criteria (such as a better error rate than human inspectors). Both computer vision and machine vision have their own limitations, and often times the limitations of one of them can be overcome by the other's strengths.

A successful machine vision system should be reliable, fast, and consistent. In the application of natural product inspection, product variability is one of the most significant challenges to machine vision engineers. Unlike the manufactured parts encountered in many industrial applications, it is nearly impossible to find two identical natural products even in the same batch. Natural products have different sizes, shapes, colors, and textures in the captured images, and the defects on the products vary in terms of severity and location. These variability problems, among other challenges, should be addressed by an integrated solution of optical, electrical, mechanical and algorithmic approaches.

2.4.2 Image representation and acquisition

As mentioned previously, a digital image is a 2D array of real numbers (or rarely, complex numbers) represented by a finite number of bits. The basic elements of a digital image are called pixels, an abbreviation of the words *picture elements*. The grayscale of each pixel is the incident light intensity at the specific location of the image. In general, any image can be seen as the camera sensed result of the interaction of a certain type of radiation and some material. By the nature of the interaction, images can be roughly classified into reflection images, such as ordinary photographs; emission images, such as thermal images; and absorption images, such X-ray images (Bovik, 2000).

In order for the images to be processed by a digital computer, they should be first converted to digital format, i.e., the images should be defined on a discrete space/time domain, and should take values from a finite discrete set of values (Bovik, 2000). This conversion process is called analogue to digital conversion (ADC), and is comprised of two steps, sampling and quantization. In a typical machine vision system, the ADC process is usually conducted by the image digitizer, with a high frame rate and sufficient quantization bits.

2.4.3 Image enhancement

In most cases, the acquired images need to be enhanced before further processing. In general, the goal of image enhancement is to strengthen certain image features for subsequent analysis (Jain, 1989). Enhancement involves noise reduction, edge sharpening, contrast enhancement, and pixel interpolation. Although the enhanced image is more

useful for display or analysis than the original, it is noteworthy that the process of image enhancement does not add any information content to the image data.

There are four types of image enhancement techniques: point operations, spatial operations, transform operations, and pseudo coloring (Jain, 1989). In practice, the most frequently used enhancement techniques are histogram modeling (a type of point operation) and spatial filtering (a subset of spatial operations).

The histogram of an image represents the relative frequency of occurrence of various gray levels in the image. By modeling the histogram, it is possible to improve the contrast of the image, or modify the shape of the histogram to meet certain requirements. One example is to perform histogram stretching, also known as contrast stretching, where the histogram is stretched to fill the full range of the grayscale, for instance, 0-255 for an eight-bit image.

Another example is histogram equalization, also known as histogram flattening, which changes the grayscales of the pixels so that a uniform histogram is obtained. It is believed that an image with a perfectly uniform histogram has the largest possible amount of information.

Spatial operations are often used to remove or suppress image noise by convolving the image with a *spatial mask*. The spatial mask is usually a small square image patch. For instance, the mask for a spatial averaging operation could be a 3 by 3 mask with value 1/9 at each pixel, suppressing noises in the original image significantly after convolution. Other than the convolution based approaches, a spatial operation called median filtering is also used to enhance images. With a median filter, the input pixel is replaced by the median of the pixels in a small neighborhood of the original pixel.

Median filtering is very effective in removing the so-called *salt and pepper* noises in images.

2.4.4 Image segmentation

Image segmentation is the process of dividing an image into a set of homogeneous regions, which is a crucial step in image processing. A homogeneous region refers to a group of connected pixels that are similar in terms of grayscale, color, or texture (Bovik, 2000). In many applications, image segmentation provides a convenient starting point for subsequent image manipulation, classification and understanding. For instance, to grade fruit quality, the fruit in an image should be segmented from the background.

Among the many segmentation techniques, threshold-based, edge-based, region-based, and texture-based methods are most frequently used. In threshold-based methods, the segmentation decision of each pixel is made by comparing the grayscale of the pixel with a threshold value. Often the threshold value is determined by examining the shape of the histogram, and the valley in the histogram could be used as the threshold value (Jain, 1989). The threshold could also be determined automatically based on the statistical distribution of the pixel grayscales. For instance, assume there are two classes of pixels, i.e., object pixels and background pixels, then the optimal threshold can be chosen to maximize the between class variance (Otsu, 1979).

Edge-based segmentation is also known as the boundary-based method, which focuses on finding the boundaries that separate different regions. Typically, the edges in an image are detected and tracked, then an edge linking process is used to obtain a tentative boundary of regions, and finally the boundaries are refined to yield a final

segmentation. In an image where objects are touching or overlapping, the boundary could be broken due to noise, which makes it difficult to segment using edge-based methods.

Region-based methods often emphasize direct detection of homogenous regions in an image according to the spatial similarities of pixels. Common techniques include region splitting, region merging, and region growing. In general, if a region is not uniform (according to a certain uniformity measure) it is split, and if two adjacent regions are similar (again based on a certain similarity measure) they are merged. In the region growing approach, a pixel adjacent to an established region is included in the region if the pixel shares similar features to the pixels already included in the region. Note that the three methods have different starting points. The region splitting starts by assigning the whole image as a single region, region merging starts by assigning each pixel as a region, and region growing starts from chosen *seed pixels*.

The goal of texture-based segmentation is to partition an image into a set of homogeneous textured regions. Texture is a ubiquitous phenomenon in nature, and there has been extensive research conducted on texture-based image analysis. Although there is not yet a widely accepted definition of the term *texture*, many methods have been proposed to grasp the texture features in images. One widely used method is filtering, which filters the image with a group of spatial filters (a filter bank), and the output of the filter bank at each pixel is used to assign a label to each pixel (Malik and Perona, 1990).

2.4.5 Blob analysis and feature extraction

The blob analysis method (Jain, et al., 1995), also known as blob coloring, blob labeling, or region analysis, is often used in machine vision image processing. A blob is a

group of connected pixels that share similar intensity, color or other properties, which can be obtained as the homogeneous region yielded by image segmentation. For each blob, features such as area, perimeter, location, and mean grayscale can be conveniently calculated. The features can then be used in a classification process to identify the nature of the object, for instance, whether it is acceptable or defective.

One of the advantages of blob analysis over pixel-based analysis is that the image processing tasks can be performed on a blob-by-blob basis, instead of on numerous pixels in an image, thus leading to the need for less computation power.

Another benefit is that the properties of the objects of interest can be conveniently represented by expected blob features, such as area and shape, which lead to efficient and effective object classification.

2.4.6 Real time implementation issues

In many, if not all, machine vision applications, real-time capability is an essential requirement. It is important for the imaging system to respond to the incoming image data in a timely manner. In practice, it is recognized that real time image processing involves three major tradeoffs: performance vs. image resolution, performance vs. data bandwidth, and synchronization vs. number of concurrent tasks (Laplante and Stoyenko, 1996). Besides adopting faster hardware, it is usually more important to use appropriate software architecture to ensure reliability and functionality of the overall system, and maximize the potential of the available hardware. Among various useful techniques, double buffering and parallel processing are briefly described as follows.

Double buffering refers to an image processing technique using two image buffers to ensure data integrity without interfering with real-time image acquisition. The basic

idea of double buffering is to process the image in one buffer while the new image data are grabbing into another buffer. It is assumed that the maximum processing time for a buffer is shorter than the minimum acquisition time for a buffer; otherwise, the image data being processed could be contaminated. Once the image processing of the first buffer is complete, the program will start waiting for the second buffer. As soon as image acquisition of the second buffer is complete, the program starts processing it, and acquisition of the first buffer begins at the same time. This simple mechanism guarantees that at any time, the image buffer being processed is not being updated by another source, thus data integrity is ensured.

In order to maximally exploit the hardware computing power, it is often helpful to process the images in a parallel manner. In a typical real-time machine vision system, image processing throughput is affect by the acquisition frame rate of the camera and image grabber, processing speed of the onboard image processor, data bandwidth for transferring the image stream to host CPU, processing speed of host CPU and data bandwidth between CPU and memory. The speed of image processors is usually faster than the data transfer speed in the data buses. If a single thread of image processing is active, sometimes the onboard processor and host CPU are idling while the mass of image data is being transferred. Computation power of the onboard processor and host CPU can be employed more efficiently if there are multiple processing threads working in parallel, and the overall throughput and responsiveness of the overall system would be improved with the same hardware.

CHAPTER 3

OBJECTIVES

The overall objective of the research is to develop a combined X-ray and laser range-imaging system for sensitive detection of physical contaminants in meat fillets on poultry processing lines. The method should be able to overcome the obstacles of traditional X-ray technology due to uneven thickness of meat, and enable accurate on-line imaging identification of bone fragments in deboned meat. Furthermore, the performance of the detection system should be equal to or better than the current industry practice of hand checking for ensuring product quality and safety. The specific goals of this research are to:

(1) Analyze X-ray absorption and determine optimal sensing parameters for detection of bones in poultry materials and to configure a high-contrast X-ray imaging system that is highly sensitive to bones and other physical contaminants.

(2) Develop a combined X-ray and laser range-imaging system to produce integrated and thickness-invariant X-ray images for enhanced sensitivity. This will include the development of image pattern recognition algorithms to identify bones and make rejection decisions.

(3) Evaluate the performance of the combined X-ray and laser range-imaging system through on-line testing.

CHAPTER 4

X-RAY IMAGING WITH SINGLE ENERGY AND DUAL X-RAY ENERGIES

In this chapter, two imaging modes were studied for deboned poultry inspection. For X-ray imaging with a single energy level, an objective measure was proposed for X-ray image quality evaluation. For X-ray imaging with dual energies, a least squares based method was studied for bone fragment detection.

4.1 Equipment overview

The inspection system consists of an X-ray imaging system and a motor-driven conveyor belt which carries poultry fillets to be inspected. X-ray images were collected by a high-resolution line-scan digital X-ray camera (Hamamatsu Photonics Systems, Japan), featuring 1280 pixels per line at 0.4 mm resolution (thus the total effective reception length is 512 mm). A Genesis imaging board (Matrox Electronic Systems Ltd, Canada) grabs the images with an onboard 12-bit digitizer. The imaging board was installed on a custom built Pentium 4 PC, which handles real-time acquisition and analysis of X-ray images, and interacts with the human operator.

4.2 Evaluation of X-ray image quality

4.2.1 Objective index for X-ray image quality evaluation

The X-ray image is the foundation of the physical contamination detection. Image quality involves many factors, including resolution, contrast, and noise, etc. Image

resolution in X-ray imaging is related to the detector pixel pitch and the focal size of the X-ray tube, both of which are unchangeable in our system. The quality evaluation and enhancement focus the contrast and noise characteristics.

The image contrast is defined as the intensity difference of the object and its background (Dowsett, et al., 1998), or

$$C= (I_1-I_0)/I_0 \quad (4.1)$$

where

C: the contrast, and

I_1: the intensity of the object and I_0 is the background intensity.

In the application of bone fragment detection, it is desirable that the difference between the bone fragment and normal meat tissue be maximized. Certainly, sufficient contrast between the bone fragment and its surrounding meat tissue is crucial in order to detect the intrusion.

Noise level in an image is characterized using the signal to noise ratio (SNR). A higher SNR leads to better image quality. SNR is defined by

$$SNR=I_{signal}/\sigma_{noise} \quad (4.2)$$

where

I_{signal}: the signal intensity, and

σ_{noise}: the standard deviation of the noise.

In the field of X-ray imaging, it is known that contrast and SNR are both related to X-ray energy. In general, better image contrast can be achieved at lower energies, while better SNR can be observed by increasing X-ray energy. Thus in order to choose a

suitable X-ray energy for a specific application, both contrast and SNR should be considered at the same time, and a tradeoff between these two is usually needed.

A measure called detective quantum efficiency (DQE) is often used to evaluate X-ray image quality by the medical radiographic community. For an X-ray image sensor, the DQE is defined as the ratio of the SNR of output signal to the SNR of input signal, or

$$DQE=SNR^2_{out}/SNR^2_{in} \quad (4.3)$$

where

SNR_{out}: the signal to noise ratio at the output of the X-ray image detector,

SNR_{in}: the SNR at the input quantum field at the detector input (Sandborg and Carlsson, 1992).

In practice, the DQE value is often calculated as

$$DQE(f) = \frac{G^2 \bullet MTF^2(f) \bullet \Phi}{NPS(f)} \quad (4.4)$$

where

f: the spatial frequency (line pairs/mm);

G: the detector gain, a constant determined by the design of the image detector;

MTF: the modulation transfer function of the imaging system;

Φ: the X-ray quanta per unit area, or quantum flux at the input of the image detector; and

NPS: the noise power spectrum in the output X-ray image (Spartiotis, et al., 2003).

Although DQE has been widely used to evaluate system performance in X-ray imaging fields, it is not used for our application. This is because of the difficulty of

measuring the X-ray quantum flux in our current X-ray setup, and the fact that DQE is not directly linked to image contrast and the SNR.

As an alternative, the author proposes an new objective image quality measure, the combined quality index (CQI), which is defined as

$$CQI = C \bullet SNR \tag{4.5}$$

Where

C: the image contrast, and

SNR: the signal to noise ratio of the output X-ray image.

This quality index is solely image based, i.e., it can be derived directly from the X-ray image. Therefore, it is not necessary to measure or estimate the input X-ray quantum flux. Another advantage of the CQI is that, it combines the effects of image contrast and SNR into one measure, thus it is possible to determine an optimal X-ray energy level by finding the maximum CQI.

4.2.2 Apply CQI to the X-ray images of sample images

In order to find the optimal X-ray energy level for deboned poultry inspection image quality needs to be evaluated at different energy levels. X-ray images of a meat sample are taken at different X-ray energies, and the CQI for each image was calculated and compared. The meat sample was a deboned chicken breast with a bone fragment inside. In the X-ray image, image contrast was calculated in a small region of interest containing the bone fragment and the surrounding meat tissue. Refer to equation 4.1, the mean intensity of the bone fragment is I_1, and the mean intensity of the surrounding meat tissue is I_0. For the SNR calculation, the mean intensity in the background was used as

signal intensity I_{signal} in equation 4.2, and the standard deviation was calculated in the same background region. After the C and SNR were calculated, the combined quality index can be determined by following equation 4.5.

4.3 Bone Fragment Detection Using Dual Energy X-ray Imaging

In order to overcome the thickness variation problem with traditional single energy X-ray imaging technology, we studied a method based on dual energy X-ray imaging. The underlying belief is that, when X-ray energy changes, the attenuation properties of different materials change differently. In this section, the physics background, experimental setup, and image-processing algorithms are presented.

4.3.1 Physics background: dual energy X-ray imaging

For certain applications, two X-ray images at different energy levels of an object are captured and analyzed, yielding valuable information not available when using only one energy level. This method is called dual energy X-ray imaging, which is widely used in security and medical fields. In some security screening systems deployed in airports or other vital locations, dual energy X-ray imaging has been used to scan travel luggage, providing enhanced image details for hazardous object detection (Tyson, 2003). In medical applications, dual energy X-ray imaging has been widely used to analyze body composition, especially for bone mineral density (BMD) measurements, and is often referred as dual-energy X-ray absorptiometry (DXA) (Genant, et al., 1996). Recently, digital dual energy X-ray technology has been applied to other medical applications such

as the chest X-ray (GE, 2003) and mammography (Lemacks, et al., 2002), and has shown significant advantages over single energy X-ray systems.

The physics basis of dual energy X-ray imaging comes from the fact that the X-ray absorption coefficient of any specific material is energy dependent. In other words, the X-ray absorption coefficient of a material is a function of X-ray photon energy. In the application of meat tissue inspection, there are two principle X-ray attenuation processes, photoelectric absorption and Compton scattering. It has been found that, at lower X-ray energies (usually below 30 keV), photoelectric absorption plays a dominant role, while Compton scattering becomes more significant at higher energies (usually above 50 keV). The dependency of X-ray attenuation coefficients of certain material upon energy can be modeled as (Lehmann, et al., 1981):

$$\frac{\mu(E)}{\rho} = a_c f_c(E) + a_p f_p(E) \qquad (4.6)$$

where

$\mu(E)$: the attenuation coefficient,

ρ : the mass density of the material,

a_c and a_p :constants determined by the atomic composition of the material, and

$f_c(E)$ and $f_p(E)$: functions describing the energy dependencies of Compton scattering and photoelectric absorption respectively.

The term $\frac{\mu(E)}{\rho}$ is often referred to as the mass attenuation coefficient. It has been proven that the mass attenuation coefficient of any material can be expressed by the linear combination of any two materials with known attenuation properties (Lehmann, et al., 1981),

$$\frac{\mu_{\xi}(E)}{\rho_{\xi}} = a_1 \frac{\mu_1(E)}{\rho_1} + a_2 \frac{\mu_2(E)}{\rho_2} \qquad (4.7)$$

where

$\frac{\mu_1(E)}{\rho_1}$ and $\frac{\mu_2(E)}{\rho_2}$: the mass attenuation coefficients of materials 1 and 2 respectively, and

a_1 and a_2 : constants determined by the composition of the materials.

Materials 1 and 2 are called basis set materials, and are used to span the "space" of materials. For instance, in the medical dual energy X-ray absorptiometry applications, the basis materials are often aluminum and Lucite, mainly because their atom numbers bracket most materials of clinical interest (Lehmann, et al., 1981).

Equation 4.7 lays the foundation for dual energy X-ray imaging, making it possible to cancel out a certain material in an X-ray image based on attenuation characteristics of two known materials. For example, suppose the basis materials are aluminum and Lucite, then the two basis images (aluminum image and Lucite image) can be derived from the dual energy X-ray images. The derivation is based on the known mass attenuation coefficients of the two basis materials at dual energy levels. Finally, a linear combination of the two basis images generates a "desired" final image, which can be expressed as(Lehmann, et al., 1981)

$$C_0 = A_1 \sin\Phi_0 + A_2 \cos\Phi_0 \qquad (4.8)$$

where

A_1 and A_2 are the two basis images,

Φ_0 is an angle which can be varied to adjust the linear combination, and

C_0 is the resultant combined image, called the *basis projection image*.

It follows that (Lehmann, et al., 1981), an unknown homogeneous sample of material can be canceled in the basis projection image by adjusting Φ_0 in an interactive manner. In the bone detection application, this means that the soft tissue of poultry meat could be canceled out of the final image regardless of thickness variations, as long as the material is approximately homogeneous.

However, the above rationale is based on the assumption that the X-ray source is monochromatic, or that it generates a "pure color" X-ray beam at a single energy level, or wavelength. Unfortunately, most of the current available X-ray sources are polychromatic, and the only adjustable parameter in this regard is the maximum energy, or equivalently, the minimum wavelength of radiation. In order to address this continuous energy issue, the X-ray beam can be filtered by a certain material, thus only a small portion of the incident X-ray spectrum can reach the image sensor. Alternatively, mathematical methods such as the Monte Carlo N Particle model (Batistoni, et al., 2003, Nilsson, et al., 2002) can be used to decompose the continuous energy distribution and obtain a fairly accurate estimate of the material composition.

4.3.2 Selection of high and low energies

To apply the idea of dual energy imaging in practice, the first step is to choose an appropriate combination of low and high energy levels. Selection of low and higher energy levels depends on the nature of the specific application, especially on the radiological characteristics of the relevant material. Theoretically, any two X-ray energies can be used in the dual-energy X-ray setup as long as both of them can lead to an acceptable image quality for the object. In our imaging system, based on image contrast

and noise level, the high energy level selected was 50 keV, and the low energy level was 30 keV.

4.3.3 System setup

In order to apply dual energy X-ray to online inspection tasks, it is necessary to collect X-ray images at different energies in real-time. Some systems have one X-ray source and two imagers binned together that are fine-tuned to different energies. Other systems have two X-ray sources working at different energy levels and two imagers, each of which is synchronized to one of the sources. Obviously the latter setup tends to be more accurate, but more expensive as well.

To explore the potential of dual energy X-ray imaging in industrial poultry inspection, we simulate the dual source dual imager setup using a single source single imager system, which is coupled with precise image registration procedures. The X-ray machine described in section 4.1 was used to conduct the dual energy X-ray imaging experiments. For each sample, X-ray images were taken at different energies, and the series of images for that sample were recorded for further analysis.

4.3.4 Image analysis

For dual energy X-ray imaging, it is essential to ensure that the sample images at different energies are registered to each other at the pixel level. Because our imaging system uses a line scan type X-ray imager, it was impossible to precisely control the location of the object in the view of each image frame. In order to determine the relative translation between image scans, an image based registration method was used. A small

reference object in the image was used to register the relative location of the start point of an image, thus the translation between different images was recovered, and X-ray images at different energy levels were registered to each other.

After registration, for each pixel, its grayscales in both high energy and low energy X-ray images were compiled to form a two-element attenuation vector, which carried information of X-ray attenuations of both energy levels at the pixel level. At this stage, a region of interest was obtained by segmenting the low energy image using a suitable threshold, so that computing power could be focused on the object instead of the image background. The attenuation vectors of all the pixels residing in the region of interest serve as inputs for the following image-processing tasks. In the classic method proposed by Lehmann, et al (1981), the vectors were fed into a nonlinear equation system, solved using an iterative fitting approach. The Compton effect tends to be more significant when the X-ray energy is high, and in an energy region where the Compton effect predominates, the attenuation is approximately material independent (Hale, 1975). It follows that, at an appropriate high energy level, the X-ray attenuation is approximately solely determined by the thickness of the object, no matter if the object consists of a homogeneous material or is a mixture of heterogeneous materials (Jamieson, 2002). Assuming that the thickness is represented by high energy attenuation, while low energy attenuation is affected by both thickness and material characteristics, a linear polynomial model of certain order can be formulated to describe the relationship between the high-low energy grayscale pair for each pixel as:

$$M_L=f(M_H)=a_0+a_1M_H+a_2 M_H^{\,2}+a_3 M_H^{\,3}\ldots+a_n M_H^{\,n} \qquad (4.9)$$

where

M_H and M_L: the log attenuations at a certain pixel using the high and low energies respectively,

f(.): the polynomial model, and

n: the order of the model, which should be determined *a priori* .

The assumption is that if the model is accurate and the material is homogenous through the object, M_L at any pixel in the region of interest can be properly estimated according to the corresponding M_H with negligible error. On the other hand, if there is any intrusion embedded in the object, it will be identified as an outlier in the model, and thus will be detected by image differentiation between the true M_L and the estimated M_L.

In practice, image noises are inevitable in the acquired dual energy X-ray images, thus a 3x3 median filter was used to smooth out most of the "salt-and-pepper" noises. After noise removal, M_H and M_L values for each pixel in the region of interest are taken from the high and low energy images respectively. Assuming that the total number of pixels in the region of interest is N, then the N pairs of (M_H, M_L) are plugged into Equation 4.4, and the model parameters $\{a_0,\ldots, a_n\}$ can be obtained using least squares method.

Given the model parameters, the estimated version of low energy image $\hat{M}_L$ is then calculated using the model expressed by Equation 4.9. The difference image is given as

$$\Delta M = |M_L - \hat{M}_L| \qquad (4.10)$$

The nonzero pixels in the ΔM image are considered as the abnormal signatures, which can be identified using various pattern recognition methods. In this preliminary study the greatest interest was whether the dual energy X-ray imaging could provide sufficient

information for thickness cancellation, so options of target detection algorithms were not explored for this application.

Test results and analysis of this dual energy X-ray approach are presented in chapter 5, nevertheless it should be noted here that thickness variation of the soft chicken tissue can only be cancelled to a limited extent. This limited capability is due to the underlying physics of X-ray attenuation, which is discussed in chapter 6.

Although it is certainly helpful to apply sophisticated machine learning algorithms such as artificial neural networks to further suppress thickness interference, it is better to have a start point where the explicit thickness information could be acquired physically. Thus, in the research of this dissertation, the combined X-ray and laser 3D imaging approach was preferred to solve the thickness variation problem.

CHAPTER 5

COMBINED X-RAY AND LASER 3D IMAGING

As mentioned previously, due to the nature of poultry bone fragment detection problem, the effectiveness of the detection method depends largely on how well the thickness variation is neutralized. In this study, a thickness estimation method based on laser 3D imaging was adopted to measure the thickness explicitly in real-time, which leads to the possibility of complete thickness cancellation.

