UNIVERSITY OF CALIFORNIA, SAN DIEGO

Performance Enhancement Approaches for a Dual Energy X-ray Imaging System

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Electrical Engineering

(Signal and Image Processing)

by

Kenneth Fu

Committee in charge:

Professor Pankaj Das, Chair
Professor Clark Guest, Co-Chair
Professor Bill Hodgkiss
Professor Jan Kleissl
Professor Truong Nguyen

2010

The Dissertation of Kenneth Fu is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Chair

UNIVERSITY OF CALIFORNIA, SAN DIEGO
2010

DEDICATION

To my family

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This thesis would not have been possible without the support and encouragement of many people. To the following people, I owe an enormous debt of gratitude.

I would like to acknowledge Professor Pankaj Das for his support as the chair of my committee and Prof. Clark Guest for his support as the co-chair of my committee and the lengthy time spending on reviewing all my writing. Through multiple drafts and many long nights, their guidance has proved to be invaluable. My thanks also goes to the other members who serve on my committee: Prof. Truong Nguyen, Prof. William Hodgkiss, and Prof. Jan Kleissl

I would also like to acknowledge my co-workers at SAIC, without whom my research would not have happened. Thanks to Barbara Gardner, Rex Richardson, James Shea, Myron Mandell, and Dale Ranta. Their support helped me in an immeasurable way. I also want to express my gratitude to the Department of Homeland Security for authorizing the use of data for my research and publications.

Chapter 4, in part, is a reprint of the material as it appears in the Proceedings of SPIE (2009), Fu, Kenneth; Ranta, Dale; Guest, Clark; Das, Pankaj "A novel algorithm for material discrimination using a dual energy imaging system." The author of this dissertation is the primary investigator and author of this paper.

Chapter 5, in part, is adapted from the material as it appears in the Proceedings of SPIE (2009), Fu, Kenneth; Guest, Clark; Das, Pankaj "Segmentation of suspicious objects in an X-ray image using automated region filling approach." The author of this dissertation is the primary investigator and author of this paper.

Chapter 6, in part, is a reprint of the material as it appears in the Proceedings of SPIE/IS&T (2010), Fu, Kenneth; Ranta, Dale; Guest, Clark; Das, Pankaj "Layer separation for material discrimination cargo imaging system." The author of this dissertation is the primary investigator and author of this paper.

Chapter 7, in part, is adapted from the material as it appears in the Proceedings of SPIE/IS&T (2010), Fu, Kenneth; Ranta, Dale; Guest, Clark; Das, Pankaj "The application of wavelet denoising in material discrimination system." The author of this dissertation is the primary investigator and author of this paper.

Working on my PhD while having a full-time job has been very challenging to me both mentally and physically. It was my family's and my girlfriend's being unconditionally supportive that helped me get through the hardest times. I could never say enough to thank them. I believe the completion of this thesis is the best way to repay the people whose support gave me the strength to go on to the last day.

VITA

1997	B.S. in Physics, Nation Taiwan University, Taiwan, R.O.C
1999	M.S. in Physics, Nation Taiwan University, Taiwan, R.O.C
1997-2001	Musician, songwriter/assistant producer, Taiwan, R.O.C
2004	M.S. in Electrical and Computer engineering, University of California, San Diego
2005	Graduate Students Researcher, Computational Neurobiology Lab, Salk Institute, La Jolla, CA.
2007-2010	Signal and Image Processing Scientist, Science and Application International Corporation (SAIC)
2010	Ph.D. in Electrical and Computer engineering, University of California, San Diego.

Publications

Fu, Kenneth; Ranta, Dale; Guest, Clark; Das, Pankaj, "Layer separation for material discrimination cargo imaging system", Proceedings of the SPIE, Image Processing: Machine Vision Applications Volume 7538, pp. 75380Y-75380Y-12 (2010). Presentation was given San Jose, CA.

Fu, Kenneth; Ranta, Dale; Guest, Clark; Das, Pankaj, "The application of wavelet denoising in material discrimination system," Image Processing: Machine Vision Applications Volume 7538, pp. 75380Z-75380Z-12 (2010). Presentation was given San Jose, CA.

Fu, Kenneth; Ranta, Dale; Guest, Clark; Das, Pankaj," A novel algorithm for material discrimination using a dual energy imaging system," Proc. SPIE, Volume 7445, pp. 744513-744513-11 (2009). Presentation was given San Diego, CA.

Fu, Kenneth; Guest, Clark; Das, Pankaj; "Segmentation of suspicious objects in an X-ray image using automated region filling approach," Signal and Data Processing of Small Targets 2009, Proc. SPIE, Volume 7445, pp. 744510-744510-12 (2009). Presentation was given San Diego, CA.

Analog to Digital Conversion Using Recurrent Spiking Neural Networks, Society for Neuroscience 34th Annual meeting, *Abstract and presentation*, San Diego, CA, USA, 2004.

ABSTRACT OF THE DISSERTATION

Performance Enhancement Approaches for a Dual Energy X-ray Imaging System

by

Kenneth Fu

Doctor of Philosophy in Electrical Engineering
(Signal and Image Processing)

University of California, San Diego, 2010

Professor Pankaj Das, Chair
Professor Clark Guest, Co-Chair

Dual energy imaging is a technique whereby an object is scanned with X-rays of two levels of energies to extract information about the object's atomic composition (Z). This technique is based on the fact that the X-ray absorption coefficient decreases with X-ray energy for low-Z materials, but begins to increase for high-Z materials due to the onset of pair production. Methods using the ratio of the attenuations for high-energy to low-energy images as an indicator of Z value have been proposed by several people. However, the statistical errors associated with the systems make those indicators unreliable. This thesis will discuss the problems associated with using a dual-energy system for high-atomic-number material (also known as high Z material) detection. We will identify the sources of noise that hinder system performance and propose solutions for noise reduction. Later chapters will deal with methods to automate the high Z detection process. We use a method called adaptive masking to identify possible high Z objects and reduce the false alarms. For objects shielded by materials common in a cargo container, we propose a layer separation approach to estimate the ratio of the high-and low-energy attenuations of the shielded objects. The approaches provided in this thesis are able to enhance the detection rate and reduce the false alarms significantly.