5.1 Overview of combined X-ray and laser 3D imaging

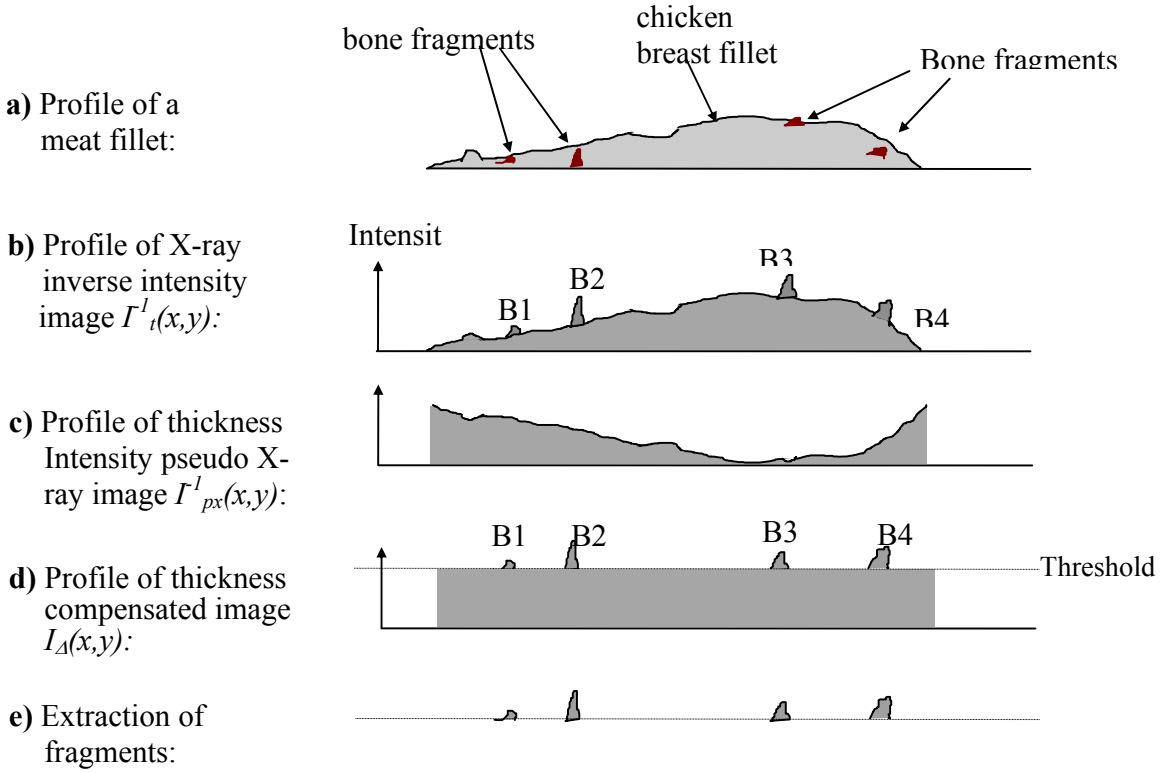


Figure 5.1 The concept of thickness cancellation for bone detection

Figure 5.1 illustrates the conceptual procedure of the sensor fusion scheme. Figure 5.1 (a) shows a chicken fillet containing bone fragments (or other physical contaminations). Figure 5.1 (b) shows the X-ray image profile. The surface topography is obtained using laser range imaging, and then the 3D thickness image is mapped to a pseudo X-ray image as shown in Figure 5.1 (c). By combining the mapped image with X-ray image, the thickness-compensated X-ray image is obtained as shown in Figure 5.1(d). Finally, the embedded bone fragments are segmented as shown in Figure 5.1(e).

Because both X-ray and laser images are function of thickness, a mapping relationship exists between them. The transformation of a depth image T(x, y) to a pseudo X-ray image $I_{px}(x,y)$ can be obtained by applying a mapping function f(.) as:

$$I_{px}(x,y) = f(T(x, y)) \tag{5.1}$$

The mapping function is determined by constructing a model between the X-ray image and laser image of a chicken meat fillet without bone inclusions (the modeling process is described in section 5.4.2). In ideal case, i.e., the meat tissue material of the sample is homogenous everywhere, and the constructed model is accurate, the mapping function results in:

$$I_{px}(x,y) = I(x, y) \tag{5.2}$$

i.e., the transformed pseudo X-ray image $I_{px}(x,y)$ would be the same as the X-ray image I(x, y), no bone fragment inclusions show in the pseudo X-ray image. Thus, the grayscale variation in the X-ray image due to the uneven meat thickness can be eliminated by subtraction:

$$I_{tc}(x,y) = I_{px}(x,y) - I(x, y) \tag{5.3}$$

where

$I_{tc}(x,y)$: the resultant thickness compensated image.

In principle, if chicken meat contains no bone fragment (or other foreign materials), $I_{tc}(x,y)$ should be close to zero for any pixel. In practice, to eliminate any residual or image noises caused by imperfect thickness compensation, $I_{tc}(x,y)$ image can be segmented with a small threshold ε. The bright spots in the segmented image would indicate possible intrusions, which will be identified by further analysis. Certainly if the segmented image is all zero, it can be concluded that the meat sample of interest is free of any bone fragment or other physical hazard.

The image data from both X-ray and laser subsystems are acquired simultaneously and then integrated to compensate for the thickness variation. The embedded bone fragment, if any, will be identified based on the final combined image.

The performance of the overall system depends on accuracy of the information provided by the individual subsystems, and the synchronization/coordination between them. Due to the unique synergism of X-ray imaging and laser 3D imaging in this study, special consideration and modification are taken to fit the application of poultry meat inspection.

5.2 Laser 3D imaging

The 3D thickness profiles of the surface of the poultry fillets are acquired online using the laser range imaging technique. In general, optical range finding can be defined as the process of determining the distance (or depth) from a given observation point to all points of consideration in a scene. The technology of optical range finders has been widely used in many applications, such as airborne remote sensing survey, medical imaging, and reverse engineering, etc., to obtain 3D description for object of interest.

Two of the most popular methods for range finding are time-of-flight (TOF) based technique (Pace, et al., 2003), which measures the travel time of light to infer distance, and structured light based technique (Sjodahl and Synnergren, 1999), which reconstructs 3D structure based on triangulation.

In this study, a structured light system (SLS) is used to obtain the 3D profiles for poultry samples, because compared with the TOF approach, structured lighting method tends to be more accurate, flexible, robust and cost effective (Chen, et al., 2000).

In general, a SLS uses laser beams to project intense and sharp parallel lines onto an object and registers the resultant light pattern with a camera. Once the relative position of the camera and the light source are known, the distance from the camera to the object can be calculated using triangulation, which leads to the depth information. The complete 3D surface of the object can then be reconstructed if the entire object has been scanned. The schematic setup of a typical system is depicted in Figure 5.2. With a cylindrical lens based optical system, a laser projects a light plane onto the object. The thickness variation of the object will be registered by the light stripe on its surface. In the demonstration of Figure 5.3, the complete 3D surface profile can be estimated from the set of light stripes posed on the object over time.

The geometry of the structured light system is shown in Figure 5.4. The coordinates of an image point are given relative to the center of the camera lens. A single illumination is represented at angles θ_x and θ_y to the z-axis in the x and y directions, respectively.

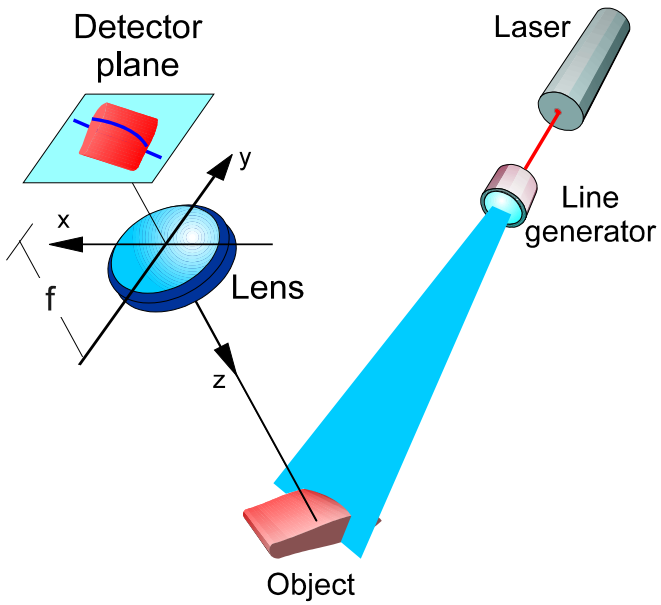


Figure 5.2 Schematic setup of a basic laser based structured light system for thickness measurement.

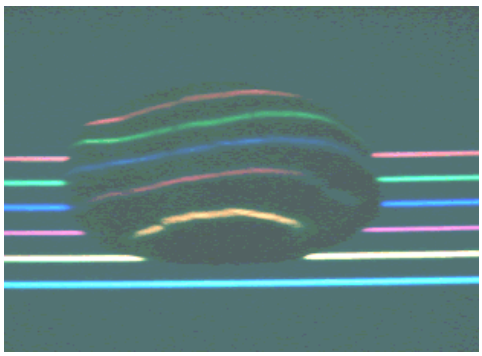


Figure 5.3 Surface structure reconstruction using multiple light stripe scanning

The stripe illuminates the object at the coordinates (x_0,y_0), which is given by

$$\begin{aligned} x_o &= x_s - (z_o - z_s)\tan\theta_x, \\ y_o &= y_s - (z_o - z_s)\tan\theta_y. \end{aligned} \qquad (5.4)$$

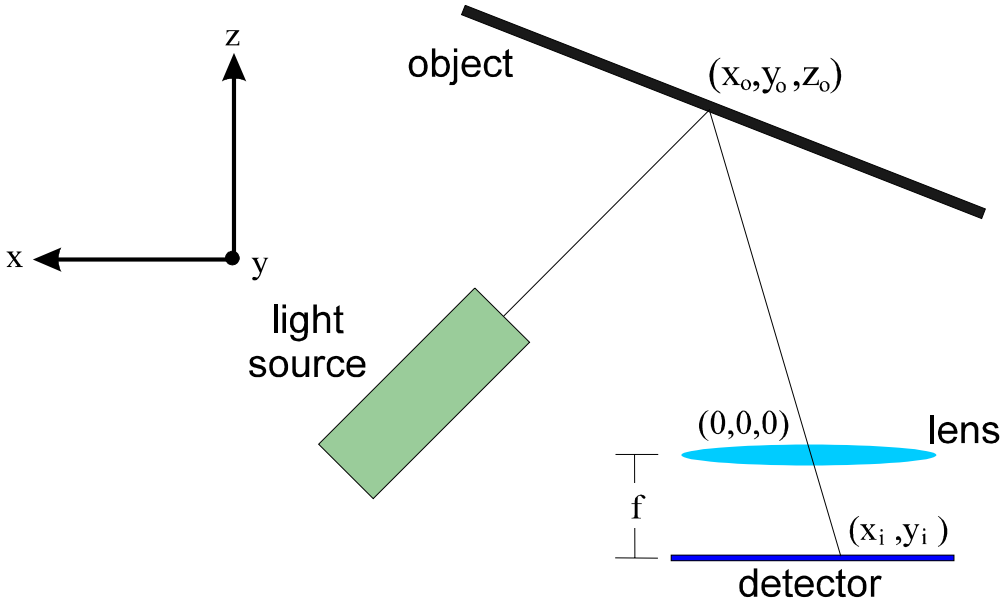


Figure 5.4 Geometric model of the triangulation procedure in structured light systems

An image of this point is formed on the detector plane at the coordinates

$$\begin{aligned} x_i &= \frac{f}{z_o} x_o = \frac{f}{z_o}\left(x_s + z_s \tan\theta_x\right) - f\tan\theta_x, \\ y_i &= \frac{f}{z_o} y_o = \frac{f}{z_o}\left(y_s + z_s \tan\theta_y\right) - f\tan\theta_y. \end{aligned} \tag{5.5}$$

The distance between the receiving lens and detector, f, is nearly equal to the focal length of the receiving lens. If this image point is compared to the image point produced by a calibration object at the reference distance z_{ref}, the displacements of the image are given by

$$\begin{aligned} \Delta x_i &= f(x_s + z_s \tan\theta_x)\left(\frac{1}{z_o} - \frac{1}{z_{ref}}\right), \\ \Delta y_i &= f(y_s + z_s \tan\theta_y)\left(\frac{1}{z_o} - \frac{1}{z_{ref}}\right). \end{aligned} \tag{5.6}$$

The x displacement can be used to calculate the object distance:

$$z_o = \frac{z_{ref}}{1 + \Delta x_i z_{ref} / f(x_s + z_s \tan\theta_x)} \qquad (5.7)$$

The laser 3D imaging subsystem constructed for this study is shown in Figure 5.5. Note that two cameras are used in this setup. Part of the cross section curve generated by the laser sheet and the meat surface could be invisible to one of the cameras, due to meat occultation and/or random light reflections. By using two cameras, as shown in Figure 5.5, the other camera can obtain a better view of the missed laser reflection, thus the robustness against occultation and lighting irregularity of the overall imaging system is improved. If in some circumstances neither of the cameras could get a clear view of the whole profile, the missed portion will be estimated using interpolation methods.

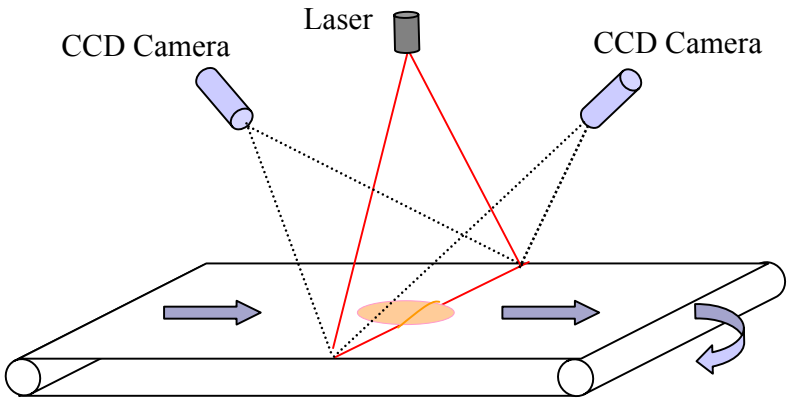


Figure 5.5 Laser 3D imaging subsystem for poultry meat thickness profiling

5.3. X-ray imaging subsystem for the combined approach

The imaging hardware of the X-ray subsystem used in the combined approach is essentially the same as the one used in the previous single X-ray and dual energy X-ray studies, except for a few but crucial modifications. One of the most important changes on

the X-ray imaging part is the triggering mechanism for image acquisition, which is described in details as follows.

5.3.1 External trigger

In order to perform thickness compensation for each pixel in the X-ray image correctly, it is crucial to synchronize the X-ray and laser imaging subsystems. For this purpose, an encoder is mounted on the shaft of the conveyor rotor and connected to the imaging system. The encoder records the movement of the conveyor belt and sends out continuous electronic pulses. The frequency of the electronic pulses is proportional to the speed of the conveyor belt movement, or more precisely, the number of pulses generated during any time period, N_{pulse} , is proportional to the distance of the conveyor belt movement, $d_{conveyor}$,

$$N_{pulse}=k_{encoder}\, d_{conveyor} \qquad (5.8)$$

where

the parameter $k_{encoder}$ is determined by the specification of the encoder.

In our system, the pulse signal from the encoder is transformed to TTL levels, divided to an appropriate frequency range (<2kHz) by a set of pulse dividers, and then is used to trigger both the X-ray and laser image grabbers, as shown in Figure 5.6.

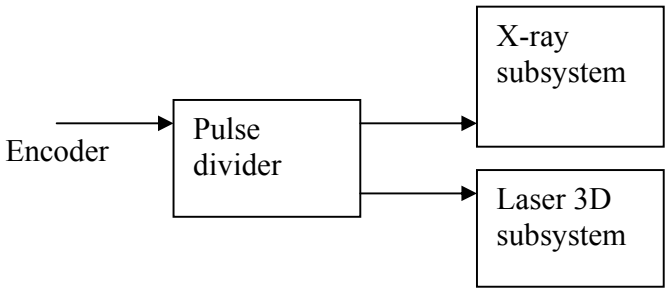


Figure 5.6 Synchronizing X-ray and laser image acquisitions using encoder

5.3.2 Image artifact

The principle of Figure 5.6 works well to synchronize the X-ray and laser subsystems, enabling the two subsystems to share a common location index for any moving object on the conveyor belt. However, it leads to significant artifacts in the X-ray images due to inconsistent exposure time, which would cause significant classification errors.

The cause of these artifacts is related to the principle of X-ray image integration and the jittering nature of the conveyor movements. Integration of the X-ray photons in the X-ray detector is a function of the integration time, or exposure time. It follows that fluctuations of the exposure time will lead to fluctuations of grayscales in the resultant X-ray image. In the combined X-ray and laser imaging system, the exposure time of X-ray image is determined by the width of encoder pulse, which is further dictated by the moving speed of the conveyor belt. Unfortunately, the speed of the conveyor belt is never constant. The cause of the speed jitter can be traced to a series of transient mechanical and electrical irregularities, which are inevitable in practice. Although some of the causing factors can be suppressed to a limited extent, the speed jitter problem can not be effectively eliminated.

Initially, a software-based post-processing method was considered to remove the artifacts, which attempted to identify and remove the artifacts in the imperfect images based on texture features. However, even when some of the artifacts were cancelled out by the algorithm, the sharpness of the X-ray image was also impaired. Furthermore, it is possible that some physical contaminations share similar texture characteristics with the

artifact pattern, which could be mistakenly smoothed out by the software method. Based on these considerations, a hardware method is designed to remove the image artifacts.

5.3.3 Hardware solution for the artifact elimination

Principle

As mentioned above, the exposure trigger for the X-ray imaging subsystem is connected to the encoder mounted on the conveyor rotor via a set of signal amplifying and pulse dividing circuitry.

When the encoder pulse jitters, the exposure time for the X-ray image line changes accordingly. In order to eliminate the negative effect of jitter problem, it is desirable to add a set of pulse conditioning circuitry between the encoder and X-ray detector, which can keep the exposure time constant even if the encoder pulse jitters.

In order to meet the requirements, an internal timing signal of the X-ray detector, HSYNC, is intercepted and combined with the encoder pulse. HSYNC stands for horizontal synchronization, and it is the signal initialing photon integration for each line. There are two image acquisition modes for the X-ray detector, i.e., internal or external modes. According to the detector design, HSYNC signal is generated by the control unit of the X-ray detector as digital pulses of a constant frequency, if the image acquisition mode is set to “internal”. The frequency of HSYNC can be specified by a human operator via a command console in this mode. If the image acquisition mode is set to be “external”, the frequency of HSYNC will totally depend on the external trigger, which could jitter from time to time. The main idea of the hardware solution is to “trick” the control unit

into the internal acquisition mode and for the detector to fabricate a new version of HSYNC which is synchronized with the encoder trigger pulse, as shown in Figure 5.7.

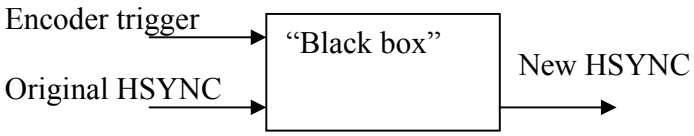


Figure 5.7 Diagram of the hardware solution for artifact removal: the black box fabricates a new HSYNC based on the encoder trigger pulse and the original HSYNC signal

Implementation

The implementation of the design is depicted in Figure 5.7. There are two main portions of the pulse conditioning circuitry, one is the pulse tracking module (PTM) implemented with a phase locked loop (PLL) and the other is the pulse sampling module (PSM) constructed by D flip-flops. The two modules are shown in Figure 5.8 and Figure 5.9, respectively.

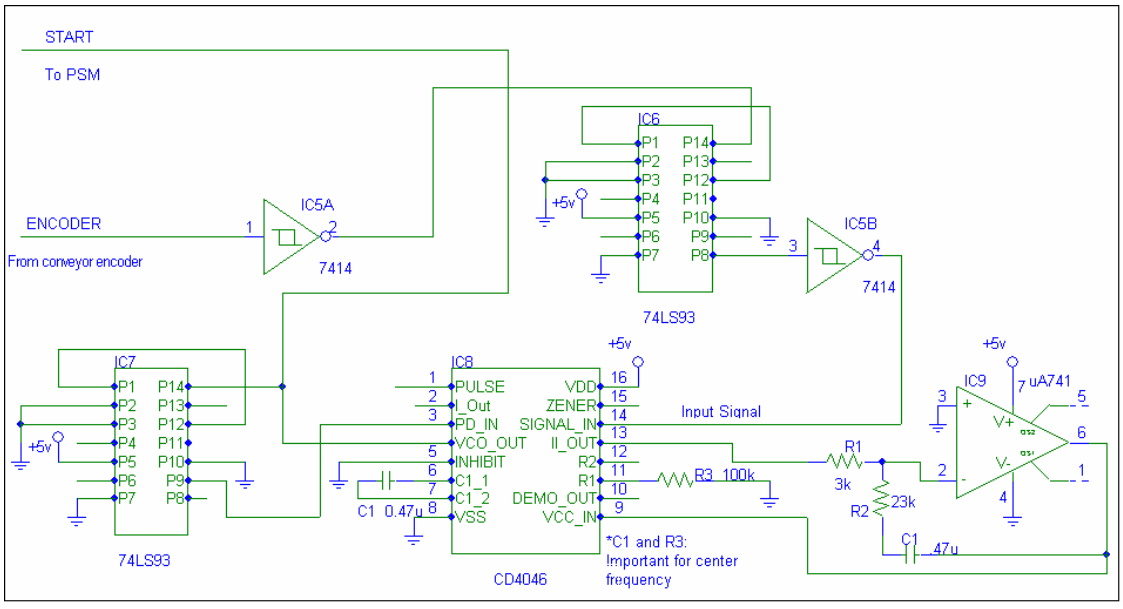


Figure 5.8 Schematics of the pulse tracking module (PTM)

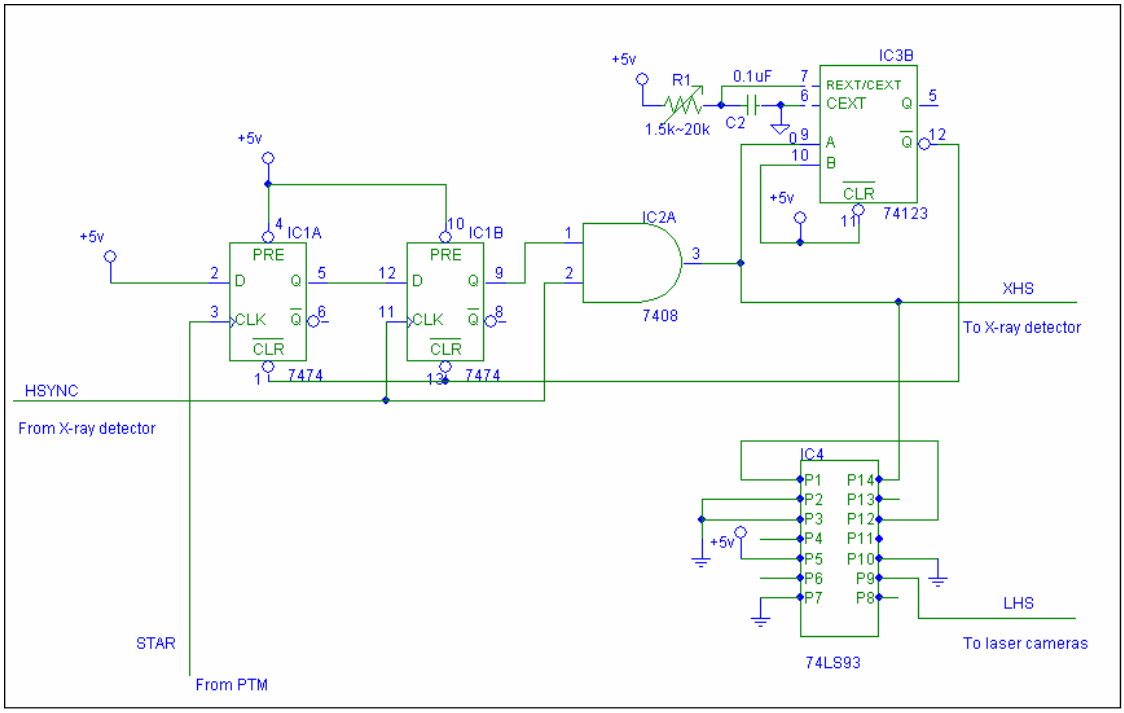


Figure 5.9 Schematics of the pulse sampling module (PSM)

As shown in Figure 5.8, the input signal of the PTM is the encoder pulse, ENCODER, and the output signal is START, which is then relayed to the PSM. The core of the PTM is a PLL IC CD 4046, which traces the ENCODER signal, and generates a stabilized signal, START. As shown in Figure 5.10, a typical PLL consists of a phase detector (PD), a loop filter (LP), a voltage controlled oscillator (VCO) and a loop divider (1/N).

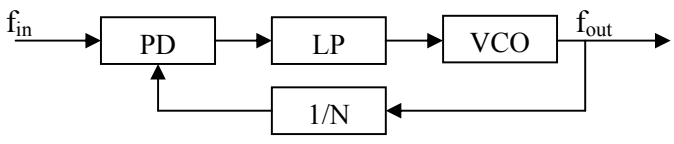


Figure 5.10 Diagram of a phase locked loop (PLL)

If the PLL is working at locked mode, we have

$$f_{out} = N\, f_{in} \qquad (5.9)$$

where

f_{in} and f_{out}: the frequencies of the input (reference) and output signal,

N: a preset positive integer.

The PLLs are a class of circuits widely used in fields including communication, instrumentation, and telemetry, for signal recovery, modulation/demodulation, and signal conditioning. For more details on principles and applications of the PLL, please refer to literature (Wolaver, 1991). In our application, the PLL was used to generate the START signal, which has the same average frequency as the ENCODER signal, but has a more stable phase than ENCODER. Refer to Figures 5.8 and 5.12, the PD and the VCO were included in the chip CD4046, the LP is implemented with a μA741 operational amplifier, and the pulse divider (1/N) was realized by a 74LS93 IC.

As shown in Figure 5.9, the START signal generated by the PTM was used as an input for the PSM. The other input signal of the PSM was the HSYNC signal from the X-ray detector, which in a sense was used to "sample" the START signal and generate a new version of horizontal synchronization signal, XHS, for X-ray image acquisition. The design ensured that, for each positive pulse of the START, there would be one and only one XHS pulse generated. It was also ensured that the XHS pulse coincides with the first HSYNC pulse that follows the rising edge of START pulse.

Refer to Figure 5.9, the working mechanism of the PSM can be described as follows. Without losing generality, assume[1] that initially the states of the two D flip-flops IC1A and IC1B were both 0, meaning the Q port of IC1A and IC1B were both at low voltage level. According to the property of the 74LS74 flip-flop, the state of IC1A will

[1] It can be shown that the function of PSM is not affected even if this assumption on initial circuit state does not hold.

change from 0 to 1 if and only if there is a rising edge at its CLK port which is driven by the START signal. Once a START rising edge comes, the state of Q port of IC1A will change from 0 to 1, and will remain at 1 level for now. If a positive pulse of HSYNC arrives at the CLK port of IC1B, the state of IC1B will change from 0 to 1, because the rising edge at its CLK port samples a 1 at its D port. Thus the 1 at port Q of IC1B will be ANDed with the positive pulse of HSYNC by IC2A, an AND gate, yielding a new positive pulse, XHS. The width of the XHS pulse would be the same as HSYNC[2]. At the end of the XHS pulse, its falling edge will trigger the one-shot oscillator IC3B, generating a negative pulse at the $\overline{Q}$ port of IC3B. As shown in Figure 5.9, the negative pulse will be relayed to the $\overline{CLR}$ ports of IC1A and IC1B, thus will clear both of the D flip-flops (i.e., force the Q ports to change from 1 to 0). After IC1A and IC1B are cleared, the PSM goes back to its initial state and becomes ready for another round of operations, which will be triggered by a new pulse of the START signal. In this way, it was ensured that there was one and only one XHS pulse generated for each START pulse, thus the location information was retained. On the other hand, because the width of the XHS pulse is constant regardless of speed of the conveyor, it is also guaranteed that the photon integration time for each line in X-ray image is constant, thus the problem of X-ray image artifact is solved.

[2] Strictly speaking, the width of XHS is narrower than HSYNC. However the difference is less than 100 ns (Texas Instruments, 1988), which is negligible comparing with the width of XHS or HSYNC, which is at least 1ms.

5.4 Integration of X-ray and laser imaging subsystems

Once the X-ray and laser imaging subsystems are ready and capable of yielding X-ray and laser 3D images with acceptable quality and throughput, the next major task is to integrate them into a coherent and efficient system. In this section, various aspects of the integration task are discussed, including:

1. Image registration, which determines the geometric relationship between the image obtained by the X-ray subsystem and the one obtained by the laser subsystem;
2. Image modeling, which generates a model describing the relationship between grayscales in an X-ray image and those in a laser 3D image, and yields a pseudo X-ray image based on the model and a laser 3D image. The thickness variation in the X-ray image will then be canceled out by using the pseudo X-ray image;
3. Defect detection, which searches for the bone fragment or other types of abnormality in the compensated image; and
4. Multithread framework, which makes it possible to finish the image processing tasks in real time.

5.4.1 Image registration

In order to perform thickness compensation for the X-ray image, it is crucial to first ensure that the pixels in the laser 3D image are mapped correctly to those in the X-ray image. The spatial relationship between the two images can be described by a geometric transformation, and the type and parameters of the transformation needed to be

determined for every pair of the X-ray and the laser images. In our system, the X-ray source, the X-ray detector, the laser, and the CCD cameras were mounted to a rigid frame, so the geometrical relationship between the components of the two subsystems are fixed. Because the parameters of the geometric transformation will keep constant over time, we can estimate the parameters offline and then apply them for real-time processing.

Affine Transform

According to the setup of the imaging systems, the geometric relationship between the X-ray and the laser images was approximated by the following affine transformation (Forsyth and Ponce, 2003),

$$\begin{bmatrix} x \\ y \end{bmatrix} = \mathbf{A} \times \begin{bmatrix} u \\ v \end{bmatrix} + \mathbf{B}$$

$$= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \times \begin{bmatrix} u \\ v \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \qquad (5.10)$$

where (x,y) and (u,v) are the coordinates of a matched pixel in the X-ray image and laser image respectively. The matrix A and vector B define the affine transformation. In the general affine model of equation (5.10), operations including translation, scaling and rotation are described. In our system, the X-ray and laser imaging devices were aligned to each other, thus only translation and scaling were considered in this registration stage, and equation (5.10) is reduced to

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} \times \begin{bmatrix} u \\ v \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \qquad (5.11)$$

Furthermore, the parameter a_{22}, the ratio of trigger frequencies of the X-ray and the laser imaging subsystems, was predetermined based on the setting of the trigger electronics

(refer to section 5.3.3). According to Figure 5.9, a_{22} is known to be set to four[3]. Thus, in the model of equation (5.11), there are only three unknowns, a_{11}, b_1 and b_2, which can be solved if given two sets of (x,y,u,v).

Linear Least Squares Method

In order to solve the three unknowns in equation (5.11), at least two pairs of matched marker pixels from the X-ray and the laser images should be selected. In practice, six pairs of the matched anchor points are used, and the affine transformation parameters were estimated using a linear least squares method for a best-fit (Forsyth and Ponce, 2003).

Bilinear interpolation

Once the affine transformation parameters are estimated, they are applied to the original laser image according to equation 5.11 to obtain the registered version of the laser image, which is used to cancel thickness variations in the X-ray image.

The scaling operation will cause "missing points" in the transformed image. A bilinear interpolation method was used to estimate the grayscale of the missing pixels by taking an average of the known values at neighboring points. The bilinear interpolation takes the weighted average of a 2x2 pixel neighborhood as the assigned value to evaluate the interpolated pixel. Weights were determined by measuring the distance from the interpolated pixel to its nearest four surrounding pixels. The value of the interpolated pixel P in (Figure 5.11) can be evaluated as follows,

[3] According to the specification of 74LS93 IC, the frequency at pin 9 is one fourth of that at pin 14.

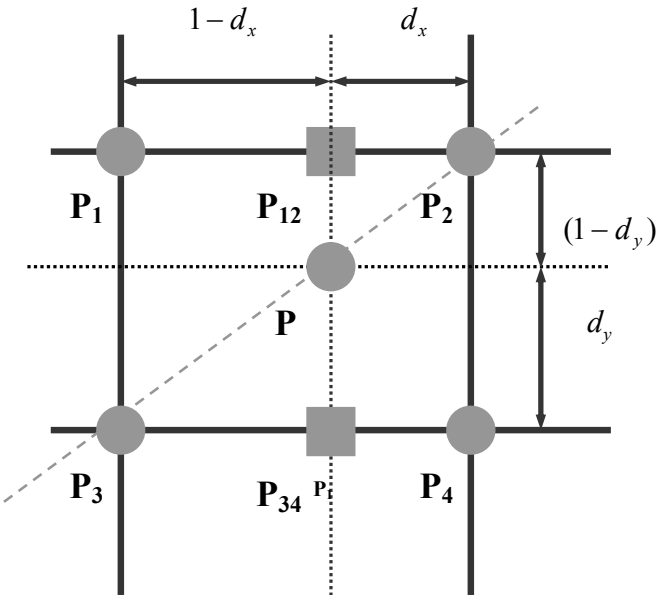


Figure 5.11 Schematic representation of linear interpolation

$$p_{12} = d_x p_1 + (1-d_x) p_2$$

$$p_{34} = d_x p_3 + (1-d_x) p_4$$

$$p = d_y p_{12} + (1-d_y) p_{34}$$

$$= d_x d_y p_1 + (1-d_x) d_y p_2 + d_x (1-d_y) p_3 + (1-d_x)(1-d_y) p_4 \qquad (5.12)$$

where p_1, p_2, p_3 and p_4 represent the pixel values from a 2x2 pixel neighborhood of the interpolated pixel P. p is the value of P. p_{12} and p_{34} are the intermediate pixel values used to derive the value of p.

5.4.2 Image modeling and thickness compensation

In the X-ray image, the grayscale of each pixel is determined by the intensity of X-ray transmission to the specific pixel, which is further determined by the total X-ray attenuation along the path of transmission. Assume that the X-ray attenuation due to free

air and conveyor belt are negligible[4], and further assume that the material of chicken meat is homogenous, then the X-ray intensity at pixel (x,y) will be

$$I(x,y)=I_0\exp[-\mu\, d(x,y)] \qquad (5.13)$$

where

I_0: the incident X-ray intensity,

μ :the attenuation coefficient of the meat tissue, and

d(x,y): the length of X-ray path inside the meat.

The model can be used to cancel out the X-ray image grayscale variation caused by thickness variations. According to equation (5.13), the transmitted intensity I(x,y) can be modeled based on the characteristics of the material, however, in practice, it is not feasible to evaluate the attenuation coefficient for each possible composition of chicken meat and X-ray energy setting. In our study, an image based approach was used to construct the model.

As mentioned before, for each sample on the conveyor belt, the laser range imaging subsystem yields a 3D thickness image T(x,y). For each pixel in the thickness image, its grayscale is proportional to the thickness of the corresponding spot of the object (i.e., either the meat sample or the conveyor belt). In other words, let the thickness profile of the object be d(x,y), then there exists a positive constant k, which satisfies:

$$T(x,y)=k\, d(x,y) \qquad (5.14)$$

In a similar manner, the sample is scanned by the digital X-ray detector, which yields I(x,y), an X-ray snapshot of the object. A thickness compensated image can be obtained as:

[4] The conveyor belt used in this application is a food grade conveyor belt with no internal texture and minimum X-ray absorption.

$$I_{tc}(x,y)=I_{px}(x,y)-I(x,y)+C(x,y) \quad (5.15)$$

where

$C(x,y)$: a constant image that keeps $I_{tc}(x,y)$ positive, and

$I_{px}(x,y)$: the pseudo X-ray image defined as (similar to equation 5.15).

$$I_{px}(x,y)=I_0\exp[-\mu\, d(x,y)] \quad (5.16)$$

where

I_0 corresponds to the original intensity of X-ray beam, which is determined by the X-ray system; and

μ : the effective X-ray absorption coefficient of the sample, which is determined by the X-ray energy and the material of the sample. Putting everything together, the compensated image $X_{tc}(x,y)$ can be calculated as:

$$I_{tc}(x,y)= I_0\exp[-\frac{\mu}{k}\bullet T(x,y)] -I(x,y)+C(x,y) \quad (5.17)$$

Ideally, if there is no foreign material in the sample, the compensated image I_{tc} will be flat; otherwise the pixels corresponding to the foreign material "stand out", which can be segmented with a simple threshold operation.

In equation (5.17), $I(x,y)$ and $T(x,y)$ were acquired in real time, and I_0 and $C(x,y)$ can be predetermined. The difference between $I(x,y)$ and $\{I_0\exp[-\frac{\mu}{k}\bullet T(x,y)]\}$ should be small enough and thickness independent, otherwise the resultant thickness compensated image $I_{tc}(x,y)$ will suffer significant noise. Theoretically the ratio $\frac{\mu}{k}$ can be determined beforehand by exhaustive experiments. However, it is hardly practical to generate a data set for the ratio, which could cover each case of the combinations of various X-ray

settings and materials. It is desirable to bypass the complex process of finding the ratio $\frac{\mu}{k}$.

Revisiting equation (5.17), the operation can be viewed as canceling I(x,y) with a function of T(x,y), and in a sense, the smaller the residual of the cancellation the better. Now the problem is reduced to finding a minimum error estimate of I(x,y) based on T(x,y), which can be readily solved by the following least squares approach.

Suppose that the X-ray image I(x,y) and the laser image T(x,y) are registered perfectly to each other, i.e., there is no displacement, rotation or scaling issue. Under this assumption, {I(x,y)} and {T(x,y)} can be treated as one dimensional vectors of the same size, and the thickness compensated equation (5.17) can be rewritten as

$$I_{tc}(n)= f(T(n))-I(n), \quad n=1,\ldots,\max(x)*\max(y) \tag{5.18}$$

A handy approach to determine the function f() is the least squares polynomial estimation. Assume a k^{th} order polynomial estimate is desired, then the f() can be defined as:

$$u=f(v)= \sum_{i=0}^{k} a_i v^i \tag{5.19}$$

where a_i , i=0,…,k are the fitting coefficients. When substituting I(n) and T(n) for u and v in equation (5.19) respectively, the a_i's can be solved using a standard least squares algorithm. Once the function f() is determined, the pseudo X-ray image $I_{px}(x,y)$ can be calculated as

$$I_{px}(x,y)=f(T(x,y))=\sum_{i=0}^{k} a_i [T(x,y)]^i \tag{5.20}$$

After the pseudo X-ray image $I_{px}(x,y)$ is obtained, the thickness variation in the X-ray image can be canceled out using equation (5.15).

5.4.3 Detection algorithm

In the thickness compensated image, pixels with nonzero grayscale represent either physical contamination or random noise. In practice, the contaminations to be detected in the application are highly diversified, ranging from soft and thin bone fragments to hard and sharp metal clips. As a result, the signatures of the hazardous materials revealed by the synergic X-ray and laser imaging system are also highly diversified, which makes it difficult to develop a universal classifier to distinguish the hazardous items from the normal meat tissue.

In order to detect the highly diversified subjects, a rule-based scheme was adopted in this study, which enables the system to make complicated classification judgments based on a predefined decision tree (Duda, et al., 2000).

As in most pattern recognition applications, prior to the classification stage, we need to define a set of subjects to be classified, and construct a feature vector for each of the subjects. In this system, prior to pattern recognition, thickness compensation was applied to each X-ray image of chicken fillet (as mentioned in the previous section), followed by segmentation of the compensated image, and then the blob analysis.

Automatic threshold selection for image segmentation

In a thickness-compensated chicken meat image, if there is no physical contamination embedded, the intensity distribution of the image pixels would be uniform. When there does exist a foreign material, such as a bone fragment, the intensity levels at the region of the bone area present higher values. Based on these properties, the image

segmentation method is applied to the thickness-compensated image in order to find the contaminants.

In an ideal case, the simplest approach to segment an image is to select a global threshold value, and classify the image into two clusters, as described in equation (5.21),

If $TC(x, y) > T$ then $TC(x, y) = 0$ else $TC(x, y) = 255$ (5.21)

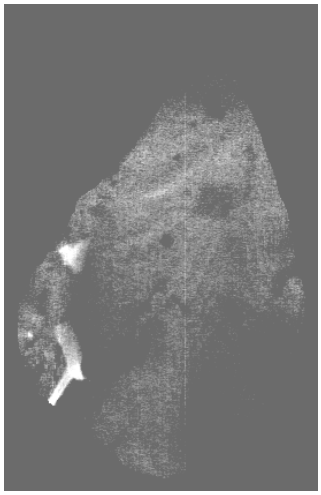
Where,

TC(x,y) represents the intensity value of a pixel at position (x,y) in image TC, and

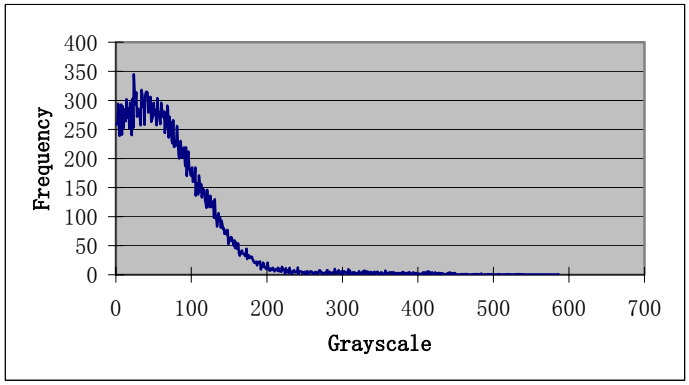
T stands for the threshold of intensity level for clustering the objects.

The threshold value (T) can be selected based on the histogram of the image. If the histogram presents a deep and sharp valley, and the two peaks represent the object and background in the image, the grayscale of the valley in the histogram can be chosen as the threshold. However, in our application, the histogram of the thickness-compensated image rarely presents a sharp valley. As shown in Figure 5.12, In Figure 5.12(b), the histogram of a typical thickness compensated image is flat and broad, imbued with noise, and presents no traceable valley.

In order to make segmentation more robust for this application, it is desirable that the threshold selection can be selected automatically. Three methods have been explored and compared in order to find the best strategy for this application.



(a)



(b)

Figure 5.12 Histogram characteristics of a typical thickness compensated image. (a) The thickness compensated image of a chicken breast (b) The histogram of the image. Note that that zero pixels are omitted in this histogram.

Otsu's method

Otsu's method (Otsu, 1979) is an automatic optimal threshold selection method based on a two-class discriminant analysis. It assumes that the pixel values in an image fit a bi-modal distribution. Under this assumption, the method was designed to select the threshold value by maximizing the between-class variance of the two groups of pixels separated by the thresholding operator.

Consider an image with L gray levels and its gray-level histogram is normalized and regarded as a probability distribution (Otsu, 1979),

$$p_i = n_i / N, \ \ where\,(i = 1 \cdots L), \quad p_i \geq 0, \quad \sum_{i=1}^{L} p_i = 1 \tag{5.22}$$

where,

N is the total number of pixels, and $N = n_1 + n_2 + \cdots + n_L$.

Assuming that we have set the threshold at *T*, and the total pixels in the image will accordingly be partitioned into two groups: the background (assume when no more than T) and the object (when larger than T). The probabilities of the class occurrence are given by:

$$P_{bg} = \sum_{i=1}^{T} p_i, \quad P_{obj} = \sum_{i=T+1}^{L} p_i, \quad where, P_{bg} + P_{obj} = 1 \tag{5.23}$$

where,

P_{bg} : the probability of background occurrence, and

P_{obj} : the probability of object occurrence.

The mean gray-level value of the background (μ_{bg}) and the object pixels (μ_{obj}) will be:

$$\mu_{bg} = \frac{\sum_{i=1}^{T} i\ p_i}{\sum_{i=1}^{T} p_i} = \frac{1}{P_{bg}} \sum_{i=1}^{T} i\ p_i, \quad \mu_{obj} = \frac{\sum_{i=T+1}^{L} i\ p_i}{\sum_{i=T+1}^{L} p_i} = \frac{1}{P_{obj}} \sum_{i=T+1}^{L} i\ p_i \tag{5.24}$$

where,

μ_{bg} : the mean of the background gray-level, and

μ_{obj} : the mean of the object gray-level.

The total mean gray-level value μ_{tot} over the whole image (grand mean) is:

$$\mu_{tot} = \frac{\sum_{i=1}^{L} i\ p_i}{\sum_{i=1}^{L} p_i} = \sum_{i=1}^{L} i\ p_i \tag{5.25}$$

The variance of the background $\sigma^2{}_{bg}$ and the object $\sigma^2{}_{obj}$ will be:

$$\sigma^2{}_{bg} = \frac{\sum_{i=1}^{T} (i-\mu_{bg})^2\ p_i}{\sum_{i=1}^{T} p_i} = \frac{1}{P_{bg}} \sum_{i=1}^{T} (i-\mu_{bg})^2\ p_i, \tag{5.26}$$

$$\sigma^2{}_{obj} = \frac{\sum_{i=T+1}^{L} (i-\mu_{obj})^2\ p_i}{\sum_{i=T+1}^{L} p_i} = \frac{1}{P_{obj}} \sum_{i=T+1}^{L} (i-\mu_{obj})^2\ p_i \tag{5.27}$$

The total variance of the whole image $\sigma^2{}_{tot}$ is,

$$\sigma^2{}_{tot} = \frac{\sum_{i=1}^{L} (i-\mu_{tot})^2\ p_i}{\sum_{i=1}^{L} p_i} = \sum_{i=1}^{L} (i-\mu_{tot})^2\ p_i \tag{5.28}$$

The total variance σ^2_{tot} can be re-written as follows,

$$\sigma_{tot}^2 = \sum_{i=1}^{L}(i-\mu_{tot})^2 p_i = \sum_{i=1}^{T}(i-\mu_{bg})^2 p_i + \sum_{i=T+1}^{L}(i-\mu_{obj})^2 p_i +$$

$$\sum_{i=1}^{T}((\mu_{bg}-\mu_{tot})^2 + 2(i-\mu_{bg})(\mu_{bg}-\mu_{tot}))pi +$$
$$\sum_{i=1+T}^{L}((\mu_{obj}-\mu_{tot})^2 + 2(i-\mu_{obj})(\mu_{obj}-\mu_{tot}))pi$$

$$= P_{bg}\sigma_{bg}^2 + P_{obj}\sigma_{obj}^2 + P_{bg}\mu^2{}_{tot} - 2P_{bg}\mu_{bg}\mu_{tot} + P_{bg}\mu^2{}_{bg} +$$
$$P_{obj}\mu^2{}_{tot} - 2P_{obj}\mu_{obj}\mu_{tot} + P_{bg}\mu^2{}_{obj}$$

$$= P_{bg}\sigma_{bg}^2 + P_{obj}\sigma_{obj}^2 + P_{bg}(\mu_{obj}-\mu_{tot})^2 + P_{obj}(\mu_{obj}-\mu_{tot})^2 \qquad (5.29)$$

Let $\sigma_w^2 = P_{bg}\sigma_{bg}^2 + P_{obj}\sigma_{obj}^2$ be the within-class variance and $\sigma_B^2 = P_{bg}(\mu_{obj}-\mu_{tot})^2 + P_{obj}(\mu_{obj}-\mu_{tot})^2$ be the between-class variance. From equation (5.29), the total variance σ_{tot}^2 can then be expressed as the sum of the within-class variance and between-class variance, i.e, $\sigma_{tot}^2 = \sigma_w^2 + \sigma_B^2$. In Otsu's method, the "goodness" of selecting T as the threshold value is evaluated by calculating the discriminant criteria measurements:

$$\lambda = \sigma_B^2 / \sigma_w^2, \quad or \quad \kappa = \sigma_{tot}^2 / \sigma_w^2, \quad or \quad \eta = \sigma_B^2 / \sigma_{tot}^2 \qquad (5.30)$$

The optimal T value is selected when λ or κ or η reaches their maximum value.

Otsu's method doesn't work well when the total pixel numbers in both classes differ dramatically, which is the case of most samples in this application. In most cases, the variances of the two classes are different. As a result, the unbalanced data inputs contribute unequal weights to the within and between variance calculations. Since the discriminant criteria are sensitive to the class data distribution homogeneity, the unbalanced data distribution will bias the optimal result.

Prior knowledge based method

Embedded bone fragments are relatively small compared with the whole chicken meat sample. In most cases bone fragments occupy no more than 5% of the total area of the chicken meat sample. Bone fragments in a thickness-compensated chicken breast sample image are bright objects against the dark background. Therefore, a prior knowledge based method can be used for automatic selection of the threshold. For instance, if the bones are known to present the brightest pixels in the compensated images, based on the histogram distribution, an appropriate threshold can be chosen by partitioning the histogram to 95% over 5%.

This straightforward method works well and fast when the bone fragments are homogenous. However, the major drawback of the method is obvious. If the bone fragments present different brightness patterns, this method may miss some of the darker bone fragments in the compensated image.

Localized auto-threshold method

This method is to partition the image into MxN sub-images, and then calculate the threshold values for each sub-image. The threshold of a sub-image is called its sub-threshold. The sub-threshold values are partitioned into two groups according to their magnitudes, and the global threshold is chosen to be the smallest value of the group that contains the higher threshold values.

For each thickness-compensated chicken meat sample image, nine sub-images were generated. Otsu's method was applied to each sub-image, and both the sub-

threshold value and the discriminant criteria measurement λ were calculated (see equation 5.30). If a sub-image contains bright bone fragments, both the sub-threshold and sub λ values are high. On the other hand, if a sub-image doesn't have any bone fragment, both the sub-threshold and sub λ values are very low. If in a thickness-compensated chicken meat sample image the bone fragments have different brightness levels, sub-images that contain the darker bone fragments have relative lower sub-threshold values compared with the ones that contain brighter bones. In those sub-images, however, the auto-threshold values and sub λ values are still high enough to be distinguished from the sub-images that don't contain any bone fragments. A detailed discussion is presented in the next chapter.