1
Introduction

With the advance of technology, the world we live in has become a global village—transportation and communication are much more convenient. Unfortunately, it also makes the world more easily flooded with various kinds of extremism and terrorism on a scale we have never seen before. It is therefore extremely important to develop security systems that inspect luggage and cargo containers at customs and other checkpoints with great efficiency. Radiographic technologies have been applied to security systems in past decades, scanning luggage with X-ray beams so that inspectors can visually detect contraband objects by looking at the X-ray images. However, some contraband materials can be made into any size and shape, making the detection of those targets a challenging task. Several material discrimination techniques have been used in the airport to detect certain materials, for example, metal detection.

During the past decade, the security industry has developed dual-energy X-ray imaging systems to discriminate certain materials, such as a metal handgun or plastic explosives in baggage. Dual energy (DE) imaging is a technique that uses two different energies of X-ray beams to scan the same object and extract information about the atomic composition of the scanned object. It is widely used in the medical field to distinguish tissues and bones from a single X-ray image. Airport security systems first used the technology to detect metal (mostly to detect handguns) and later extended that technology to detect organic materials to look for plastic explosives.

Not until recently has there been interest in extending such technology to high-energy (mega volt) X-ray scanning for cargo containers. One reason for this is that there seemed to be no need for it. The other reason is that the technology was not ready to meet the challenges. Some theoretical analyses were published to discuss possible methods for cargo container content material discrimination. However, all of them were doomed to fail because of the huge statistical errors of the systems.

The objective of this thesis is to provide solutions to the problems of cargo inspection imaging systems using digital signal and image processing approaches, especially for the realization of a practical dual energy imaging system for material discrimination.

1.1 Motivation

After September 2001, many countries started to pay more attention to national security and possible terrorist attacks. An advanced security system to detect contraband material in airports and sea ports became a crucial factor for ensuring the safety of our nation. The security industry started to exploit the possibilities of using mega-voltage dual-energy X-ray imaging systems for material discrimination in cargo container inspection in customs. There are two major issues regarding this technology: first, the processing speed must be fast to be practical. The second is the accuracy of the system. Those two objectives usually contradict each other. A fast scanning time usually means fewer samples are taken during the data acquisition process, which usually leads to a higher error rate and smaller signal to noise ratio. Also, resolution is another key factor in the performance. Low resolution will not only make fine objects indiscernible but also introduce a greater number of errors in images. On the other hand, a sophisticated image processing algorithm is usually time-consuming. Therefore, a high processing speed usually sacrifices the image quality and accuracy in material discrimination. However, a long processing time, including scanning the object at a very slow speed, using highly sophisticated signal and image processing algorithms to denoise the image, to improve the image quality, and to automatically segment suspicious objects, will make the task unacceptably inefficient and

impractical. Thousands of containers are waiting at customs each day. Thirty seconds to a couple of minutes is probability the time available to scan the cargo and process the data to keep traffic flowing at an acceptable rate. Therefore, we are interested in looking for approaches that can achieve high speed and high accuracy without suffering from their tradeoffs. At least, a goal should be set to find an optimal combination of processing time and accuracy.

Using a dual energy system to extract the exact atomic make-up of the content in a cargo container is nearly impossible in practice. Depending on the nature of the containers, it is possible that X-rays cannot pass through all of the objects. In this case, no information about the object is attainable. Also, due to statistical errors that hinder the precision of measurement, it is difficult to obtain the exact effective atomic numbers for the objects, since atomic number is calculated from the measurements of X-ray attenuation which is very sensitive to small errors. However, dual energy imaging still has value in cargo imaging systems because the exact atomic information is actually not necessary. Radioactive materials that could be made into weapons all have very high atomic number. Also, they are usually shielded by lead to avoid the leakage of radioactive rays. A commercial cargo container usually won't have its items made up of material having atomic number greater than steel (mostly iron, with a small percentage of other alloying elements such as magnesium, chromium, vanadium and tungsten). Other possible common metals inside a commercial cargo container are nickel and copper. All the common metals have atomic number (Z) less than the atomic number of lead (Z=82) or tungsten (Z=74). Therefore, the task becomes much easier if the goal is just to discriminate objects that contain a high percentage of metals with atomic numbers higher than that of lead. Those materials are defined as high Z materials in this thesis. The task of a DE imaging system can be simply to divide the image into two classes: high Z and non-high Z materials.

Even though the task can be simplified, it remains very challenging. System noise, mostly Poisson noise, will introduce uncertainty in the counts of photons. Consequently, false classification can be made. Some non-high Z metal, because their

atomic numbers are so close to that of high Z materials, will be classified as high Z material as result of noise. Also, digital imaging artifacts and the geometry of objects will also contribute to false classification. Recently, a great amount of resources have been designated in research for using dual-energy image in nuclear-active material detection. Different teams have proposed different approaches for determining the atomic numbers of scanned objects. All of them, however, are facing problems described above.

1.2 Objective

Our objective in this thesis is developing image processing algorithms to enhance the quality of the image and increase the high Z detection rate. The ideal goal is to reach a 99% detection rate (for each one hundred scans of the same high Z object, less than one time do we have a miss). Conversely, the false alarm rate is computed by counting the number of events out of one hundred scans, a non-high Z object is classified as a high Z object. Although we would rather have a false alarm than miss, we still want to keep the false alarm rate to a minimum (less than 3 %). That is a very ambitious goal at the current time. To achieve such a goal, the dosage of X-rays must be high enough to provide significant information. Also, a very delicate detector that is sensitive and can provide a fine resolution is desired. Those factors will compromise the objective of radioactive safety and efficient processing time. Regardless of whether the obstacles created by the hardware can be resolved by an improved design, we have tried to provide possible software solutions for most of the problems that may or may not have hardware solutions.

1.3 Contributions

Dual-energy imaging for cargo inspection is a recently emerging technology not many publications about high-energy dual-energy imaging systems for the application of cargo imaging systems can be found. Most of the publications only discuss the physical principles of the system without mentioning too much of the engineering aspects. Most papers that do discuss the engineering aspects of this topic either address the problems without providing a solution, or provide solutions that only work under a very ideally controlled condition. The solutions provided in this thesis are therefore preliminary but original. Over the past two years, the technology of dual-energy imaging for cargo inspection has faced many criticisms. The critics have doubted about the applicability of using such technology in screening for the nuclear threat, and asked for a cheaper and simpler substitute [93]. In responding to those criticisms, we believe that the development of any technique takes time. Our work has shown that there is still a great potential in such technology, though many challenges remain as well. We will discuss the relevant issues in the later chapters. The major contributions of this thesis can be summarized as follows:

1. A probabilistic model using a global ratio threshold for high Z material discrimination has been developed. Several methods using dual-energy X-ray imaging system for high atomic number (Z) material detection are proposed. Until now, only a few methods for discriminating different Z number materials for cargo container inspection have been proposed. Among those, our methods give a very high detection rate that is better than other approaches. Most proposed methods look at the high to low energy attenuation ratio to determine the effective atomic number (effective Z) of scanned objects. We calculate the probability of a material having very high atomic number from the measured ratio, and make the determination based on the likelihood that the calculated ratio is actually above a threshold ratio value. The selection of the threshold ratio can be global or adaptive, depending on the properties of the imaging

system being used. A global threshold is used when the measured ratio is not a function of object thickness. In this case, the ratio of a material remains approximately the same as long as X-ray is able to penetrate the object. The selected ratio threshold is a value higher than the ratio of steel.

One problem with this global threshold probabilistic method is the determination of a threshold of probability, above which the target is called high Z. When a high Z object is buried in a non-high Z background, the effective ratio will decrease as the thickness of background increases. Consequently, the calculated probability will decrease, which can cause a detection miss. We are first to propose an approach that can resolve this issue. The proposed "adaptive masking approach" processes the X-ray image on a region by region basis. It does not use a fixed threshold of probability for the whole image, but determines the threshold adaptively based on the relative probability an object has to its regional background. A modified thresholding approach for image segmentation has been developed and employed. Not only does our approach resolve the problem of global threshold of ratio, it segments objects in all regions and uses them as "masks". Those masks are used to eliminate false alarm pixels caused by noise. The proposed masking approach is a very powerful tool in false alarm elimination. Our experimental results have shown a very high accuracy of automated high Z detection in simple test setups. The details of this approach will be provided in Chapter 4.

2. A simple and efficient method for data clustering is proposed. Image segmentation is another pillar of our project. By employing image segmentation, we can then focus on regions requiring attention, which saves a lot of processing time. A good data clustering method helps make any kind of segmentation more efficient. This thesis discusses the most commonly used segmentation approaches. It also discusses a new segmentation method we propose for detection of suspicious objects in an X-ray image. A modified clustering method, a mixture of leader clustering and k-mean, has been

developed and employed in our high Z detection system, which will also be discussed in Chapter 5 of this thesis.

3. An algorithm for region filling is proposed in Chapter 5. This is an algorithm to fill in the holes in an image. The standard way is to select a starting point, called the seed, and then to grow from that seed point until the hole is completely filled. Automated region filling is always a hard problem. The major challenge is to define the seed point from which to start the region growing process. We have developed an algorithm for region filling without the need to find the seed point but that still achieves the goal of automation.

4. An algorithm to detect high Z material hidden behind low Z material is proposed. Organic materials such as paper, or inorganic materials such as water, are classified as low Z materials. Based on our experiments, we found that the probability will not only be affected by the shielding metals, which are classified as mid-Z materials, but will also be greatly affect by low Z materials. The calculated probability of a high Z target will be significantly lower than it should be if it is put behind a box of drinking water. We propose an algorithm to determine the locations of low-Z- surrounded objects and modify the calculated probability. This algorithm greatly improves the accuracy of high Z detection. This low Z background problem has not been reported in any publication, we are the first to define and resolve this problem. A model of using an adaptive threshold for the high Z ratio was also developed to replace the global threshold to improve the system performance.

5. In the end, a complete method for hidden high Z object detection is developed, including noise reduction, edge enhancement, data clustering and background layer separation. This complete scheme enables the design of DE system to achieve 99% of detection rate with very limited false alarms.

1.4 Structure of this thesis

This thesis presents several approaches to enhance the performance of a cargo dual energy system. Depending on the requirements and objectives of a system, some of the approaches are substitutable alternatives to others. Some of them are an integral part of others. For example, an idea we call “masking” is a key part to the complete scheme. Figure 1.1 shows the flow chart of the high Z detection scheme. The two X-ray images obtained from the system go through a series of stages to obtain the final result:

1. Noise reduction: enhance the quality of images and enhance the accuracy.

2. Classification: the process of generating ratio information to classify the materials of objects.

3. Discrimination: discriminate high Z material from non-high Z material.

4. Post processing: clean up the false positives generated from the previous stage. Machine error and noise caused by random scattering are dealt with this stage.