Blob generation

When the global threshold value is automatically selected, a thresholding operation, as shown in equation (5.31), is applied to partition the thickness-compensated image into two groups, and a binary image is generated as

$$BIN(x,y) = \begin{cases} 1 & TC(x,y) > T_g \\ 0 & others; \end{cases} \qquad (5.31)$$

where

TC(x,y): the thickness-compensated image, and

BIN (x,y): for the binary image, and

T_g : the selected global threshold value.

The pixels that represent bone fragments are segmented and assigned value "1" in the binary image BIN(x,y), so are those representing the noise. Blobs are then formed by

grouping the adjacent “1” pixels. In order to eliminate the blobs that are caused by noise, morphological operations such as the erode filter and the dilate filter are used.

The morphological filters are Boolean filters that are applied to a binary image. The most common Boolean operations are AND and OR. The binary erode filter is defined by

$$E(x,y) = \text{AND}\ [\ BIN(x,y)] \tag{5.32}$$

where,

E(x,y): the resultant eroded image, and

BIN(x,y): the binary image defined in equation (5.31).

Similarly, the dilate filter is defined by OR operation applied to the binary image BIN(x,y),

$$D(x,y) = \text{OR}\ [BIN(x,y)] \tag{5.33}$$

where,

D(x,y): the resultant dilated image.

The erode filter shrinks the size of the object blobs. Usually, the isolated object blob with very small sizes will be eliminated by applying this operation. Therefore, small dot-like noise caused by auto-thresholding can be removed.

The morphological operations are simple, yet fast and powerful tools suitable for machine vision online applications. In this application, the erode filter is applied first to BIN(x,y) to obtain erosion image E(x,y) and to eliminate many small noise related blobs. Then, a dilate filter is applied to E(x,y) to expand the remaining blobs (if there is any) close to their original sizes. The D’(x,y) image is generated accordingly,

$$D'(x,y) = \text{OR}\ [\ \text{AND}\ [BIN(x,y)]\ ] \tag{5.34}$$

By applying the erode and the dilate filters sequentially to the binary image BINS(x,y) we can effectively eliminate the small noise blobs, and smooth the boundaries of other blobs as well. Since the two operations are not inverse operations of one another, the resultant D'(x,y) is different from B(x,y). However, since the operations are performed within the same size of binary windows, to blobs that are big enough to remain in D'(x,y), the changes of their sizes and boundaries are trivial, and won't affect the following operations. Beside the bone fragment blobs, there are still relatively "large" noise blob left in D'(x,y). In order to reduce the noise effects to identify only the bone fragments, further feature extraction methods are applied.

Blob analysis

Once the blobs are generated, useful information is extracted to help the machine vision system to "recognize" each blob by analyzing each of them. In our application, the useful basic blob information includes size (area), mass center, and position range. Also, the binary blob image D'(x,y) served as the mask to original thickness-compensated image TC(x,y), and some gray level image properties are obtained. The blobs are first labeled to identify them from each other, and then processed with blob analysis operations. Blob size is defined by

$$Bsize(b(n)) = Bsize(b(n))+1, \quad \forall\, D'(x,y) = '1',\ (x,y) \in b(n) \qquad (5.35)$$

Where,

b(n) : the blob labeled pixel group.

Blob mass center is calculated by moments. The X position of the center of gravity of a blob is defined by

$$X(b(n) = \text{MOMENT\_X1\_Y0}(b(n)) / \text{Bsize}(b(n)), \quad \forall D'(x,y) = '1', (x,y) \in b(n) \tag{5.36}$$

Similarly, the Y position of the center of gravity of a blob is defined by

$$Y(b(n) = \text{MOMENT\_X0\_Y1}(b(n)) / \text{Bsize}(b(n)), \quad \forall D'(x,y) = '1', (x,y) \in b(n) \tag{5.37}$$

In both equations (5.36),(5.37), MOMENT_X1_Y0(b(n)) and MOMENT_X0_Y1(b(n)) are syntaxes of moments of a blob labeled as b(n). MOMENT_Xn_Ym (b(n) are defined as:

$$MOMENT\_Xn\_Ym(b(n)) = \sum_{i=1}^{T(b(n))} x_i^n y_i^m p_i \tag{5.38}$$

where,

T(b(n)): the total pixel number within blob b(n), and

x_i : X coordinate of pixel i,

y_i : the Y coordinate of pixel i, and

p_i : the pixel value (in binary image p_i equals zero or one) of pixel i.

All the coordinates used are relative to the image origin (top-left corner). The blob position range is defined by the minimum and the maximum values of its X and Y coordinates.

To use blob image D'(x,y) as a mask applied to thickness-compensated image TC(x,y) is defined by

$$TC'(x,y) = \begin{cases} TC(x,y) & D'(x,y) = 1 \\ 0 & others; \end{cases} \tag{5.39}$$

where,

$TC'(x, y)$ is the resultant gray level blob image.

For each labeled gray level blob, mean (average grayscale) and standard deviation are calculated.

Based on the basic calculation of these characteristics, additional information can be derived. For instance, the distance between any two blobs is obtained by calculating the Euclidean distance between blob mass centers. All the direct and indirect characteristics of each blob will be used to formulate a set of feature vectors for further feature extraction.

Feature Extraction

The composition of the feature vector was flexible, as well as its organization. A typical composition of the feature vector is shown in Figure 5.13. Note that the parameters are adjustable and only limited information is covered in this example. In practice, our system uses a feature vector consisting of more detailed characteristics of a blob, such as shape, location, orientation, distance to other blobs, grayscale in the original X-ray image, etc.

AG: Average Grayscale	SD: Standard Deviation	A: Area

Figure 5.13 The feature vector of a blob

Classifier design

Classifier is used in many pattern recognition applications to sort objects into different groups. In this application, the purpose was to detect the hazardous objects from the chicken meat, and all hazardous materials should be removed from the meat. Therefore, to classify the hazardous objects doesn't seem critically important. However,

in practice, in order to analyze the detection abilities of the system to handle various types of hazardous objects, as well as to further eliminate noise, classifier design is necessary.

In the computer vision field, many classification techniques have been further explored for different applications. Those methods are categorized into supervised or unsupervised, parameter and non-parameter estimations. For this online application, parameters of the classification model should be minimized. The supervised training approach, which has been found quite robust in the machine vision environment, should be used with prior knowledge of the detection problem. In light of these requirements, the classifier in this application should be simple and fast. In this machine vision application, some elements in the feature vector present logical relationships which can be described using certain rules, and are very helpful in sorting. The rule-based method implemented using a decision tree, which has been widely used in other online machine vision systems, has been found effective in this application. An example of the decision tree structure is shown in Figure 5.14.

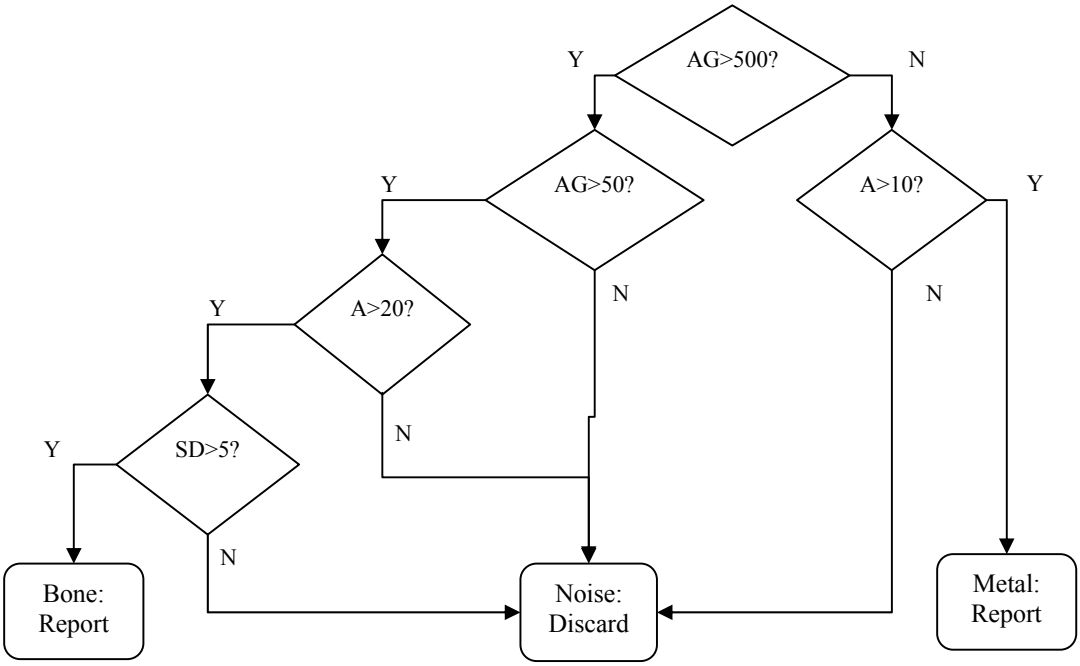


Figure 5.14 Decision tree for a rule-based classifier.

Overall algorithm

The overall procedure of intrusion detection is shown in Figure 5.15. For each meat sample, a pseudo X-ray image was formulated based on the true X-ray image and the laser image, and the compensated image was the difference between the pseudo and the true X-ray images. Blobs in the compensated image were analyzed and the features of each of the blobs were fed to the rule-based classifier, where each of the blobs was classified as either normal (meat tissue) or abnormal (bone, metal, etc.) based on a set of predefined rules.

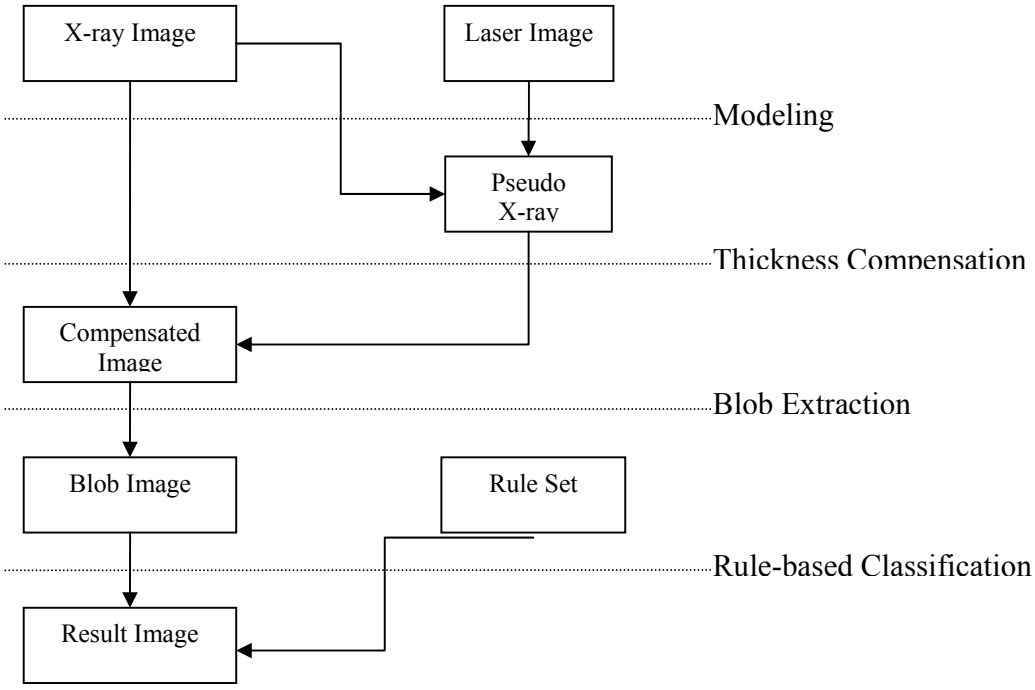


Figure 5.15 Algorithm block diagram of the synergic inspection system.

Once bone fragments or other types of hazardous objects were found, the system would record the locations of those defects along with the location of the meat. The location information can be reported to a computer screen for human intervention and/or can be forwarded electrically for product tracking and handling.

5.4.4 Multithread framework

Besides accuracy, real-time capability of a detection system is also vital for industrial applications. As mentioned above, tasks to be carried out by the imaging system involve complex operations, including synchronizing X-ray and laser subsystems,

grabbing and formulating the X-ray and laser range images, compensating the X-ray image using the laser range image, identifying subjects in the combined image, and presenting the results via an appropriate user interface. If the tasks were conducted in a traditional sequential manner for meat streams on a conveyer, the system must finish all the imaging and analysis processes for one object before another object can be started. If the processing time for an object is long, there could be a substantial waiting time for each of the objects, which leads to undesirable low throughput and inefficient use of computation power. A much better approach is to conduct the tasks in a parallel manner, i.e., the system handles multiple objects simultaneously, and thus the process stage of each object could be different. In this study, a parallel processing framework using a multithread approach was developed, which enabled concurrent system control and image analysis, and maximized the use of imaging hardware, host CPU and I/O channel.

Single thread vs. multithread

Figure 5.16 Behavior model of a single thread real-time image processing system

In a real-time image-processing environment, images are continuously presented to the processor frame by frame endlessly, as shown in Figure 5.16, where *processor* refers to a set of hardware and software dedicated to image processing. The processor needs to respond to the input image flow and make decisions accordingly, in a timely manner. The crucial time constraint can be represented as

$$T_f > T_p \tag{5.40}$$

where,

T_f: the time interval between the start of the current frame and the start of its successor, and

T_p: the processing time for the current frame.

The processor must finish the work on one frame before the next frame begins. Otherwise the new incoming frame will either be skipped by the processor or be held until the processor is free; both of which are undesirable scenarios. The disadvantage of skipping means a loss of information and the hold-and-wait approach can lead to an accumulating delay between input and output. In this case, a sizeable "waiting room" buffer is needed to hold all of the waiting frames, which leads to another difficulty, i.e., determining a sufficient and efficient buffer size.

If the workload of image processing per image frame is not heavy, or the data rate of the input image flow is low, the constraint of Equation (5.40) would pose little or no real restriction, depending on the specific application. However, in our case, since the workload was heavy due to the combined X-ray and laser imaging algorithms, and the input data rate was high, a multithread model was designed, as shown in Figure 5.17, in order to ensure timely response.

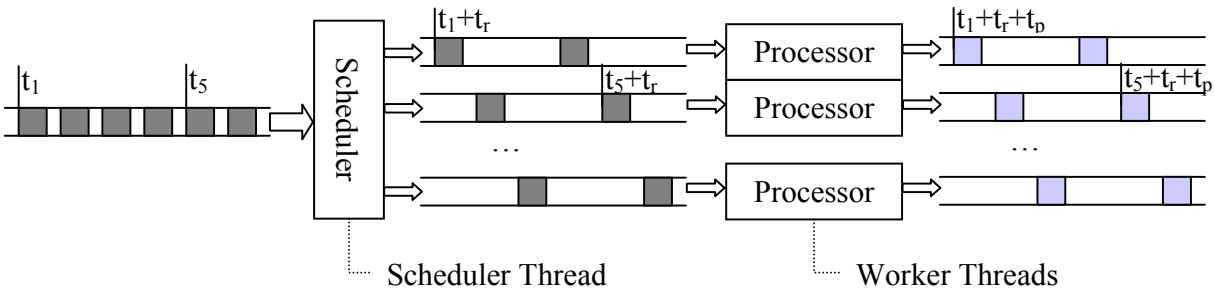


Figure 5.17 Behavior model of a multithread real-time image processing system

In this multithread model, the tasks are shared by different units (or threads), i.e., a scheduler thread and a set of worker threads. When a new image frame arrives, the scheduler will respond first if it is ready, pass the job to a *free* worker thread, and then get ready for the next frame. When a worker thread is called by the scheduler, it will enter the *busy* state and process the assigned image frame. After the frame is finished, the worker will feed the processed result back to the scheduler and return to the *free* state. The time constraint is now

$$T_f > T_r \tag{5.41}$$

where

T_f: the time interval between the start of the frame being processed and the start of its successor, and

T_r: the scheduler response time for the current frame, which is far less then T_p.

The advantage of this multithread model is that the scheduler thread can always respond in time to the incoming image frames as long as there are sufficient free worker threads. Certainly this space-for-time approach will not guarantee to suffice an infinite

throughput requirement; however, comparing with the single thread approach, it promotes real-time capability and system resource utilization significantly.

Implementation of Multithread Scheme

Basic building blocks of the multithread scheme are threads and messages (also known as events). A *thread* is a path of execution within a process (i.e., application, program, etc.)(Silberschatz, et al., 2003). Various threads of a process share the same resources of the process, while retaining certain operational independence of their own. A *message* (or event) is a carrier of information used for thread communication and synchronization. The meaning of a message can be flexible. Depending on its specific definition, a message can be used to declare that a scheduler is ready for another image frame, notify a worker thread to start or stop working, or declare that a worker is free or busy, etc.

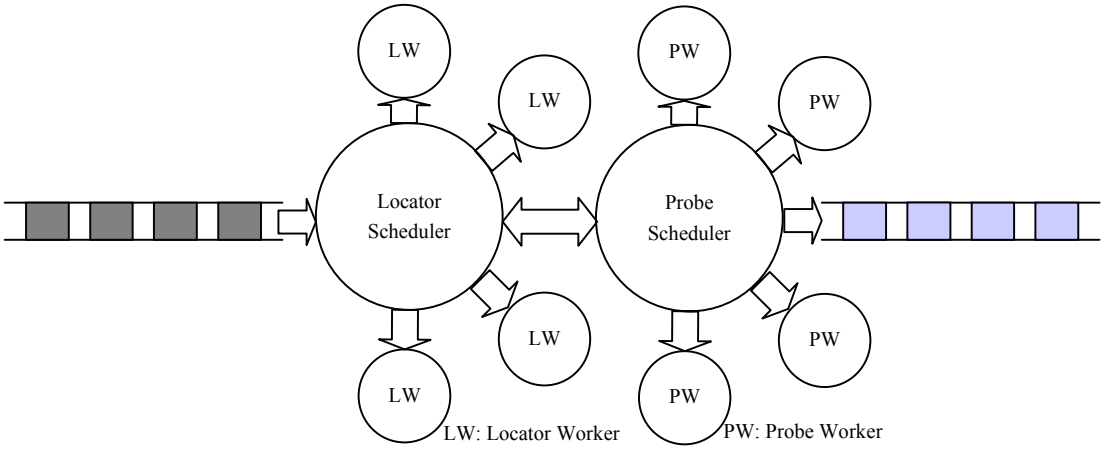
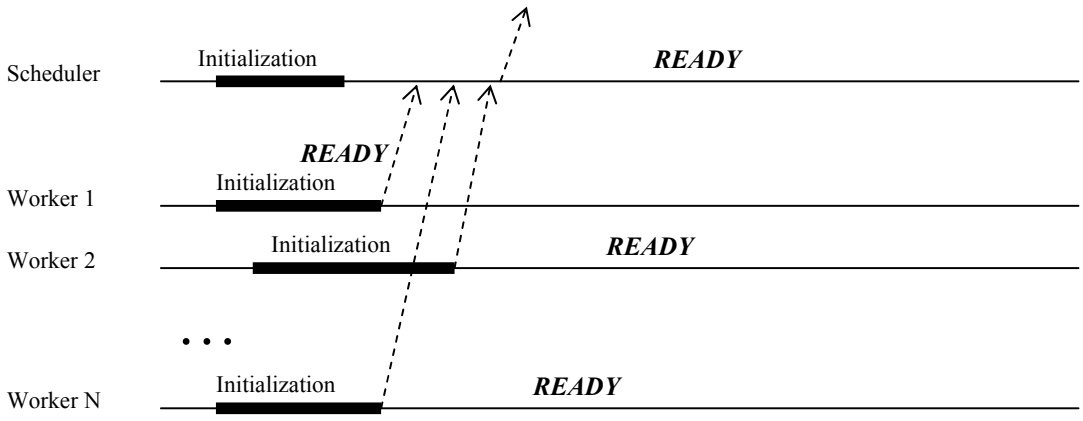


Figure 5.18 Topology diagram of the multithread system

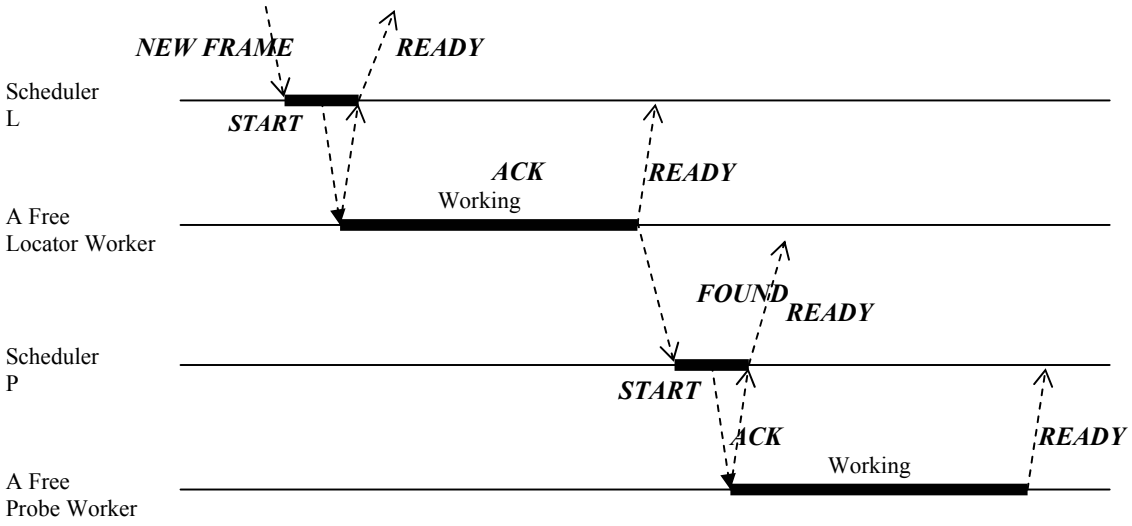
In our multithread scheme, the threads were organized in a nested manner. As shown in Figure 5.18, the threads are grouped into two sets, namely a *locator* group and

probe group, each of which consists of one scheduler thread and multiple worker threads. For the worker threads in both groups, the scheduler thread serves as an interface passing the input and output data, and as a coordinator managing workloads. The function of the locator group is to determine the location of each object, i.e. chicken fillet, in terms of the X-Y coordinates and dimensions. The probe group is set to search for physical hazards in each object passed by locators, based on the information embedded in the X-ray and laser range images. Note that besides the threads mentioned here, there are additional threads responsible for tasks such as image grabbing, result display, and user interface, etc., which are less technically relevant.

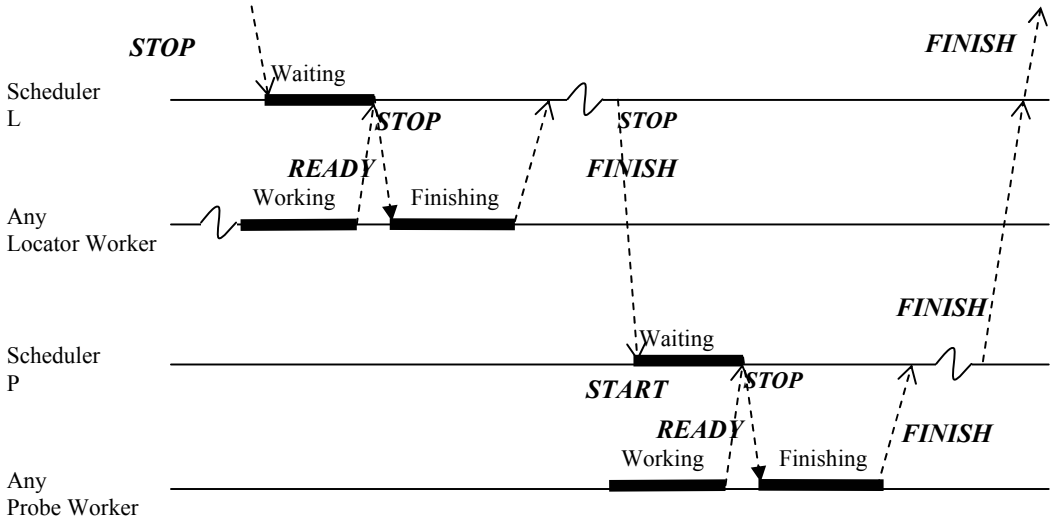
The control mechanism of the multithread system can be described as the following three phases, as shown in Figure 5.19. For convenience, we number the scheduler in the locator group as Scheduler L, and that in the probe group as Scheduler P.



(a) Initialization phase



(b) Processing phase



(c) Termination phase

Figure 5.19 Message flow of the multithread system at different working phase

Initialization phase, as shown in Figure 5.19(a):

1. Each worker thread finishes its own initialization and sends a READY message to its scheduler;
2. Scheduler thread signals READY to controller when all worker threads are ready.

Processing phase, as shown in Figure 5.19 (b):

For locator group

1. When a new image frame comes, scheduler L immediately selects a free locator worker, prepares the private working buffer for the worker and sends a START message to it. Then the selected locator worker starts working on the image frame, and scheduler L becomes ready for the next incoming frame once it is confirmed that the locator worker has started working.
2. Once a locator finishes working on an image frame, it will register the result information, send a READY message to scheduler L, and start waiting for the next job. If a new object is found, the locator thread will send a FOUND message to the scheduler in the probe group (scheduler P) before taking a break.