5. Decision making: A binary image we call “High Z map” or “Z-map” is produced to provide the detection result.

Chapters that discuss the relevant approaches are labeled in the block diagram.

Chapter 1 is the introduction of our objective, and a summary of our achievements. Chapter 2 of this thesis discusses the physical background of X-ray imaging. The background of dual energy imaging systems is also discussed at length. Chapter 3 provides the general background of imaging processing, including terminologies and commonly used approaches for image denoising and segmentation. Chapter 4 covers the first solution we developed for discriminating high Z materials for cargo imaging system. An algorithm including image enhancement, object segmentation and false alarm elimination is discussed. Chapter 5 goes one step further in image segmentation. A new data clustering approach and a new segmentation

method are proposed and discussed. Chapter 6 provides a solution for detecting threats behind different shielding. This is an important step that completes the whole picture. Chapter 7 introduces a new approach to image processing—wavelet domain processing. Chapter 8 draws conclusion from the current work and describes the future direction.

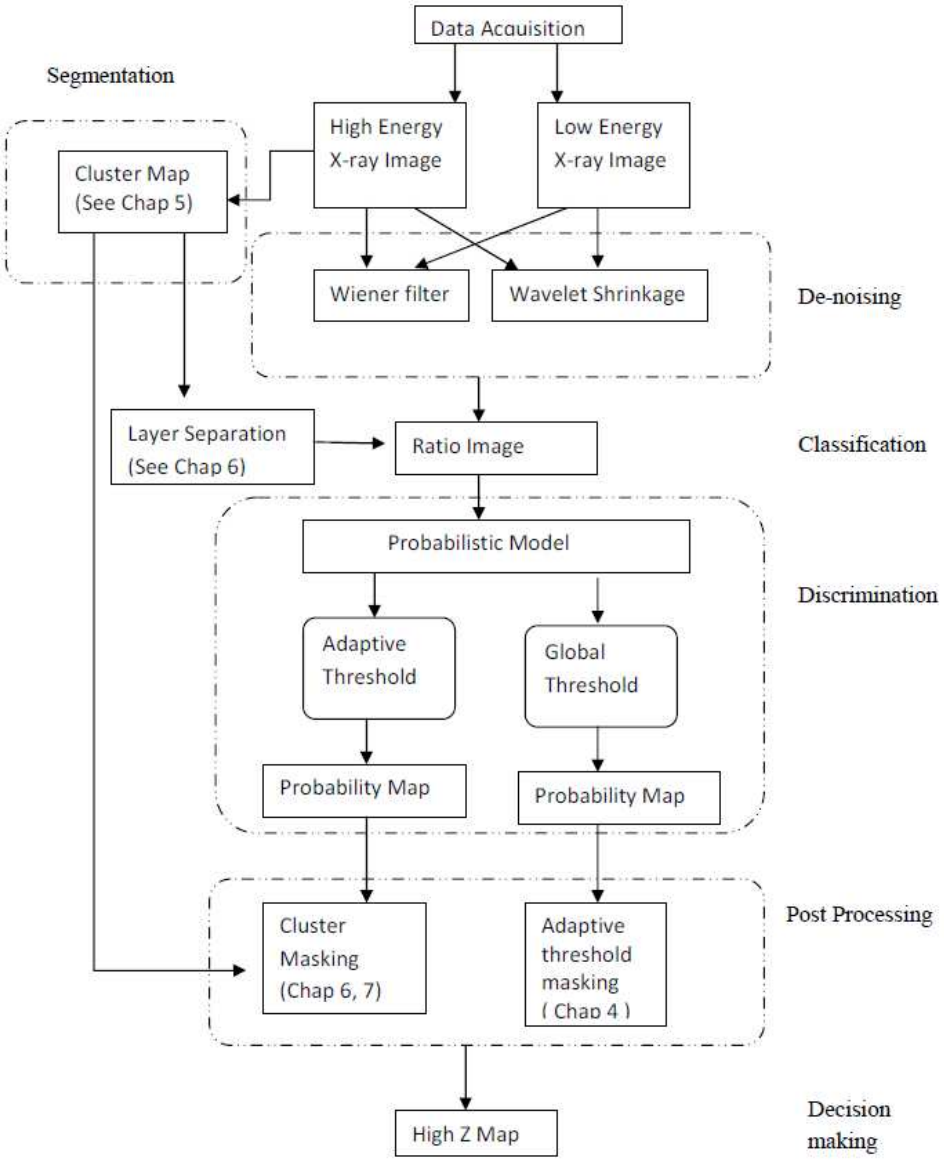


Figure 1.1 Performance enhancement scheme.

2

Background of Dual Energy Imaging

In this chapter, the theoretical background of dual energy X-ray imaging system is reviewed and the issues of such systems are addressed. The discussion is introduced with the physical background of X-ray imaging systems in Section 2.1. The theory and application of dual energy imaging are discussed in Section 2.2. And, in Section 2.3, the issues and challenges of designing a dual energy X-ray imaging system for material identification is discussed.

2.1 X-ray imaging

X-ray imaging works by placing an object between the X-ray source and the receptor (a film or a detector.) The physics of X-ray imaging can be explained by the exponential law for a narrow spectrum beam:

$$N = N_0 e^{-\mu t} \qquad (2.1.)$$

where N is the number of photons passing through the object per unit of area and N_0 is the number of photons in the incident X-ray beam, which can be obtained by measuring the number of photons passing through free air, μ is the attenuation coefficient, which is dependent on the material and the X-ray energy, and t is the thickness of the object being scanned in the beam path. The intensity of the X-ray beam in the equation is defined as the number of photons in the beam, which can be

obtained by counting the number of photons received by the detectors. As a beam of X-rays passes through an object, some of its photons will be absorbed by the probed material due to collisions between the photons and the molecules. As a result, the number of photons passing through the material will be less than the number of photons in the incident X-ray beam.

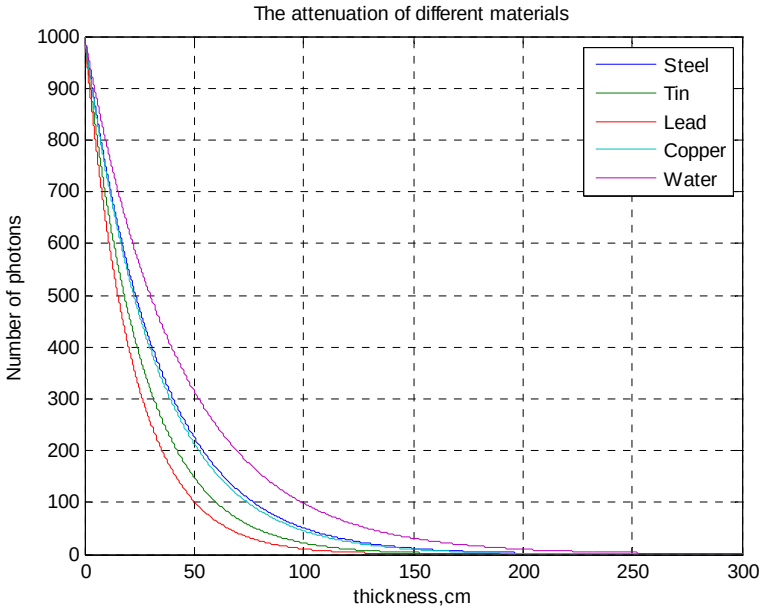


Figure 2.1 Attenuation of several materials. (Data obtained from NIST website)

Figure 2.1 shows the relation between the thickness of an object and the remaining number of photons after passing through several materials. The energy of X-rays here is 9MeV. The initial number of photons is N=1000. The attenuation coefficients of the selected materials are: $\mu_{steel}=2.34*10^{-1}$/cm, $\mu_{tin}=2.49*10^{-1}$/cm, $\mu_{lead}=5.16*10^{-1}$/cm, $\mu_{copper}=2.75*10^{-1}$/cm, and $\mu_{water}=2.31*10^{-2}$/cm. It is usually the case that materials with greater densities have greater attenuation coefficients. The attenuation coefficients determine the degree of X-ray absorption. In Figure 2.1, 30% of the 1000 photons pass through 50 cm of water. However, for 50-cm-thick lead, only one tenth of the photons can pass.

More specifically, Equation (2.1) is expressed as

$$N = N_0 e^{-\mu(E,Z)t}. \tag{2.2}$$

The attenuation coefficient μ is a function of density of the scanned material(ρ), the chemical composition of the material (atomic number, Z), and the photon energy, *E*. We can express this relationship with the following equation:

$$\mu = \rho * \tau(E,Z) \tag{2.3}$$

where τ is the mass attenuation (cm^2/g) of the material.

Transparency or transmittance is defined as the ratio between the intensities of the X-ray beam before and after the penetration of an object. It is a measurement of the transmission of X-rays. It can be expressed as a function of E, t, and Z.:

$$Tr(E,t,Z) = \frac{I}{I_0} = e^{-\mu(E,Z)t}. \tag{2.4}$$

Some people prefer to define transparency as the logarithm of Equation (2.4). To avoid confusion, we call that definition the logarithm transparency which is defined as

$$\ln Tr(E,t,Z) = \ln \frac{I}{I_0} = -\mu(E,Z)t. \tag{2.5}$$

The converse concept of transmission is attenuation. It defines the ability to absorb photons of a material. The inverse value of logarithm transparency is called attenuation or absorption. The attenuation $m(E,Z,t)$ is expressed as:

$$m(E,Z,t) = \ln\left(\frac{1}{Tr(E,Z)}\right) = \mu(E,Z)t\ . \tag{2.6}$$

Equation (2.6) indicates that the greater the attenuation coefficient μ or the thicker the object, the more photons will be absorbed in the path.

So far, our discussion is only focused on narrow band X-rays. In practical radiography systems, the X-rays are not monochromatic but have a continuous spectrum.(polychromatic) The spectrum of an X-ray beam is a distribution between

zero and the maximum photon energy(end point energy). Therefore, Equation (2.1) should be modified as

$$I(E,t,Z) = \int_0^{E_0} N(E)\, e^{-\mu(E,Z)t}\, P_d(E) E dE \tag{2.7}$$

where $N(E)$is the number of photons in the spectrum with a peak energy E_0 at energy E, and $P_d(E)$ is the photon probability density function (i.e. the probability that a photon has energy *E*.)

Traditional X-ray imaging works by placing object between X-ray source and a photosensitive film. The intensity of the X-rays is modified by the absorption of the object. X-ray images are negative images, similar to developed photographic films. An X-ray image gives information about the relative thickness and density of the object being scanned. Thicker or denser portions of an object will have greater opacity (less transparency), the X-ray intensity on the corresponding part of the film or detector will be less, which makes that region of the image appear brighter.

Currently, most X-ray imaging systems are digital. For digital X-ray images, two methods are used: (1) by digitizing the X-ray film, or (2) by having the X-rays passing through the object fall directly onto a device that converts X-ray to light (such as a phosphor screen), and then the light signal is captured by a light-sensitive digitizing system.