For probe group

1. When a FOUND message is received by Scheduler P, it immediately selects a free probe worker and prepares the private working buffer for the worker and send a START message to it. The selected probe worker will

start working on the object, while scheduler P becomes ready for the next object.

2. Similar to the locator case, once a probe worker finished working on an object, it will register the resulting information, and send a READY message to Scheduler P. If the probe finds one or more physical hazardous items, such as a bone fragment, or a metal clip, etc., it will report the existence and location of each detected item.

Termination phase, as shown in Figure 5.19 (c):

1. User sends STOP message to the scheduler of the locator group (i.e., scheduler L).
2. Scheduler L sends STOP message to each of the locator workers if the worker is ready (i.e., has no job in hand)
3. Scheduler L sends STOP message to scheduler P after all the locator workers are stopped.
4. Scheduler P sends STOP message to each of the probe workers if the worker is ready.
5. Scheduler P sends FINISH message to scheduler L after all the probe workers are stopped, and scheduler L will then send FINISH message to controller.

As demonstrated in our application, with appropriate control and synchronization mechanism in place, a multithread system can work reliably and achieve much higher efficiency than a single thread system.

CHAPTER 6

RESULTS AND DISCUSSION

In this study, a combined machine vision system with both X-ray imaging and laser 3D imaging was developed and on-line tested for deboned poultry product inspection. As preliminary studies, single energy X-ray and dual energy X-ray imaging were also explored. In this chapter, the experimental results of different methods are presented and can be categorized into three result sections,

1. X-ray imaging with single energy
2. X-ray imaging with dual energy, and
3. Combined X-ray and laser 3D imaging

6.1 X-ray inspection with single energy

In order to study the impact of the X-ray energy level on image quality, a series of images were taken at different energies of a meat sample with a bone fragment. Because the image quality was unacceptable if the energy was lower than 22 keV, and the X-ray generator in our imaging system cannot provide an energy level over 50 keV, the X-ray energy levels ranging from 22 keV to 50 keV, with a one keV step, were considered in this image quality test.

As mentioned previously, the image quality was evaluated using the CQI, a combined quantitative index composed of image contrast and SNR information. The experimental results are summarized in Figures 6.1, 6.2 and 6.3 for the 29 energy levels.

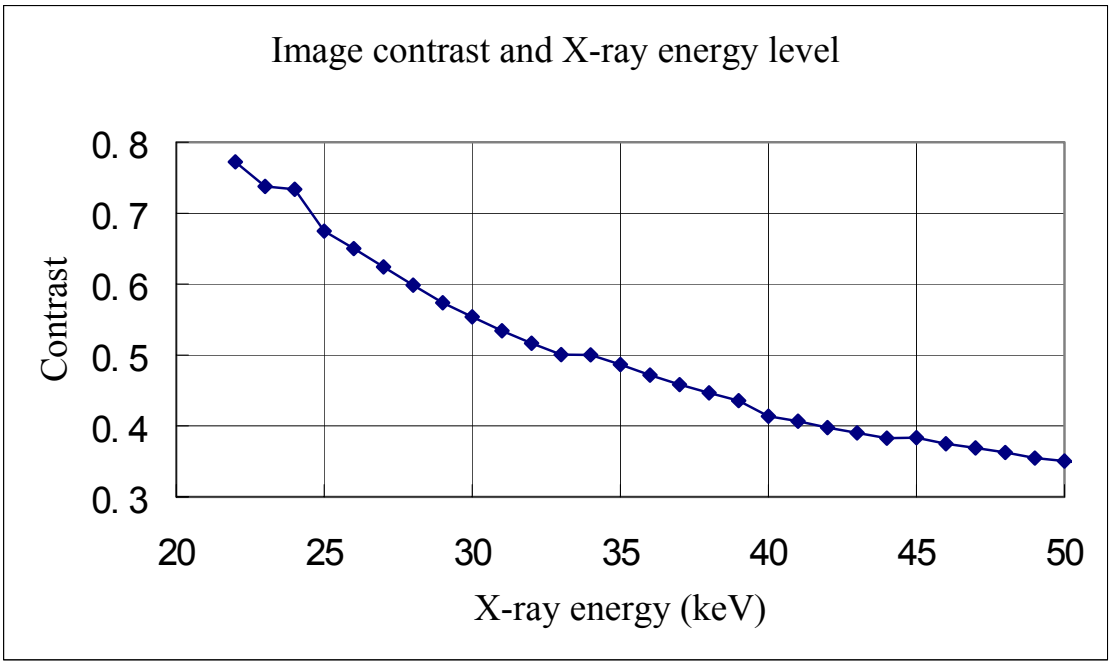


Figure 6.1 Relation between image contrast and X-ray energy level

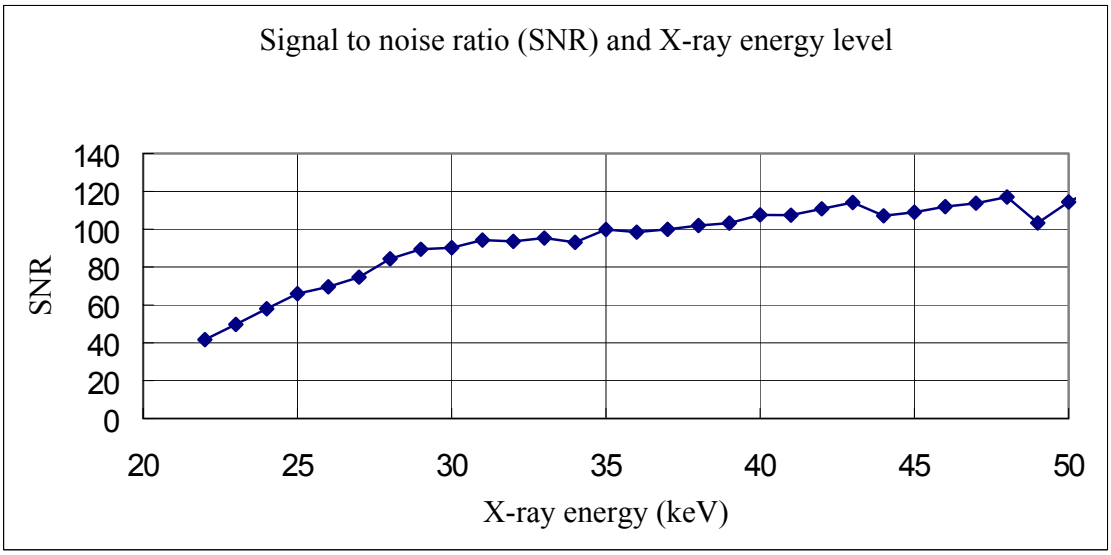


Figure 6.2 Relation between SNR and X-ray energy level

As shown in Figure 6.1, the image contrast between the bone fragment and the surrounding meat tissue is at its highest level when the X-ray energy is 22 keV, the

lowest energy, and decreases as the X-ray energy level reaches 50 keV. In Figure 6.2, as the X-ray energy increases, the SNR also increases from 41.97 at 22 keV to 114.39 at 50 keV.

Despite a few exceptions in both curves, one can conclude that as X-ray energy increases, image contrast decreases while SNR increases monotonically.

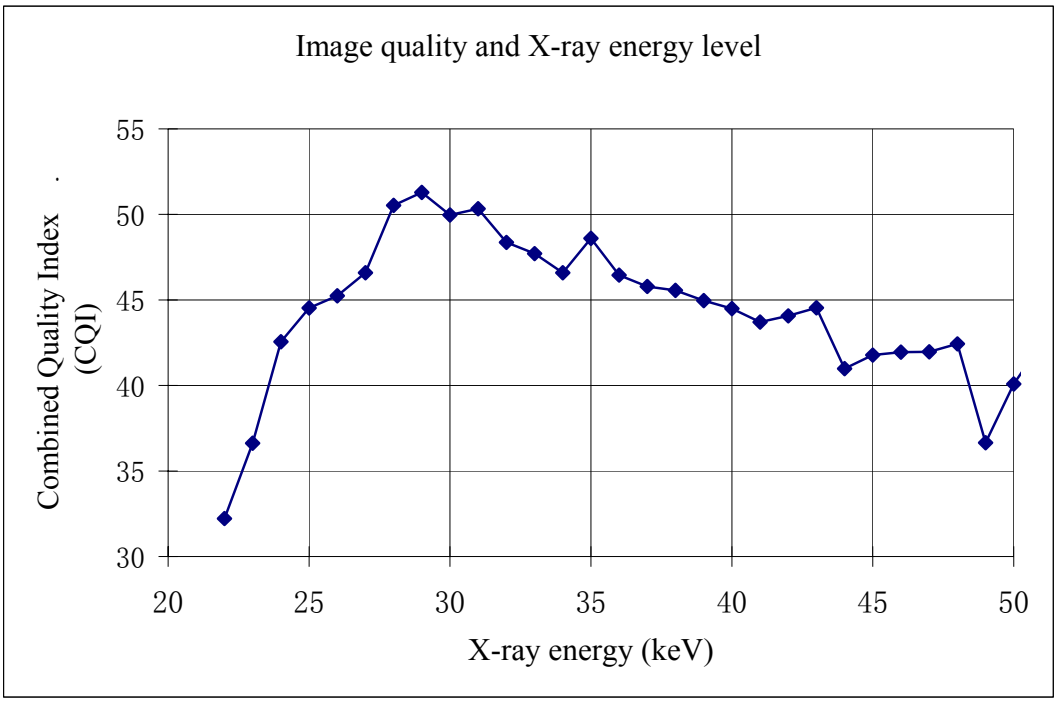


Figure 6.3 Relation between image quality and X-ray energy level, where the image quality is evaluated using the combined quality index

.

The overall image quality measured by the CQI is shown in Figure 6.3, where the peak of the CQI curve resides between 28 keV and 31 keV. This finding implies that a valuable tradeoff between image contrast and SNR can be achieved in this energy range.

For instance, the images of a chicken breast at 30 keV and 50 keV are shown in Figures 6.4(a) and (b) respectively, and the contrast in the 30 keV image is better than the 50 keV image. In this case, using 30 keV to scan the meat sample makes it easier to differentiate the bone fragment from the meat tissue.

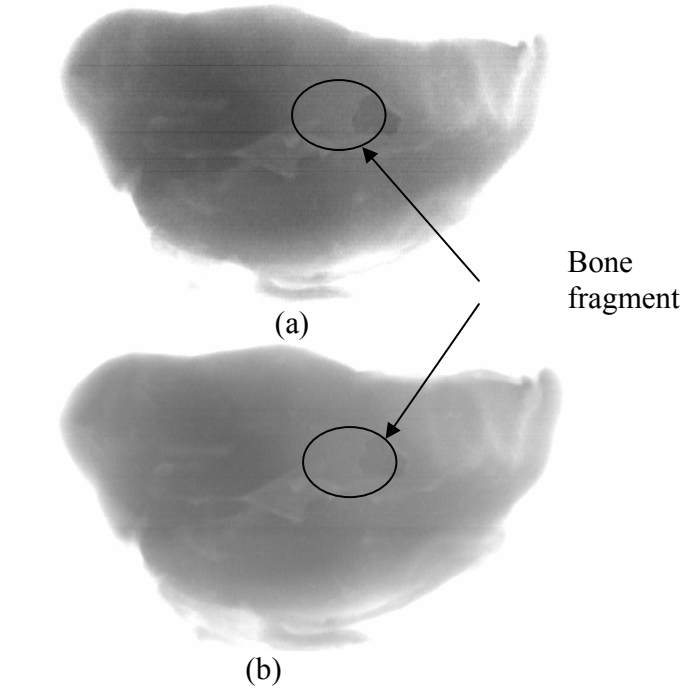


Figure 6.4 X-ray images of a chicken breast with a bone fragment at different X-ray energies. (a) Image captured at 30keV. (b) Image captured at 50keV.

6.2 X-ray inspection with dual energy

The potential of using on-line dual energy X-ray imaging for deboned poultry meat inspection is also explored in this study. The research results show that by

combining images at two different energies, it is possible to cancel the thickness variation of meat tissue, and make it easy to detect any embedded bone fragment.

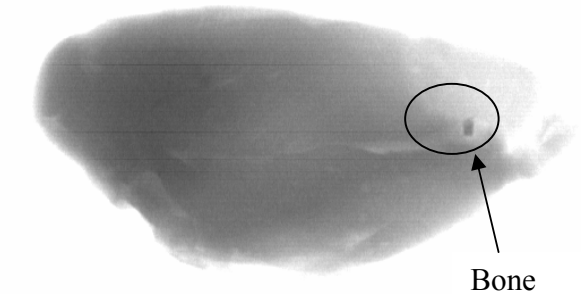
In the experiment here, two X-ray energy levels were 30 keV and 50 keV were selected for use. X-ray images taken at 30 keV have good overall image quality (see section 6.1) in terms of bone to meat contrast and signal to noise ratio. However, X-ray images taken at the 50 keV energy level lead to the smallest bone to meat contrast but the best signal to noise ratio. In other words, by differentiating the 30 keV and 50 keV images, the result is likely to compensate for the meat thickness variation and extract the bone information.

An example of bone detection using dual energy X-ray is shown in Figure 6.5. The meat sample is a chicken breast with an embedded bone fragment. The bone fragment is a small piece of fully calcified rib bone. In this example, the contrast difference between the 30 keV and 50 keV images is obvious, thus making it possible to compensate for the thickness variation in the 30 keV image by using the 50 keV image. After thickness compensation, as shown in Figure 6.5(c), the bone fragment is much brighter than the other parts of the meat sample, and can be easily extracted as shown in Figure 6.5(d).

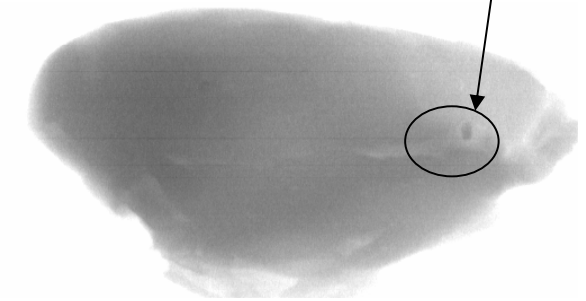
Another example is shown in Figure 6.6, which is a chicken breast with a piece of fan bone. Compared with rib bones, fan bones tend to be thin and less calcified. The contrast difference between the 30 keV image and 50 keV image is significant. However, the compensation is not as successful as in Figure 6.5. Thus, after the thickness compensation shown in Figure 6.6(c), the fan bone still can not be differentiated from the remaining thickness variation patterns, as shown in Figure 6.6(d). The difficulty of fan

bone detection can be explained as follows. Fan bones tend to be less calcified than ordinary bones, thus the X-ray attenuation difference between a fan bone and normal meat tissues is not as significant as that between a calcified bone and meat tissue. Also, fan bones tend to be thin, which make them even more difficult to differentiate from the surrounding meat tissues with X-ray imaging methods.

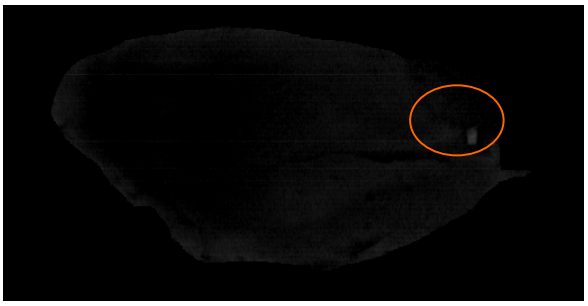
The previous two examples show both the potential and limitation of dual energy X-ray imaging technology in the application of poultry meat inspection. It was evident that by using two or more X-ray energies, it was possible to infer more information about the meat sample, which could lead to more accurate and robust detection of bone fragments. However, there are several inherent difficulties in this approach. First, theoretically, a monochromatic X-ray source is needed to cancel the thickness variation completely (Lehmann, et al., 1981), while in reality almost all of the commercially available X-ray sources are polychromatic. Thus in most cases, thickness compensation using dual energy can only be approximated to a certain extent, depending on the X-ray spectrum and materials of interest. Secondly, effective thickness cancellation requires a sufficient Compton scattering component for at least one of the X-ray energies. Compton scattering dominates at high X-ray energies, i.e., over 50keV (Dowsett, et al., 1998). However, this energy would be too high to produce sufficient image details for deboned poultry products. Finally, because of the explained physics limitations, a neural network seems to be the appropriate method to extract the subtle information about thickness variation and bone existence from the dual energy X-ray images. However, in this case the performance of the overall inspection system could be sensitive to the training procedure of the neural network (Graves and Batchelor, 2003).



(a) X-ray image at 30keV



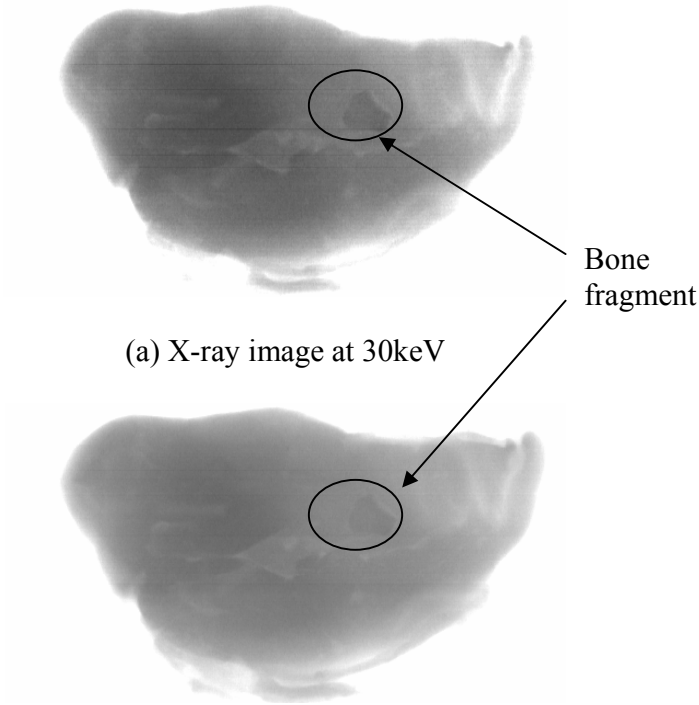
(b) X-ray image at 50keV



(c) Compensated image generated by the modeling process

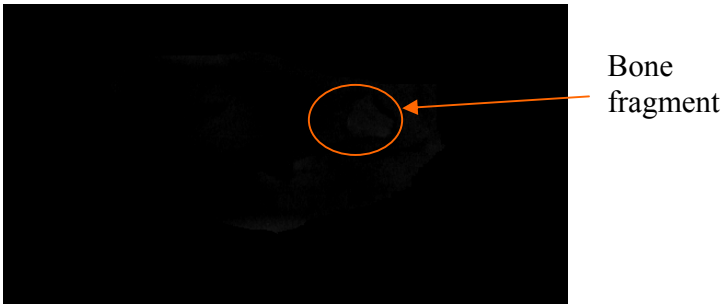
(d) Segmented image showing the bone fragment

Figure 6.5 Bone fragment detection using dual energy X-ray.

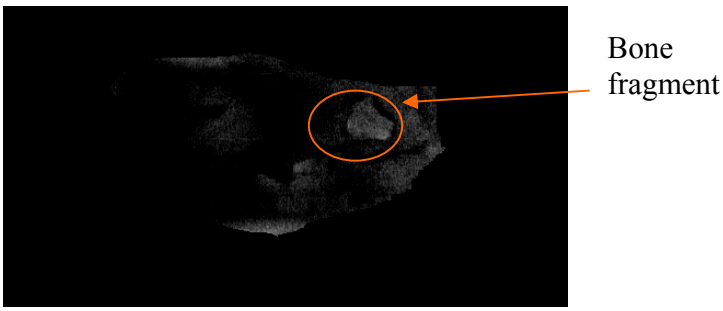


(a) X-ray image at 30keV

(b) X-ray image at 50keV



(c) Compensated image generated by the modeling process



(d) Segmented image showing the bone fragment

Figure 6.6 Limitation of bone fragment detection using dual energy X-ray.

6.3 Combined X-ray and laser 3D imaging

In this study, we adopted a laser 3D imaging subsystem to provide thickness information for each meat sample in real time, thus making it possible to solve the thickness variation problem for X-ray imaging.

As mentioned in the last chapter, given an X-ray image and laser 3D image of a meat sample, the integration process of these two images involves a series of major steps, namely,

1. Preprocessing, which removes noise and defines a region of interest;
2. Registration, which establishes geometric correspondences between pixels in the two images;
3. Modeling, which determines a model to describe the relationship of X-ray intensity and thickness of each meat sample and constructs a pseudo X-ray image based on the model and the laser 3D image;
4. Compensation, which compensates for the thickness variations in the X-ray image using the pseudo X-ray image;
5. Segmentation, which highlights the "suspicious" bright regions in the thickness compensated image;
6. Feature extraction, which extracts certain features for each "suspicious" region using blob analysis; and
7. Classification, which makes a decision about each blob based on its features and a set of classification criteria and reports the results of the classification.

The findings in each of the above stages are summarized in the following sections, and the overall performance of the combined X-ray and laser 3D imaging system is also evaluated based on sample test statistics.

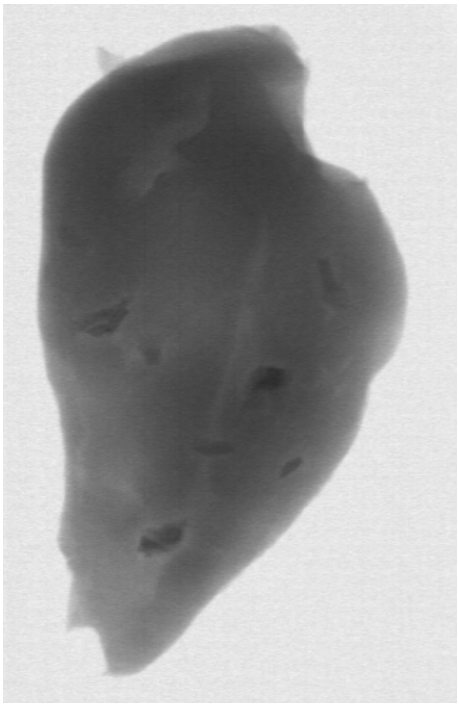
6.3.1 Preprocessing

For X-ray images, preprocessing involves noise removal and region of interest selection. The noises in the X-ray image are random noise and are modeled as Gaussian. In order to suppress the noise, several filters are compared, as shown in Figure 6.7. The 3x3 averaging and 3x3 Gaussian filters have the best performance in terms of noise suppression and detail preservation.

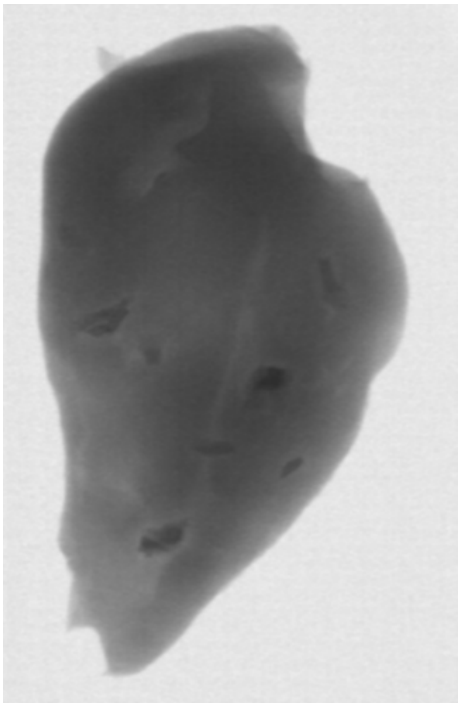
As shown in Figure 6.7, the 5x5 averaging and 5x5 Gaussian filters can provide better noise suppression than their 3x3 counterparts. However, the 5x5 filters lead to more blurring effects in the resultant images, which compromise the effective resolution of the X-ray image. Compared to the 3x3 averaging and 3x3 Gaussian filters, it was found that the Gaussian filter can provide better performance than the averaging filter.

6.3.2 Registration

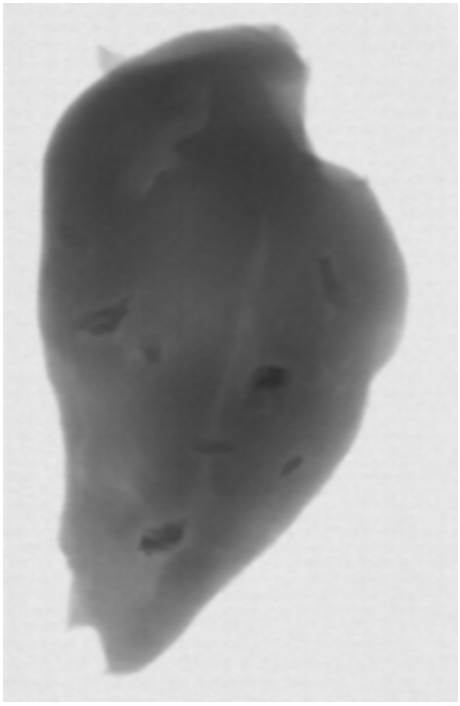
As mentioned previously, pixels in X-ray and laser images were matched to each other using the affine transform model. The experiments showed that the model was appropriate and sufficient to describe the geometric relationship between the two images. After the coefficients of the affine transform model are determined, the inverse transform is applied to the laser image to obtain a pair of registered X-ray and laser image of the same dimensions.



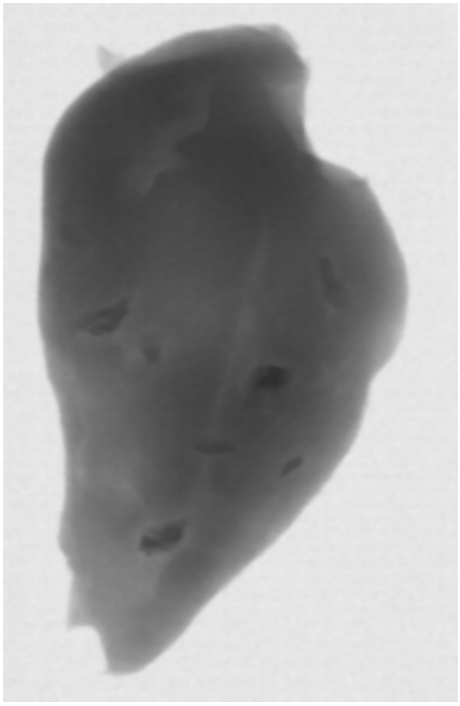
(a) The original X-ray image



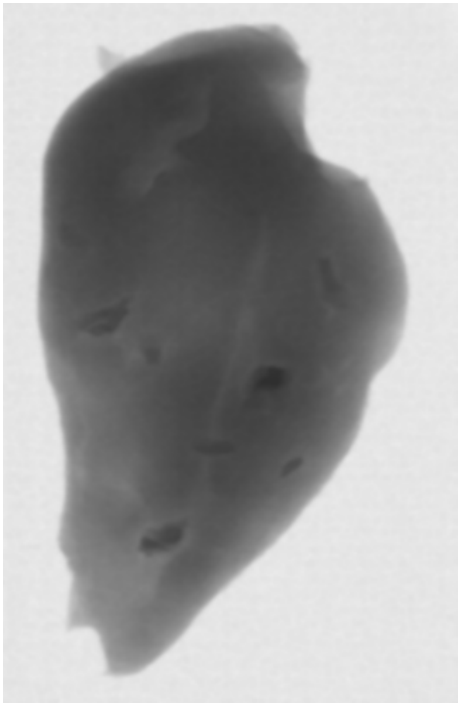
(b) Filtering using 3x3 average filter



(c) Filtering using 5x5 average filter

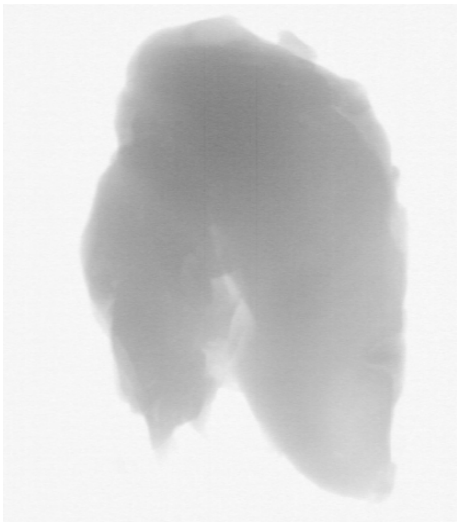


(d) Filtering with 3x3 Gaussian filter

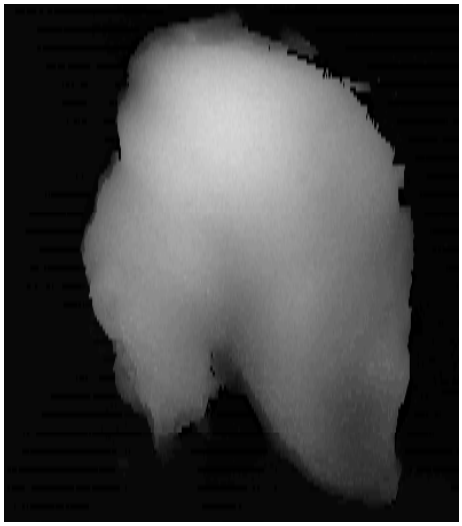


(e) Filtering with 5x5 Gaussian filter

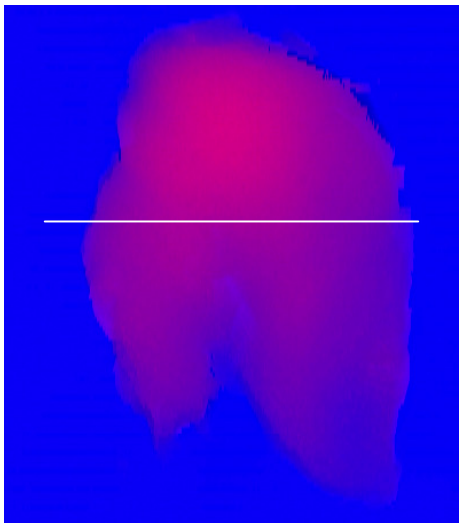
Figure 6.7 De-noising the X-ray image using averaging filters and Gaussian filters



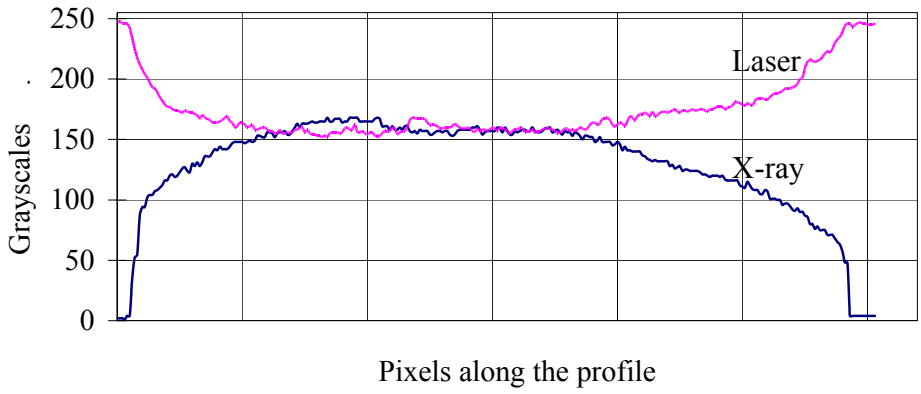
(a)



(b)



(c)



(d)

Figure 6.8 Registration of the X-ray and laser range images. (a) The original X-ray image of a chicken breast. (b) The registered laser 3D image of the chicken breast. (c) The superimposed image with laser image frame in the red channel and X-ray image frame in the blue channel. (d) The grayscale profile along the white line in Figure (c).

As shown in Figure 6.8(d), after transformation, the two images are matched to each other geometrically. Note that the parameters of the geometric transform model were estimated beforehand (see section 5.4.1), and the inverse geometric transform was applied to the laser 3D image during the real-time image acquisition.

6.3.3 Modeling

Modeling is the process of determining the grayscale relation between the X-ray and laser 3D images (see equation 5.20). Once the model is determined, a pseudo X-ray image can be generated based on the laser 3D image and the difference between the pseudo X-ray and true X-ray images is minimized. In this section, the modeling performance and calculation time were evaluated in regards to the model order and decimation issues.

As mentioned before, the model was estimated using linear polynomial regression (see chapter 5.4.2). The fitness of the regression is evaluated by the R^2 value, which is defined as,

$$R^2=1-\sum(X_{regression}-X_i)^2/\sum(X_i-\overline{X})^2 \qquad (6.1)$$

where,

$X_{regression:}$ the estimated response variable, and

X_i: is the explanatory variable, and

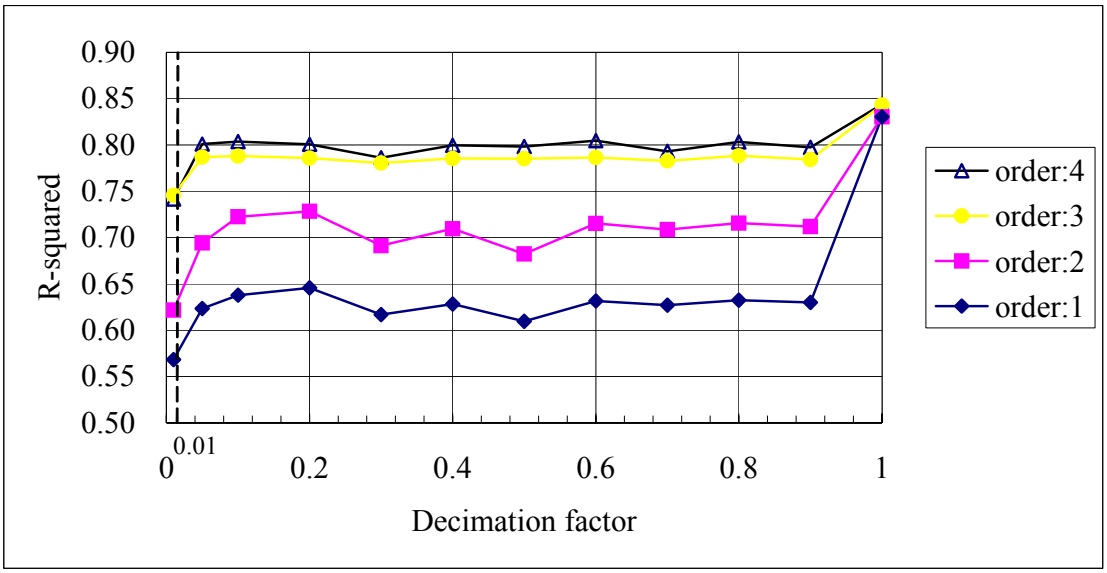
$\overline{X}$: the mean of the explanatory variables.

In our case, the explanatory variable is the pixel grayscale in the laser 3D image, the response variable is the pixel grayscale in the X-ray images, and the regression variable is the pixel grayscale in the pseudo X-ray image.

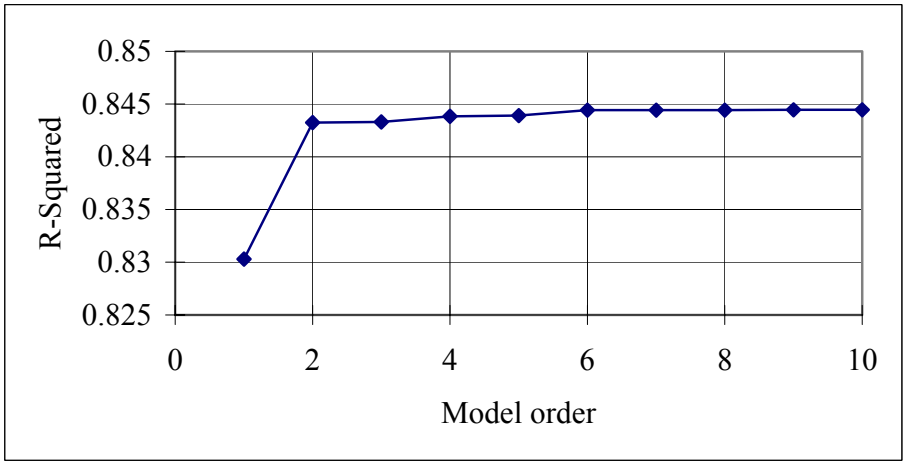
The performance of the regression process was particularly affected by the order of the model, and a decimation factor. It is found that the performance of the model is better for higher orders; however, higher order models require more time to compute, which could be undesired in real time applications.

Note that we can generate the model using decimated (also known as down-sampled) versions of the images, thus the pixels actually used in the modeling process can be significantly reduced. The benefits of this decimation approach include saving of computation time, and image noise suppression due to decimation. The drawback of this approach is the sacrifice of performance, because not every pixel in the original images participates in the formation of the model.

As shown in the Figures 6.9, 6.10 and 6.11, the performance can be reduced if the image is decimated; however the computation time is also reduced. The tradeoff between performance and speed is needed here. As shown in Figure 6.9(a), the R^2 value can be improved by increasing model order however, the pace of improvement slows down after the model order reaches 3, as shown in Figure 6.9(b).



(a)



(b)

Figure 6.9 Relation of regression performance, decimation factor and model order. (a) Plot of regression performance vs. decimation factor for model orders up to 4. (b) Plot of regression performance vs. model order, with decimator factor being 1.

The computation time for the modeling process of the same image was compared at different model orders, as shown in Figure 6.10, where the computation time increases

monotonically with increasing model order. Considering the computation time with decimation factors, as shown in Figure 6.11, it was confirmed that the less pixels involved in the modeling process, the less time it takes to compute the model. It was also verified that, higher order models take more time to compute.

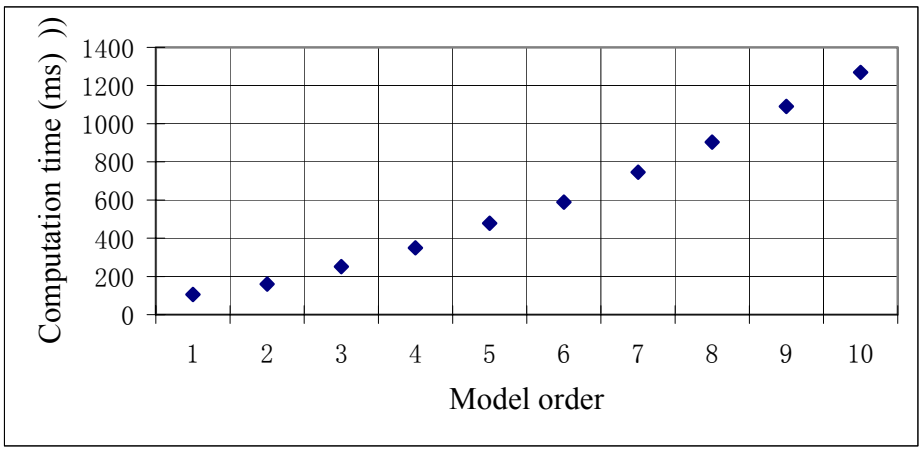


Figure 6.10 Computation time for generating the model at full resolution (i.e., the decimation factor is 1) with different model orders

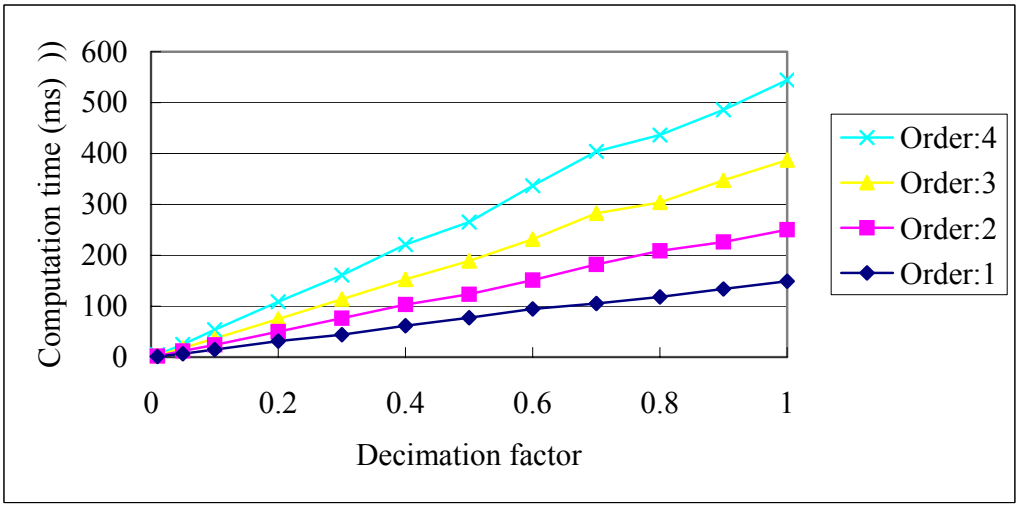
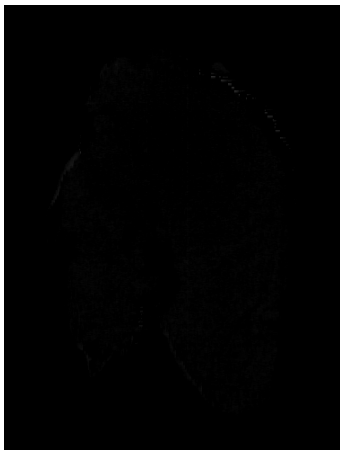


Figure 6.11 Relation between computation time with the decimation factor for model orders 1 to 4.

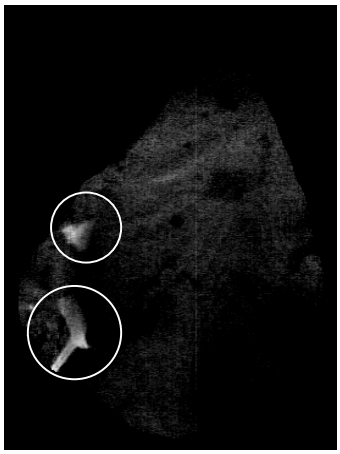
6.3.4 Segmentation and blob analysis

After thickness compensation, the thickness variations of the meat sample are canceled out, and the remaining bright regions in this compensated image represent either physical contaminants or image noises. A typical compensated image is a uniform image with most pixels approaching a zero value. An automatic threshold selection method is applied to the thickness compensated image for image segmentation. However, due to brightness differences among hazard objects, such as different types of bone fragments or metal objects, automatic selection of a global threshold value for different situations becomes complicated. In this study, the localized Otsu's method is used. The original thickness-compensated image is partitioned into sub-images, and Otsu's method is applied to each of the sub-image. For each sub-image, a sub-threshold value and sub-discriminant ratio λ is calculated. The sub-threshold values and sub-discriminant ratios are then partitioned into two groups. One group contains the lower sub-threshold values and lower discriminant ratios, while the other contains the remaining higher values. The global threshold is determined by choosing the lowest value from the second group. In order to illustrate the effectiveness of this method, images that represent different on-line situations were selected.



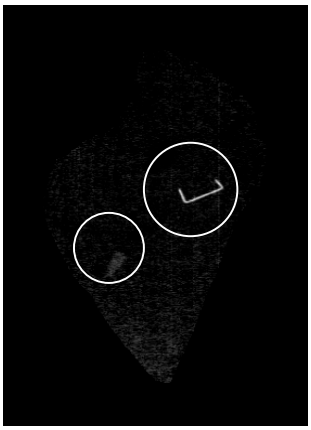
(a)

Image One



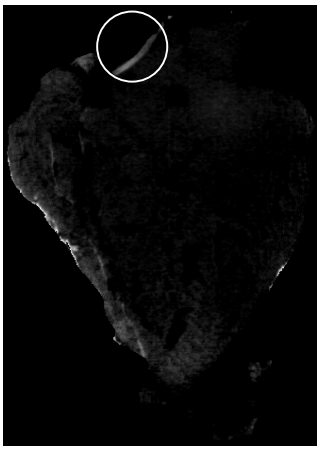
(b)

Image Two



(c)

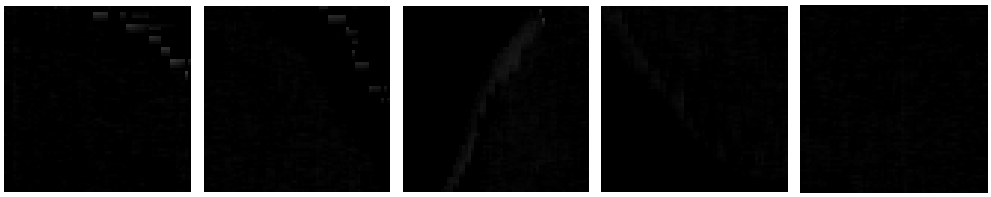
Image Three



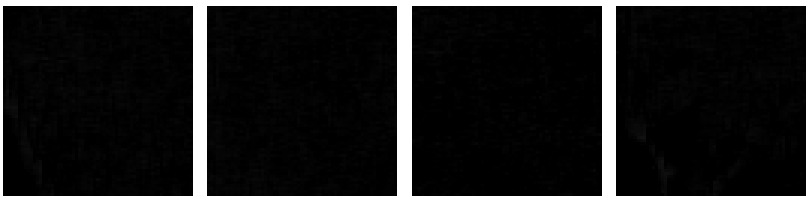
(d)

Image Four

Figure 6.12 Thickness-compensated images for different situations: (a) no hazardous objects in the chicken meat image, (b) One fan bone and one rib bone, as highlighted, existed in the chicken meat image, (c) One fan bone and two metals, as highlighted, presented in the chicken meat image, (d) One rib bone, as highlighted, presented in the chicken meat image.

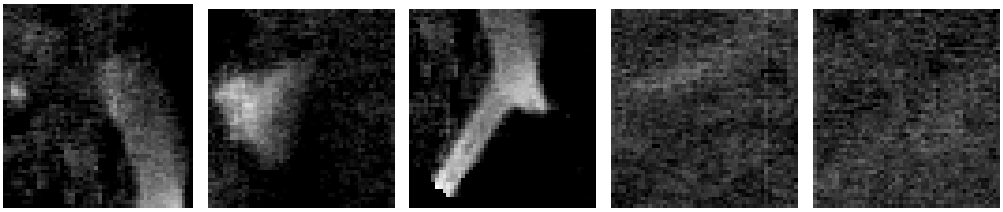


Sub 1 Sub2 Sub3 Sub4 Sub5

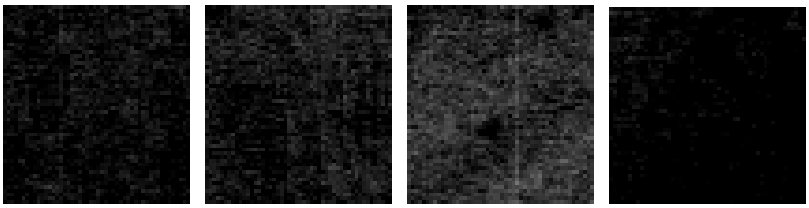


Sub 6 Sub7 Sub8 Sub9

(a) Selected nine sub-images of Image One

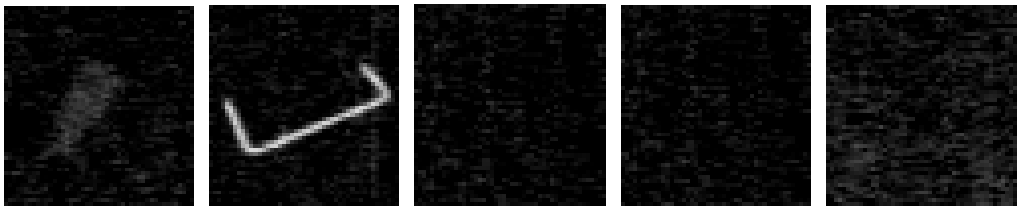


Sub 1 Sub2 Sub3 Sub4 Sub5

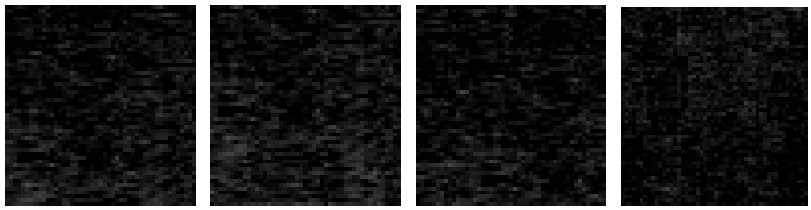


Sub 6 Sub7 Sub8 Sub9

(b) Selected nine sub-images of Image Two

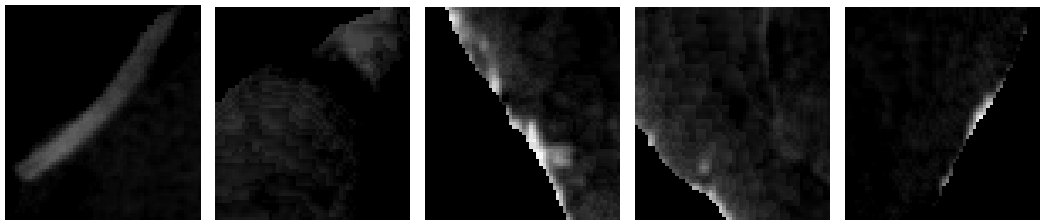


Sub 1 Sub2 Sub3 Sub4 Sub5

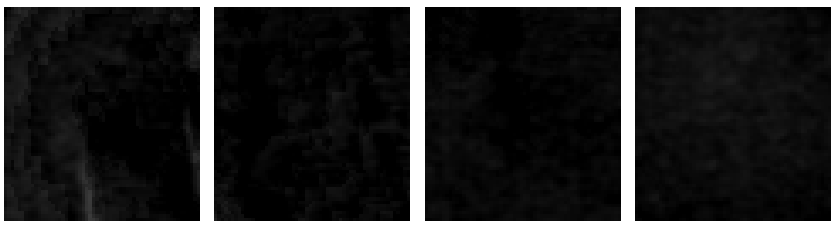


Sub 6 Sub7 Sub8 Sub9

(c) Selected nine sub-images of Image Three



Sub 1 Sub2 Sub3 Sub4 Sub5



Sub 6 Sub7 Sub8 Sub9

(d) Selected nine sub-images of Image Four

Figure 6.13 Selected nine sub-images for Image One, Image Two, Image Three and Image Four shown in Figure 6.12.