2.2 Cargo X-ray imaging

2.2.1 General Description

In Section 2.1, we have discussed the general principles of X-ray imaging. In this section, we will discuss one application of X-ray imaging—the cargo imaging system. A good reference for cargo imaging system can be found in Chen's "Understanding Cargo Imaging [1]". Figure 2.2 shows the arrangement of a cargo X-ray imaging system. A typical cargo imaging system has a source, usually mobile, that generates X-ray photons; a detector array, usually L-shaped, that records the X-ray

photons; and a signal processor unit doing signal and image processing computations such as analog to digital conversion, contrast enhancement, noise reduction or other higher level image processing. The operator display presents the output to the users.

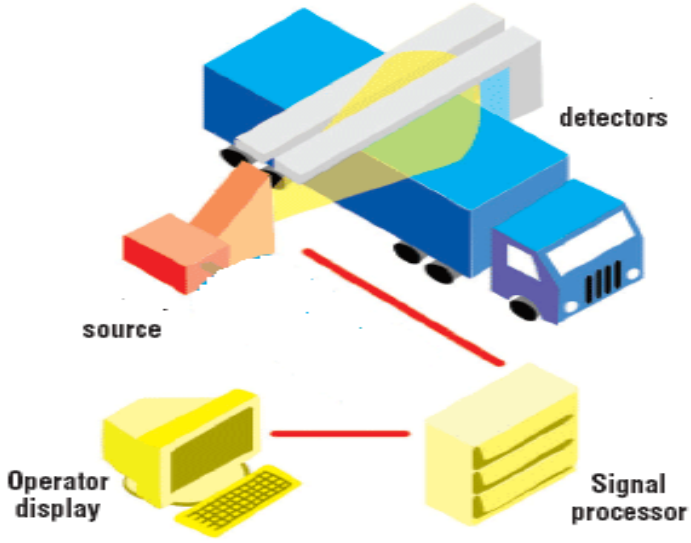


Figure 2.2 Setup of a cargo imaging system.

More detailed discussions of the components of a cargo imaging system are given in the following:

***Source*:** The source is a device that generates X-ray photons. The energy of the source refers to the maximum energy a photon can have when emitted from the source. A pulsed source generates X-ray beams at a certain frequency; each beam generated is a pulse. For a cargo X-ray imaging system, the X-rays need to have enough energy to penetrate cargo containers made of thick steel. For X-ray energy up to 500 keV, the source is usually an X-ray tube – a DC potential type electron accelerator plus a tungsten target. For higher energy, a Linac (linear accelerator) is used to accelerate electrons that generate X-rays at a tungsten target. A higher energy system has more intense X-rays. For the purpose of material discrimination, the source emits X-ray pulses at two energy levels. The high and low energy sources will be placed in an interleaved arrangement. This is called the dual energy system which will be discussed in later sections. For safety concerns, the X-ray energy must not exceed 9

MeV. Higher energy sources will have significant neutron production and photonuclear reactions can be a concern.

Detectors: most cargo imaging systems use scintillator –photodiode detectors. They can be Caesium Iodide (CsI) or Bismuth Germanate **(**BGO) for low energy and low dose rate systems, but Cadmium Tungstate (CWO) is the usual choice in high energy systems because it has a relative high photon yield. When the photodiode is properly protected from the direct beam, such detectors have proven to work for many years without noticeable performance degradation.

A cargo imaging system usually uses a linear array of detectors. As shown in Figure 2.2, the detectors are aligned in an L-shape detector tower. Each detector represents a pixel in the digital X-ray image. Detector size is a compromise of the desire to display smaller features and the need for a larger area to capture more photons. The detector usually needs to be thick enough to absorb and detect most of the X-rays reaching it. For the systems with the purpose of material discrimination, each pixel often has two detectors, one selectively records lower energy photons and the other selectively records higher energy photons. This information is used to identify the material's effective atomic number.

Data acquisition electronics include detector amplifiers, sampling and holding and analog to digital conversion. They usually have 16-bit or higher resolution and the electronic noise needs to be small enough to give good overall system dynamic range.

***Signal Processor*:** A computer equipped with an image processing algorithm to process data from detectors is the heart of the system. Processing may include detector normalization, using calibration data to convert intensity of detected photons to thickness relative to the calibration material. Steel and copper are the most commonly used calibration materials; source variation is corrected by calculating the standard deviation of the calibration data, rejection of bad data, image manipulation and analysis. A practical system must be efficient to keep the commerce flow at a reasonable speed. So far, no standard is given for the processing speed. However, according to the statistics given by the American Association of Port Authorities (AAPA) website[94], about 300.000 to 370,000 containers will be loaded a month in

Long Beach port. With about 10.000 containers waiting to be loaded per day, a processing time of less than 10 seconds per container is desirable. .

***Operator display*:** The result of processing is displayed to the operator in the chosen format. Operators can inspect items inside the truck by looking at the processed X-ray image. They can also process the image for their specific purpose with preset functions. For example, they should be able to enhance the contrast of the image for their convenience, or to zoom in on any particular area in the image that is suspicious to them. For the purpose of material discrimination, an overlay of alarmed regions and the X-ray image is provided so the inspectors can use that information to take further actions.

Other important components include the collimator and shielding. Collimation can be done in both source's side and detector's side. The shielding is necessary for safety concerns.

Collimator: The collimator is used to control photon scattering. The source generates X-ray photons that travel in all directions. It is desirable to have at least one source collimator to form a slice of beam. A detector collimator may also be used to reduce detection of scattered X-rays and therefore to improve system dynamic range. In many cases, especially when there is no closed shielding, a beam filter is desirable to remove these unwanted photons –the lower energy portion of the X-ray spectrum.