Figures 6.12(a) – (d) show different types of situations. Figure 6.12(a) shows an image with no hazardous objects inside, (b) is an image with uniformed brightness bone fragments, (c) shows an image with bright metal clips (big and small) and relatively dark fan bone fragments and (d) is an image with high edge noise and a big hard bone fragment. Each of the four images was partitioned into sub-images. Figure 6.13 shows the selected nine sub-images for each of the four images. When Otsu's method is applied to the sub-images, the background pixels are excluded. Table 6.1 lists the results of the localized Otsu's method.

Table 6.1 Sub-threshold values and sub-discriminant ratios of test images

	Image One		Image Two		Image Three		Image Four	
	Threshold	λ	Threshold	λ	Threshold	λ	Threshold	λ
Sub1	15	14	47	856	48	348	51	892
Sub2	13	7	69	1240	79	1138	31	62
Sub3	7	11	68	2401	10	107	71	1122
Sub4	2	3	46	261	10	59	69	162
Sub5	2	2	45	251	19	159	69	199
Sub6	2	4	14	92	17	137	12	64
Sub7	2	3	21	174	20	137	6	19
Sub8	2	2	42	256	15	116	5	14
Sub9	2	4	13	110	12	120	11	18

Table 6.1 shows that, if the pixel values in the sub-images present clear bimodal distributions, the discriminant ratios λ in Otsu's method are very high and the threshold

values are selected accurately, as in Image (b) Sub2 and Sub3. However, if the bimodal distribution is not obvious, such as the case for most of the noise sub-images, the values of λ are low. On the other hand, when the pixels in the sub-images are bimodal, but the two classes have very unequal pixel numbers, the λ value would also drop to very low, as shown in Image (d) Sub4.

Table 6.2 Partition result of sub-threshold values and λ ratio

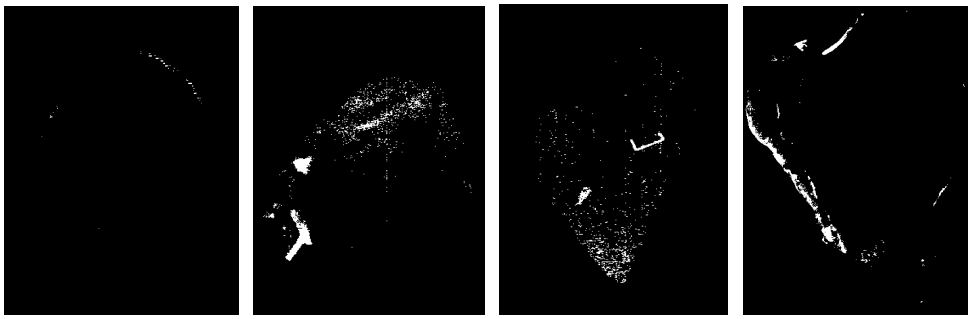
Group Num	Image One			Image Two			Image Three			Image Four		
	Sub-index	Thres.	λ	Sub-index	Thres.	λ	Sub-index	Thres.	λ	Sub-index	Thres.	λ
Group Two	Sub1	22	14	Sub1	47	856	Sub1	48	348	Sub1	51	892
	Sub2	19	7	Sub2	69	1240	Sub2	79	1138	Sub3	71	1122
				Sub3	68	2401				Sub4	69	162
										Sub5	69	199
Group One	Sub3	7	6	Sub4	46	261	Sub3	10	107	Sub2	31	62
	Sub4	2	3	Sub5	45	251	Sub4	10	59	Sub6	12	64
	Sub5	2	2	Sub6	14	92	Sub5	19	159	Sub7	6	19
	Sub6	2	4	Sub7	21	174	Sub6	17	137	Sub8	5	14
	Sub7	2	3	Sub8	42	256	Sub7	20	137	Sub9	11	18
	Sub8	2	2	Sub9	13	110	Sub8	15	116			
	Sub9	2	4				Sub9	12	120			

After partitioning the sub-threshold values into two groups (the λ ratio is used when the auto-threshold value is around the valley of the two peaks), the partitioned

result is shown in table 6.2, where the global threshold values selected for each of the four images are highlighted. After thresholding, the resultant binary images are shown in Figure 6.14. In contrast, binary images obtained by using global Otsu's thresholding method and the prior knowledge based thresholding method are presented in Figure 6.15.

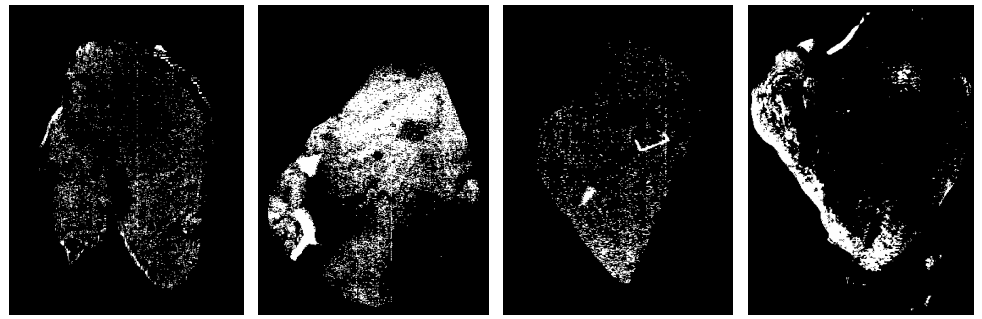
Morphological operations were applied to binary images of Figures 6.14 and 6.15 to obtain blob images, as shown in Figures 6.16-6.18. It shows that global Otsu's method breaks down in most of the cases. Prior knowledge based method works better. However, in an image where hazard objects present obvious difference of brightness levels, the prior knowledge method will lead to miss-detection of bones with lower intensities, such as in Image(c). Also, a potential problem of this method is that when the total area of hazard objects is large, this method tends to cause miss-detection. The localized Otsu's method shows the best performance in the sampled images. However, more parameters are to be determined, such as the sub-image size and the two group partition rules (ratios). In this application, the sub-image size was 60-by-60, and two group partition rules were:

1. Partition the sub-threshold values by calculating the distances between two extreme sub-thresholds.
2. For the sub-threshold values around the middle of the two extreme sub-thresholds, judge the λ value to determine its group number.

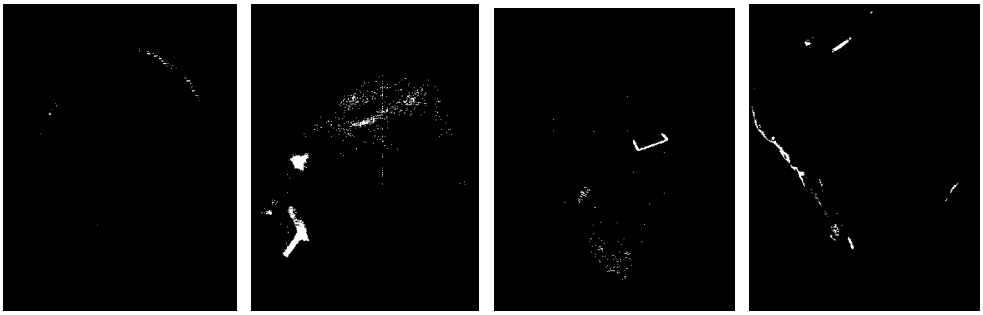


(a) (b) (c) (d)

Figure 6.14 Binary Images obtained by using local Otsu's method for global threshold selection. (a) Binary image of Image One. (b) Binary image of Image Two.(c) Binary image of Image Three. (d) Binary image of Image Four.

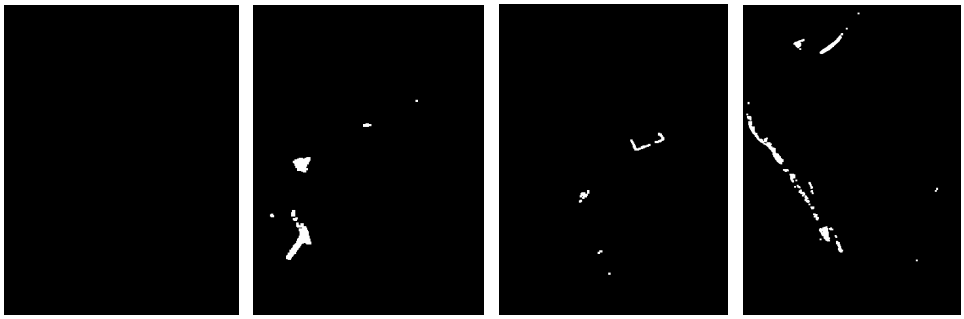


(a) (b) (c) (d)



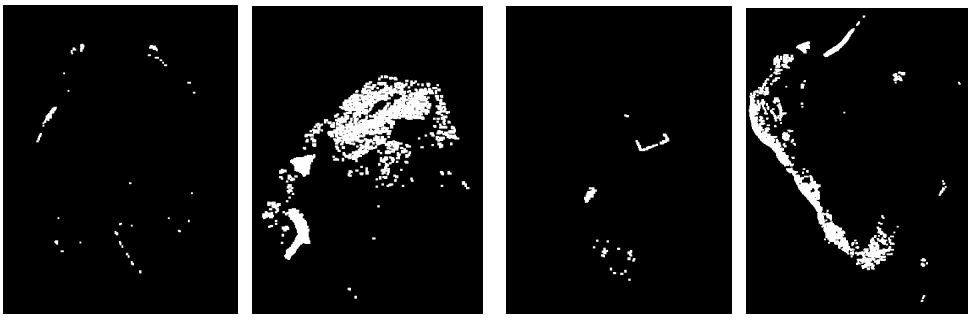
(e) (f) (g) (h)

Figure 6.15 Binary Images obtained by using Otsu's method and prior knowledge based method for global threshold selection. (a)-(d) Using Otsu's method to obtain binary images. (a) Binary image of Image One. (b) Binary image of Image Two.(c) Binary image of Image Three. (d) Binary image of Image Four. (e)-(f) Using prior knowledge method to obtain binary images. (e) Binary image of Image One. (f) Binary image of Image Two.(g) Binary image of Image Three. (h) Binary image of Image Four.



(a) (b) (c) (d)

Figure 6.16 Blob Images generated based on the binary images obtained by using local Otsu's method. (a) Blob image of Image One. (b) Blob image of Image Two. (c) Blob image of Image Three. (d) Blob Image of Image Four.



(a) (b) (c) (d)

Figure 6.17 Blob Images generated based on the binary images obtained by using global Otsu's method. (a) Blob image of Image One. (b) Blob image of Image Two. (c) Blob image of Image Three. (d) Blob Image of Image Four.

(i) (j) (k) (l)

Figure 6.18 Blob Images generated based on the binary images obtained by using prior knowledge based method. (a) Blob image of Image One. (b) Blob image of Image Two. (c) Blob image of Image Three. (d) Blob Image of Image Four.

6.3.5 Classification

The blobs are examined based on their feature vectors and a set of classification rules. In this final stage, a rule-based classifier was used to make an accept/reject decision on each blob. In this section, four groups of sample image are used to illustrate the detection process, and then statistics of the overall detection system is presented.

Sample 1 is a piece of chicken breast without a bone fragment inside. As shown in Figure 6.19(b), the bumpy thickness profile of the sample leads to dramatic grayscale variations in the X-ray image. If only the X-ray image is considered in the inspection process, some of the bumpy regions, such as the circled area, would probably be misclassified as abnormal, such as a bone fragment or some other type of physical hazard. Fortunately, the thickness information of the whole sample was captured in the laser image (Figure 6.19(a)), and a pseudo X-ray image (Figure 6.19(c)) could be formulated, which was similar to the true X-ray image. The compensated image (Figure 6.19(d)), the difference between Figures 6.19 (b) and 6.19 (c), is flat, except for some minor noise on

the sample boundary. Since there was no qualifying blob in the compensated image, the rule-based classifier finds no abnormities in the sample, so the resultant image (Figure 6.19(e)) is empty, i.e., labeled zero everywhere.

The second sample was a piece of chicken breast with two bones inside, as shown in Figure 6.20(e). The triangle shaped abnormality is a thin fan bone (thickness less than 0.5mm), and the boot-shaped one is a pulley bone. Similar to the last sample, a pseudo X-ray image (Figure 6.20(c)) was formulated to yield a compensated image (Figure 6.20 (d)), where the thick meat areas were canceled out and the bones emerge. There is still some noise in the compensated image. Some of the noises were removed by standard filtering and/or morphology operations and the rule-based classifier will eliminate the remaining noises and identify the true abnormalities. As shown in the resulting image (Figure 6.20 (e)), the two bones were highlighted and all of the noises are successfully suppressed.

As shown in Figure 6.21, the third sample contains a small rib bone and a fan bone, along with a meat protrusion. If we look at only the X-ray image (Figure 6.21(b)), the meat protrusion would be classified as a large bone fragment. In the laser 3D image(Figure 6.21(a)), the thickness variation was reflected. Comparing the true X-ray image (Figure 6.21(b)) with the reconstructed pseudo X-ray image (Figure 6.21(c)), the meat protrusion appears in both images, while the bone fragments appear only in the true X-ray image. Similar to the previous samples, thickness variations were compensated (Figure 6.21(d)) and the bone fragments were extracted (Figure 6.21(e)).

The last sample presented in Figure 6.22 was a piece of chicken breast with a small bone fragment along with four metal clips. Two of the metal clips was very small

and could be missed with image noises. In this example, the compensated image (Figure 6.22(d)) was obtained the same way as the previous examples, and then it was segmented into blobs. The mean values of metal clip blobs were much higher than those of bone or image noise, thus the metal clips would be identified even if its size were very small. In the final image (Figure 6.22(e)), four of the metal clips were extracted, and the bone fragment was also identified.

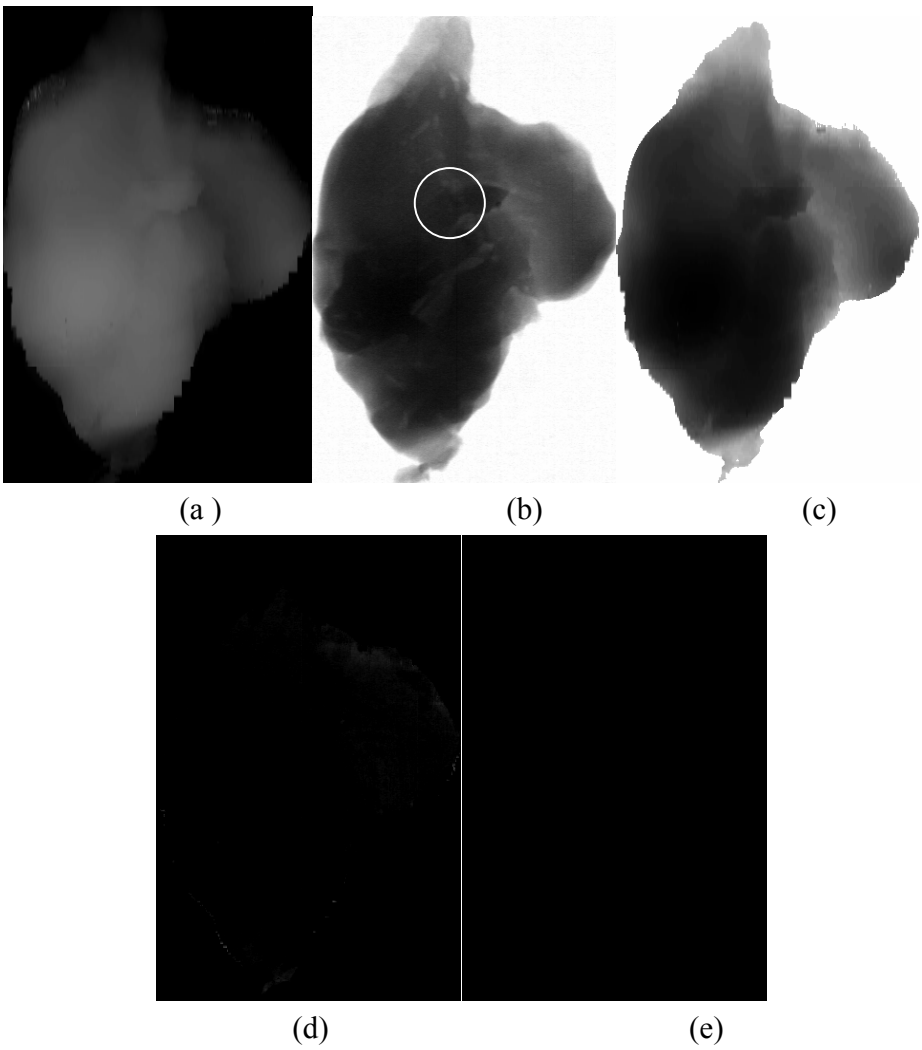


Figure 6.19 Images of a sample without bone. (a)-(e) are respectively the laser 3D image, X-ray image, pseudo X-ray image, compensated image and the result .

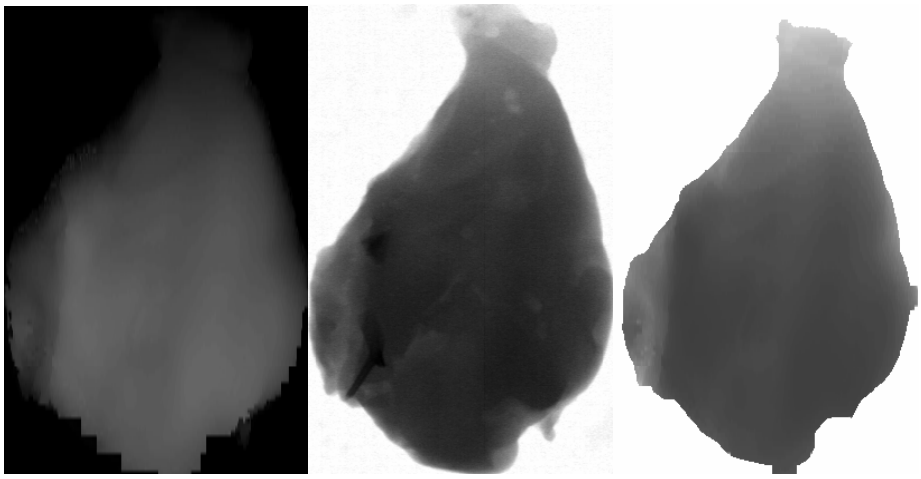


Figure 6.20 Images of a sample with two bones. (a)-(e) are respectively the laser 3D image, X-ray image, pseudo X-ray image, compensated image and the result image.

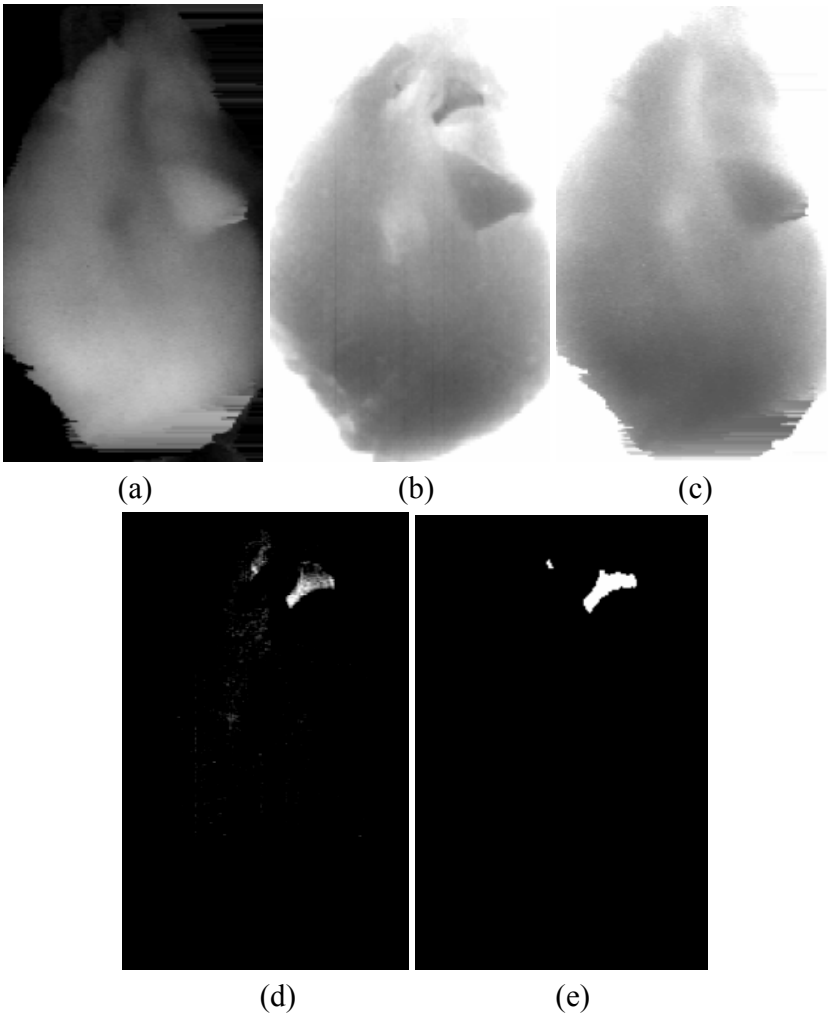
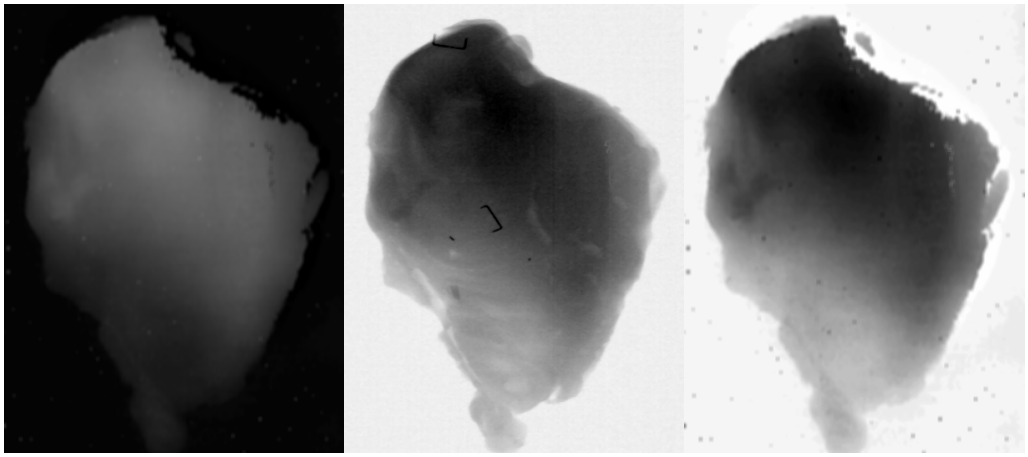
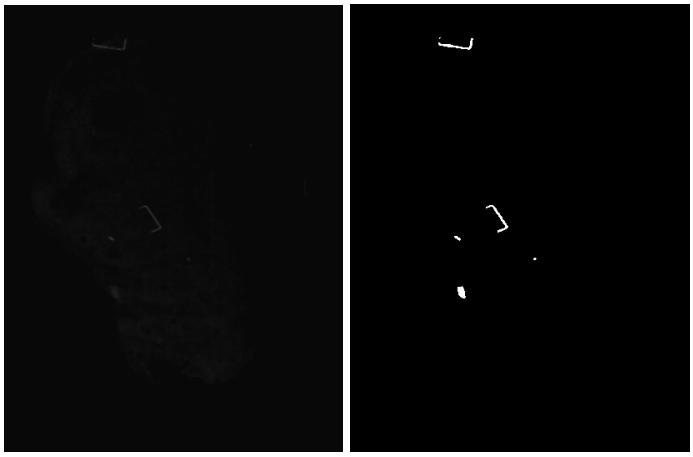


Figure 6.21 Images of a sample with a fan bone, a small rib bone and a meat protrusion. (a)-(e) are respectively the laser 3D image, X-ray image, pseudo X-ray image, compensated image and the result image.



(a) (b) (c)



(d) (e)

Figure 6.22 Images of a sample with a bone fragment and four metal clips. (a)-(e) are respectively the laser 3D image, X-ray image, pseudo X-ray image, compensated image and the result image.

The classification results were different for different types of physical contamination. We tested the inspection system with 131 pieces of chicken breast samples. According to a post test destructive inspection, there were altogether 152 pieces of rib bones and 53 pieces of fan bones in the samples. Rib bones are fully calcified while fan bones are only partially calcified, which makes fan bone less detectable by the imaging system.