***Shielding*:** For safety reasons, shielding is needed to prevent photon from leaking. Shielding varies from a few millimeters of lead in a 160 keV system to a massive steel and concrete installation in a 9 MeV system. External radiation levels must meet local regulations. For 9 MeV systems, steel and/or concrete is necessary because there is also a need to shield from the neutron byproducts. The higher the X-ray dose, the heavier the shielding will be. Heavy shielding will slow down the speed and/or consume more energy, and requires more physical space. It is therefore preferable to find the balance between the size of shielding and X-ray dose to give the maximum efficiency. The dose delivered to cargo is limited by beam collimation and filtering. An interlock system is included to prevent accidental radiation exposure.

2.2.2 Terminology

In this subsection, we introduce several often used terminologies in cargo imaging systems. The performance of a system is defined by those terminologies.

***Half-value layer* (HAL)**: It refers to the thickness of a specified material that reduces the intensity of the incoming radiation by half. HAL refers to the thickness of the first half-value layer. The subsequent HALs are the amount of specified material that will reduce the radiation rate by one-half after material has been inserted into the beam that is equal to the sum of all previous half-value layers. A good imaging system can have X-ray penetration up to fourteen half-value layers.

***Pixel noise*:** Error caused the distortion of information of a pixel. The sources of the noise can either be X-ray system noise, such as quantum noise, or the quantization noise when forming a pixel.

***Maximum penetration**:* This is one of the most important performance parameters. In cargo imaging, this is usually defined as the maximum steel thickness behind which a lead block can still be seen. For objects thicker than that, the steel would be completely dark in the image, which is usually referred to as a non-penetrable region. After penetrating that maximum thickness steel, the remaining signal must not be washed out by pixel noise. Under such conditions, the pixel noise sources are primarily electronic noise, uneven scatter, and artifacts, such as those generated by inadequate normalization in either direction. X-ray quantum noise, or statistical noise, is insignificant in such a situation because of the low photon counts. The maximum penetration of cargo imaging systems ranges from about 25 mm for 160 keV systems to more than 400 mm for 9 MeV systems. Fig. 2.3 shows the typical maximum penetration of different systems.

***Signal Contrast*:** This is the visibility of items against a background, and is highly operator dependent. To visually detect an item with a background, the signal of that item must have enough signal contrast. In other words, a sufficient signal to noise ratio makes the item not buried by pixel noise. Generally speaking, for an area object, the signal contrast usually needs to be at least one-third of the standard deviation of

pixel noise. For line objects and point objects, the contrast needs to be one and four times the pixel noise, respectively.

Contrast sensitivity: This is defined as the ability to see a thin steel plate behind a thick steel plate, usually in the range of half of the maximum penetration. For example, a typical 450 keV system has 100 mm maximum steel penetration and its contrast sensitivity is 1–2 mm behind 50 mm steel. Contrast sensitivity measures the ability to unveil details in the real world. Signal contrast is the attenuation by the thin plate. Pixel noise is predominantly X-ray quantum noise; artifacts and electronic noise are usually negligible in such condition. Wire contrast defines the ability to detect a thin steel wire behind a thick steel plate. A typical 450 kV system can detect a 3 mm diameter steel wire behind 50 mm of steel.

Display resolution: Display resolution (pixel size) is determined by detector size and image magnification. It defines the ability to see small features under high signal contrast, low noise conditions. In cargo imaging, display resolution is usually a few millimeters and is tested with line pair grids made of 2 mm thick steel, with no background object (in air).

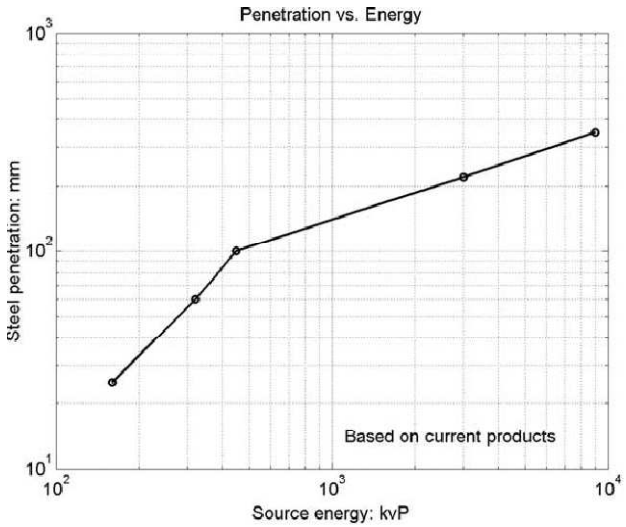


Figure 2.3 Typical maximum steel penetration of cargo imaging systems (image obtained from [1])

***Wire resolution*:** This is the ability to detect a small wire in air. The signal contrast is the attenuation of the wire, usually diluted because its shadow may not cover a whole detector element. The pixel noise is from X-ray quantum noise and artifacts. Low energy systems with small detectors often have 0.125 mm wire resolution and high energy systems usually have around 1 mm wire resolution.

2.2.3 Image Formation

The image is formed in three steps: object scanning, data recording, and information visualization.

Object Scanning is the first step of cargo imaging which allows X-rays to pass through the object. According to Equation (2.1), the number of photons is attenuated through the object. To obtain information about something inside the cargo container, there must be enough remaining photons after the beam travels through the object, and any item seen in the image must cause a noticeable difference in X-ray intensity, or else the contrast of the image will be too low to provide any information. As discussed before, the random scatter of photons must be controlled so that it does not add a strong background to cover weak signals, which means a low signal to noise ratio (SNR). Besides collimators, a beam filter can be used for scatter rejection. The best filter material for 9 MeV is polyethylene. For 3MeV X-rays, the best filter material is very high atomic number material, such as lead or uranium.

For lightly loaded cargo or small baggage, a 160keV system will provide adequate penetration and excellent contrast. At the other extreme, a fully loaded Arab boat (called a Dhow) requires a 9MeV system. However, due to the strong penetrating nature of 9 MeV X-rays, a small piece of explosive or a 0.125 mm diameter wire, which could be detected in a 160 kV system, would look almost transparent in the high energy system, and is very likely to be overlooked.