The detection rates of the two types of bone fragment are summarized in Figure 6.23 and 6.24. For rib bone detection, 95.37% detection rate can be achieved for bones larger than 114 pixels. In this system, the pixels 0.4mm by 0.4mm squares, so the bone fragment of 114 pixels in the image plane would occupy an area of 18mm^2 in real world. For fan bones, it is found that for bones larger than 109 pixels, or 17.5 mm^2, the detection rate can reach as high as 91.67%. This fan bone detection rate is significant, because fan bones have much lower detectability than rib bones under X-ray imaging since they are usually thin and not fully calcified. The detection performance for rib bones, fan bones and their mixture is further summarized in table 6.3.

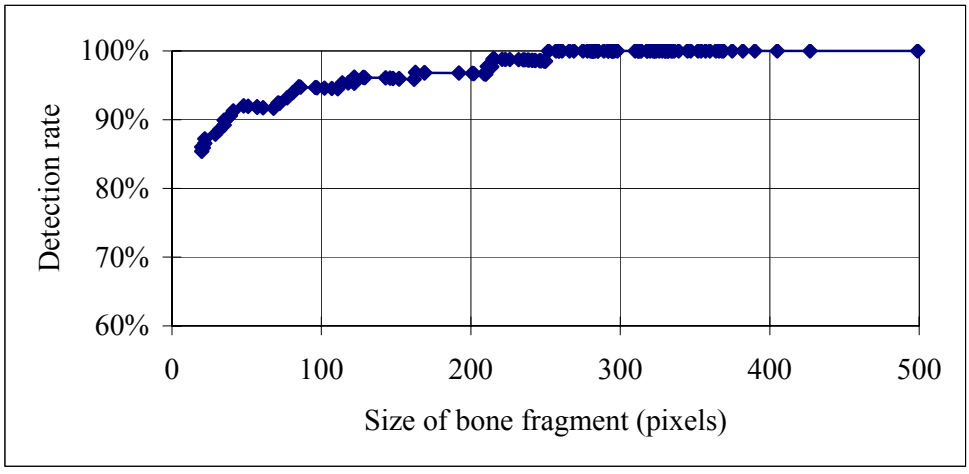


Figure 6.23 Detection rate of the rib bone fragment

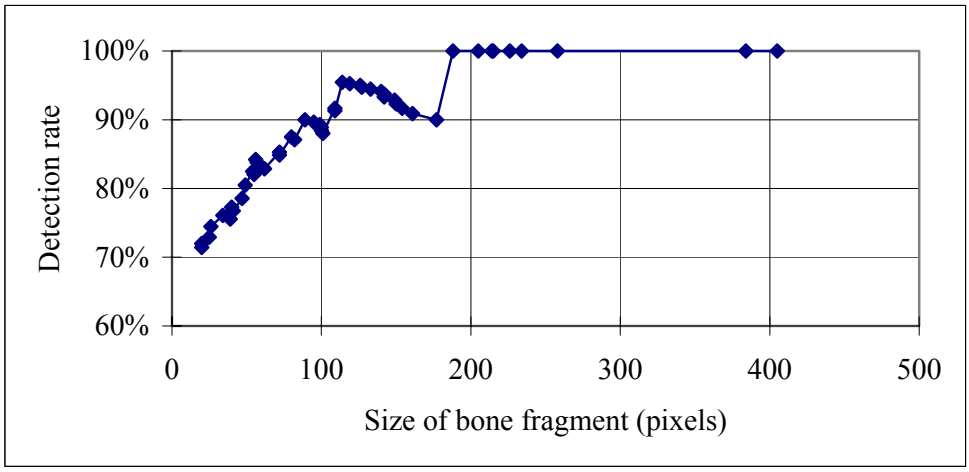


Figure 6.24 Detection rate of fan bone fragment.

Table 6.3 Statistical results of the rib bone and fan bone detection

Size of bone(pixels)[5]	# of bones			Detected bones			False negative rate (%)		
	rib bone	fan bone	mixed	rib bone	fan bone	mixed	rib bone	fan bone	mixed
≥300	39	2	41	39	2	41	0	0	0
≥200	92	8	100	89	8	97	3.26	0	3
≥70	119	34	153	110	29	139	7.56	14.71	9.15
≥20	137	50	187	117	36	153	14.60	28	18.18

Another important measure of the inspection system was the overall type II error rate, or false positive rate. According to the experiment results, the type II error rate of the inspection system was 9.78%. The cause of the false alarms is twofold. Firstly, the laser scattering effect on the meat surface could cause information loss in the images. As shown in Figure 6.25, scattering effects can cause artifacts on the meat boundary in the resultant combined image, and these artifacts are not differentiable from bone fragments. Secondly, the random image noises could also lead to false patterns in the thickness

[5] There are 15 rib bones and 3 fan bones with size smaller than 20, all of which were not detected by the imaging system and not included in this table.

compensated image. The sources of the image noises include thermal noise of the X-ray and laser imaging devices, random X-ray photon emission, and the scattering effects of the laser.

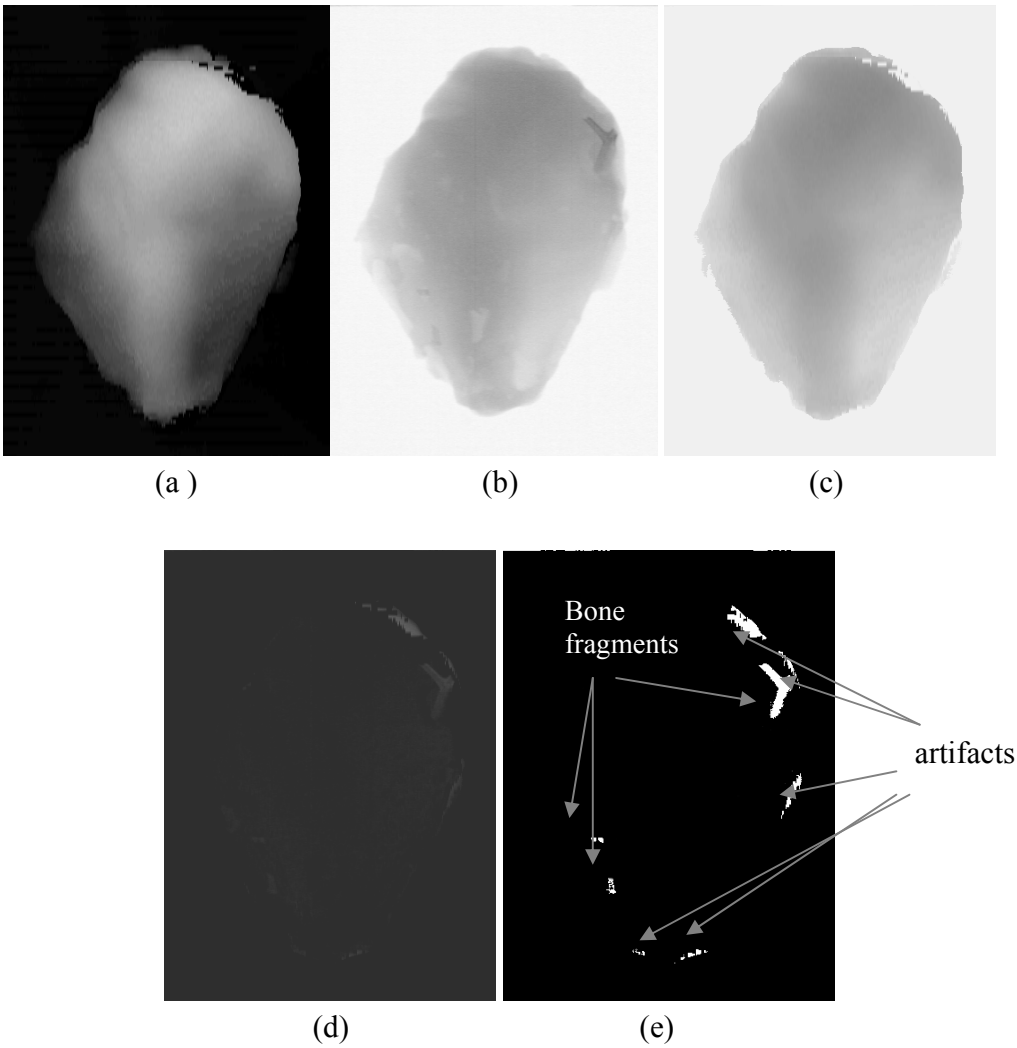


Figure 6.25 False positive caused by scattering effect in laser 3D image. (a)-(e) are respectively the laser 3D image, X-ray image, pseudo X-ray image, compensated image and the result image.

CHAPTER 7

CONCLUSIONS

It was shown that the combined X-ray and laser 3D imaging method can successfully overcome the thickness variation issue and detect the embedded physical contaminations in deboned poultry meat. A synergic inspection system was developed using this combined method for the challenging problem of deboned poultry product inspection, which can detect rib bones, fan bones and other intrusions such as metal clips.

The optimal X-ray energy range suitable for deboned poultry product inspection has been determined to be 28 keV to 31keV, which leads to satisfactory contrast and a satisfactory signal to noise ratio (SNR) with the X-ray imaging system used in this study. Although the optimal X-ray energy range would be different for a different X-ray system, or for a different detection problem, the idea of evaluating the energy level by considering both contrast and the SNR is still valid.

Results showed that the dual energy X-ray method can address the thickness issue to a certain extent; however, the inherent physical limitations make it difficult to cancel out the thickness variation using the dual energy method. In this sense, the combined X-ray and laser 3D approach is better because the thickness information of the meat was acquired directly by the system, and can be used to completely compensate for the thickness variation.

It is shown in this study that, in order to integrate the X-ray and laser 3D parts into a coherent functioning system, certain key challenges have been overcome. The

registration between the two subsystems was crucial for the data fusion tasks, and results confirm that the X-ray and laser subsystems have been correctly registered to each other, or in other words, the coordinates association between the two imaging subsystems has been established. Another important task was determining the grayscale relationship between pixels in the X-ray and laser 3D images, and it was shown that a least squares based method can be used to estimate the relationship, which makes it possible to compensate for the thickness variation in the X-ray images.

It has been shown that this combined X-ray and laser 3D imaging system has real-time capability. A multithread architecture was constructed in this system, which can initiate and synchronize various image processing jobs at the same time. The multithread framework can enhance the efficiency of the available computing hardware and data bandwidth in the system, promote the overall throughput, and improve system responsiveness.

CHAPTER 8

FUTURE STUDIES

In order to further improve the combined X-ray and laser 3D technology, future studies can be conducted in areas including:

1. Updating the X-ray image sensor and/or the X-ray source, which can lead to lower noise levels in the X-ray images.
2. Exploring more robust and effective algorithms for image segmentation, feature extraction and pattern classification.
3. Managing errors that come from laser scattering effects.

REFERENCES

Abramson, N. 1991.Time reconstruction in light-in-flight recording by holography. *Appl. Opt*. 30:1242-1252.

AS&E. 2003. AS&E®Introduces Z®Backscatter Van (ZBV) X-ray Screening System, [Online] available www:http://www.as-e.com/ inthenews/052803_zbackscattervan.html.

Batistoni, P., S. Rollet, Y. Chen, U. Fischer, L. Petrizzi and Y. Morimoto. 2003. Analysis of dose rate experiment: comparison between FENDL, EFF/EAF and JENDL nuclear data libraries. *Fusion Engineering and Design* 69(1-4):649-654.

Blake, G. M., and I. Fogelman. 2002. Dual energy X-Ray absorptiometry and its clinical applications. *Seminars in musculoskeletal radiology* 6(3):207-217

Bovik, A. 2000. Handbook of image and video processing. San Diego:Academic Press.

Carlsson, T. E. 1993. Measurement of three dimensional shapes using light-in-flight recording by holography. *Opt. Eng*. 32:2587-2592.

Carlsson, T. E. , B. Nilsson, and J. Gustafsson. 2001. System for acquisition of three-dimensional shape and movement using digital Light-in-Flight holography. *Opt. Eng*. 40(1):67-75.

Casasent, D. A. Talukder, P. Keagy and T. Schatzki. 2001. Detection and segmentation of items in X-ray imagery, *Transactions of the ASAE* 44(2): 337-45.

Casasent, D. and X. Chen. 2003. New training strategies for RBF neural networks for X-ray agricultural product inspection. *Pattern Recognition* 36(2):535-547.

Chen, F., G. M. Brown, and M. Song. 2000. Overview of three-dimensional shape measurement using optical methods. *Opt. Eng*. 39(1):11-22.

Dobbins, J. T. III, H. MacMahon, and J. Ricke. 2003. Dual energy imaging user perspectives. [Online].available www: http://www.gemedicalsystems.com/rad/xr/radio/products/de_perspectives.html.

Dowsett, D. J., P. A. Kenny and R. E. Johnston. 1998. Physics of diagnosis imaging. London: Chapman & Hall Medical.

Duda, R. O., P. E. Hart and D. G. Stork. 2000. Pattern Classification. New York, NY: Wiley.

Fainberg, A. 1992. Explosives Detection for Aviation Security. *Science* 255:1531-1537.

Fann, S., Y. Jiang, H. Hwang. 2003. An innovative a-Si:H p-i-n based X-ray medical image detector for low dosage and long exposure applications. *Applied Surface Science* 212-213:765-769.

Forsyth, D. A., and J. Ponce. 2003. Computer vision. Upper Saddle River, NJ: Prentice Hall.

General Electric Company. 2003. Dual energy imaging. [Online] available www:http://www.gemedicalsystems.com/rad/xr/radio/products/xqi/dual_energy.html.

Genant, H. K., K Engelke, T Fuerst, C. C. Gluer, S. Grampp, S. T. Harris, M. Jergas, T. Lang, Y. Lu, S. Majumdar, A. Mathur, and M. Takada. 1996. Noninvasive assessment of bone mineral and structure: state of the art. *Journal of Bone and Mineral Research* 11(6):707-730.

Godin, G. , J. A. Beraldin, J. Taylor, L. Cournoyer, M. Rioux, S. El-Hakim, R. Baribeau, F. Blais, P. Boulanger, J. Domey, and M. Picard. 2002. Active Optical 3D Imaging for Heritage Applications. *IEEE Computer Graphics and Applications* 22(5):24-36 .

Graham, D. T. and P. Cloke. 2003. Principles of radiological physics, 4th ed. New York: Churchill Livingstone.

Graves, M., Smith A., Batchelor B. G. and Palmer S.C. 1994. Design and analysis of X-ray vision systems for high-speed detection of foreign body contamination in food. In *proc. SPIE* 2347: 80-92. Bellingham, WA: The International Society for Optical Engineering.

Graves, M., and B. Batchelor. 2003. Machine vision for the inspection of natural products, London: Springer.

Grimson, W. E. L. 1981. From images to surfaces : a computational study of the human early visual system. Cambridge, MA: MIT Press.

Haff, R. P., and T. F. Schatzki. 1997. Image restoration of line-scanned X-ray images. *Opt. Eng*. 36(12):3288-3296.

Hale, J. 1975. The fundamentals of radiological science. Springfield, IL: Charles C Thomas.

Hallowell, S. F. 2001. Screening people for illicit substances: a survey of currentportal technology. Talanta 54:447-458.

Hamamatsu. 2000. Transmission mode soft X-ray tube N7599, N7599-1. [Online]. available WWW: http://usa.hamamatsu.com/hcpdf/parts_N/N7220.pdf.

Hartman, L. R. 2001. X-ray inspection boosts confidence through the pipeline. Packaging digest 2001:54.

Horn, B. K. P., and M. J. Brooks. 1986. The variational approach to shape from shading. *Computer Vision and Image Understanding* 33:174-208.

Imaginis. 2000. [Online] available www: http://imaginis.com/xray/xray_why.asp.

Jalkio, J. A., R. C. Kim, and S. K. Case. 1985.Three dimensional inspection using multistripe structured light. *Opt. Eng*. 24(6):966-974.

Jamieson, V. 2002. Physics raises food standards. [Online] available www: http://physicsweb.org/article/world/15/1/3.

Jain, Anil K. 1989. Fundamentals of digital image processing. Englewood Cliffs, NJ: Prentice Hall.

Jain, R., R. Kasturi, and B. G. Schunck. 1995. Machine Vision. Columbus, OH: McGraw-Hill.

Johns, P. C., D. J. Drost, M. J. Yaffe, and A. Fenster. 1985.Dual-energy mammography: initial experimental results. *Med. Phys.* 12(3):297-304.

Kim, S. and T. F. Schatzki. 2000. Apple watercore sorting system using X-ray imagery: I. algorithm development. *Transactions of the ASAE* 43(6): 1695-1702 .

Laplante, P. A. and A. D. Stoyenko. 1996. Real-time imaging: theory, techniques, and applications. New York, N.Y.: IEEE press.

Lehmann, L. A., R. E. Alvarez, A. Macovski, W. R. Brody, N. J. Pelc, S. J. Riederer and A. L. Hall.1981. Generalized image combinations in dual KVP digital radiography. *Med. Phys.* 8(5):659-667.

Lemacks, M. R., S. C. Kappadath, C. C. Shaw, X. Liu, G. J. Whitman. A dual-energy subtraction technique for microcalcification imaging in digital mammography--a signal-to-noise analysis. *Med Phys*. 2002 Aug;29(8):1739-51.

Lubbert, D., T. Baumbach, J. Hartwig, E. Boller, and E. Pernot.2000. μm-resolved high resolution X-ray diffraction imaging for semiconductor quality control. *Nuclear Instruments and Methods in Physics Research B* 160:521-527.

Malik, J., and P. Perona. 1990. Preattentive texture discrimination with early vision mechanisms. *J. Opt. Soc. Am. A* 7(5): 923-932.

Molloi, S., A. Ersahin and Y. Qian. 1993. CCD camera for dual-energy digital subtraction angiography. *IEEE Transactions on Medical Imaging.* 14(4):747-752.

Morales, R. A. 2002. Changes in identification and control of physical hazards since PR/HACCP. *USDA/FSIS Technical Conference on foreign material contaminants, prerequisite programs, and validation*. Omaha, NE.

Morita, K., S. Tanaka, C. Thai and Y. Ogawa. 1997. Development of soft X-ray imaging for detecting internal defects in food and agricultural products. *In Proceedings from the Sensors for Nondestructive Testing Int. Conf.*, 305:315. Orlando, FL.

Moses, S. 2000. Dual Energy XRAY Absorptiometry.[Online] available www:http://www.fpnotebook.com/RHE104.htm.

Nilsson, H. E., E. Dubari, M. Hjelm and K. Bertilsson. 2002. Simulation of photon and charge transport in X-ray imaging semiconductor sensors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 487(1-2): 151-162 .

Otsu, N. 1979. A Threshold selection method from grey-Level histograms. *IEEE Trans. Systems, Man, and Cybernetics* 9(1):377-393.

Pace, G., M. Cacciani, P. G. Calisse, A. D. Sarra, G. Fiocco, D. Fua, L. Rinaldi, and S. Casadio. 2003. Observation of polar stratospheric clouds with the ABLELIDAR during the APE-POLECAT fight of January 9, 1997. *Aerosol Science* 34:801-814.

Papanicolopoulos, C., J. C. Wyvill, W. Dayley and W. Owens. 1992. A feasibility study relative to the development of an on-line X-ray system that accurately detects bone and cartilage fragments in poultry meat. *Final report to Southeastern Poultry & Egg Association*, *SPEA research project #80*.

Papanicolopoulos, C. D., J. C. Wyvill, W. D. R. Daley, W. R. Owens. 1995. X-ray monitoring system. U.S. Patent 5,428,657.

Ratnam, M. M., C. P. Lim and A. N. Khor. 2001. Three-dimensional object classification using shadow moiré and neural network. *Opt. Eng*. 40(9):2036-2040.

Roh,Y. J., W. S. Park and H. Cho. 2003.Correcting image distortion in the X-ray digital tomosynthesis system for PCB solder joint inspection*, Image and Vision Computing*. 21(12):1063-1075.

Rosenfeld, A. and A. C. Kak. 1982. Digital picture processing. San Diego, CA: Academic Press.

Sandborg, M. and G. A. Carlsson. 1992. Influence of X-ray energy spectrum, contrasting detail and detector on the signal-to-noise ratio (SNR) and detective quantum efficiency(DQE) in projection radiography. *Phys. Med. Biol.* 37(6): 1245-1263.

Schatzki, T. F., R. Young, R. P. Haff, J. G. Eye and G. R. Wright. 1996. Visual detection of particulates in X-ray images of processed meat products. *Optical Engineering* 35(8):2286-2291.

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. In *Proceedings from the Sensors for nondestructive Testing Int. Conf.*, 161-171. Orlando, FL.

Selman, J. 1993. The fundamentals of X-ray and radium physics. Springfield, IL: Charles C Thomas.

Shahin, M. A., E. W. Tollner, M. D. Evans and H. R. Arabnia. 1999. Watercore features for sorting red delicious apples: a statistical approach. *Transactions of the ASAE* 42(6): 1889-1896.

Shahin, M. A., E. W. Tollner, R. D. Gitaitis, D. R. Sumner, and B. W. Maw. 2002A. Classification of sweet onions based on internal defects using image processing and neural network techniques. *Transactions of the ASAE* 45(5): 1613-618.

Shahin, M. A., E. W. Tollner, R. W. McClendon, and H. R. Arabnia. 2002B. Apple classification based on surface bruises using image processing and neural networks. *Transactions of the ASAE* 45(5):1619-627.

Silberschatz, A., P. B. Galvin, and G. Gagne. 2003. Operating system concepts. 6th edition. New York, N.Y.: John, Wiley & Sons Inc.

Sjodahl, M. and P. Synnergren.1999. Measurement of shape by using projected random patterns and temporal digital speckle photography. *Appl. Opt*. 38(10):1990-1997.

Smith, D. P., 2001. Defects of pre- and post-deboned broiler breast. *Journal of Applied Poultry Research* 10:33-40.

Sorgel, W., and R. J. Schalkoff. 1997. Efficient and accurate structured light algorithm for direct measurement of cylindrical surface parameters. *Opt. Eng*. 36(1):73-74.

Spartiotis, K., J. Pyyhtiä, T. Schulman, K. Puhakka and K> Muukkonen. 2003. A directly converting high-resolution intra-oral X-ray imaging sensor, *Nuclear Instruments and Methods in Physics Research Section* 501 (2-3):594-601.

Speller, R. D. , G. J. Royle, F. A. Triantis, N. Manthos, P.F. Van der Stelt, M. di Valentin. 2001. Digital X-ray imaging using silicon microstrip detectors: a design study. *Nuclear Instruments and Methods in Physics Research A* 457(3):653-664.

Texas Instruments. 1988. Quadruple 2-input positive-and gates. [Online]. available WWW: http://www-s.ti.com/sc/ds/sn74ls08.pdf.

Thomas, P., A. Kannan, V. H. Degwekar and M. S. Ramamurthy. 1995. Non-destructive detection of seed weevil-infested mango fruits by X-ray imaging, *Postharvest Biology and Technology* 5(1/2):161-165.

Tollner, E. W. 1993. X-ray technology for detecting physical quality attributes in agricultural produce. *Postharvest News and Information* 4(6):3-10.

Toyooka, S. and Y. Iwasa. 1986. Automatic profilometry of 3D diffuse objects by spatial phase detection. *Appl. Opt*. 25:1630-1633.

Tuffanelli, A., S. Fabbri, A. Sarnelli, A. Taibi and M. Gambaccini. 2002. Evaluation of a dichromatic X-ray source for dual-energy imaging in mammography. *Nuclear Instruments and Methods in Physics Research Section A* 489(1-3):509-518.

Tyson, J. 2003. How Airport Security Works. [Online] available www:http://travel.howstuffworks.com/airport-security2.htm.

USDA, 1996. Food consumption, prices, and expenditures, 41. Economic Research Service. Washington, D.C.: USDA.

USDA. 2002. United States classes, standards, and grades for poultry. Washington, D.C.:USDA.

Varian. 2003. Digital Imaging: PaxScan 4030.[Online] available www:http://www.varian.com/sip/dig102.html.

Wolaver, D. H. 1991. Phase-locked loop circuit design. Englewood Cliffs, NJ: Prentice Hall.

Xu, B., S. Lin and N. Jiang. 2001. Measurement of yarn densities in fabrics using phase-shifted digital moiré. *Opt. Eng*. 40(8):1476-1480.

Zheng, Q. and R. Chelleppa. 1991. Estimation of illuminant direction, albedo, and shape from shading, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13(7):680-702.

Zwiggelaar, R., C. R. Bull, and M. J. Mooney. 1996. X-ray simulations for imaging applications in the agricultural and food industries. *J. agric. Engng Res.* 63:161-170.

Zwiggelaar, R., C. R. Bull, M. J. Mooney, and S. Czarnes. 1997. The detection of "soft" materials by selective energy Xray transmission imaging and computer tomography. *J. agric. Engng Res*. 66:203-212.