The source intensity, or beam current, determines the initial amount of photons. The number of photons emitted by the source is designed to be received by each of the detectors in the detector tower. Through calibration, the number of photons

each detector will be getting without attenuation is determined. After photons probe the objects, the remaining number of photons for each detector will be counted and later converted to pixel intensity through signal processor.

***Data Acquisition*:** This is the second step, which detects the X-rays that pass through the X-ray-probed objects. Detectors are placed in an array; each of them records information about the corresponding area. Therefore, the detector resolution sets the upper limit of spatial resolution that the imaging system can show, which is determined by the detector size. For systems such as airport security, the resolution can be as small as 1mm. Figure 2.4 illustrates how the pixels of the image are formed in vertical direction. The source emits a fan-shaped beam of photons through the object of height h. Without any attenuation, the photons will mostly be received by the N detectors. The number of photons received by each detector will be recorded as counts and converted to a normalized value by an A/D converter. As the source moves along the object, it generates pulses of beams, the number of unabsorbed photons of each pulse is recorded. The raw data is the record of all pulses of all detectors. In the horizontal direction, the raw data from all pulses within the bin width of resolution (usually the same as h/N) will be averaged or integrated by a chosen function to form a pixel.

Since the number of photons received by different detectors will not be the same, and each detector channel has different amplification, each detector must be normalized individually. For a pulsed X-ray source, the intensity of each pulse of X-rays should be normalized with reference detectors. The remaining error of the normalization causes vertical artifacts in the image, while the horizontal artifacts are caused by pulse variation compensation (i.e. averaging over pulses) in the image. Artifacts can also be introduced by discontinuities in detector layout (such as between detector modules) or other structures; this kind of artifact is the cause of what is called “edge effect” which will be addressed in later sections and chapters. Artifacts contribute to the overall pixel noise and should be reduced whenever possible.

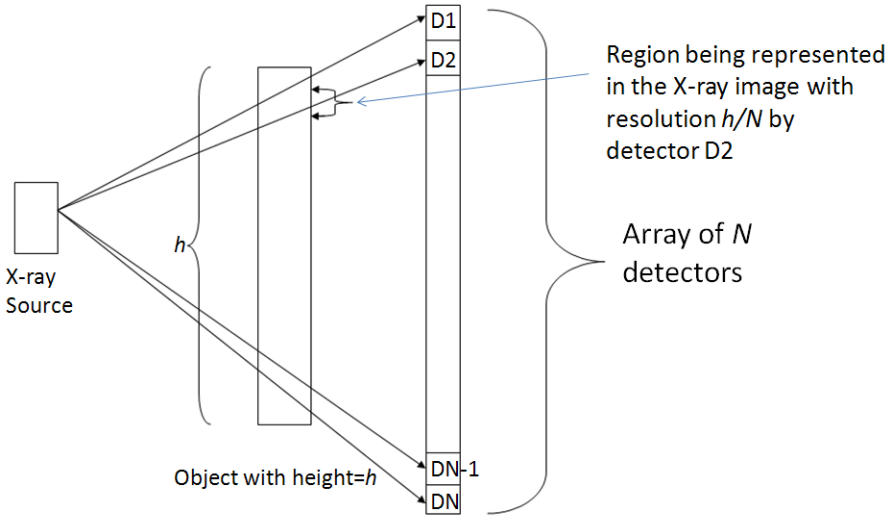


Figure 2.4 Digitization of the X-ray image. The X-rays from the source go to the *N* detectors. In the X-ray image, each pixel represents about *h*/*N* wide of an area.

***Visualization*:** Visualization refers to the process of presenting the information graphically to the system operator, which is the last step and final goal of the system. The task here is done by the signal processor unit of the system. The image data usually have 16 bits, or 64K levels. Gray scale display uses only 8 bits, or 256 levels and it cannot present all image information at one time. Pseudo-color uses 8 bits for each of the three basic colors, or 16 M levels. In this situation, the effectiveness is limited by the ability of human eyes to tell subtle color differences.

The intensities in an X-ray image can be presented in two ways: linearly and logarithmically. A linear display normalizes pixel data and linearly maps the data to the gray levels. As shown in Equation (2.1), the X-ray signal decreases exponentially with thickness. Most parts of the image would appear dark and it would be hard to discern any details. A better way to presents the data might be to map the logarithmic intensity, which is proportional to the thickness of the object, to the gray levels. However, the logarithm display does not show thin items very well compared with linear mapping.

Another way is to take the square root of pixel data and map them to 256 gray levels. By doing so, the contrast between pixels will be enhanced. Square root display is an equal-noise display method. That is, whether the signal is at 250 or 25 on the gray scale, the X-ray quantum noise corresponds to the same number of gray levels.

Noise reduction is usually built into these nonlinear transfer processes. That is, the logarithm or square root lookup table can be modified to suppress noise. The single most helpful process is AHE, or adaptive histogram equalization. After performing some non-linear transformation, the image is divided into smaller sections. In each section, pixels are assigned a new value so that the number of pixels that fall into each gray level is equal. Bilinear AHE usually provides the best overall appearance of the object, but true object thickness information is lost. Different areas may have very different gray level mapping. Fig. 2.5 demonstrates the effect of linear, logarithm and AHE displays. For the best result in a specific region, imaging software usually allows the operator to adjust display contrast. Due to the cluttered nature of cargo, complexity of the object, and the usually high pixel noise, edge enhancement techniques have very limited effectiveness in cargo imaging. For the same reason, automatic image analysis does not work very well either. There remain many challenging tasks for image processing and computer vision.

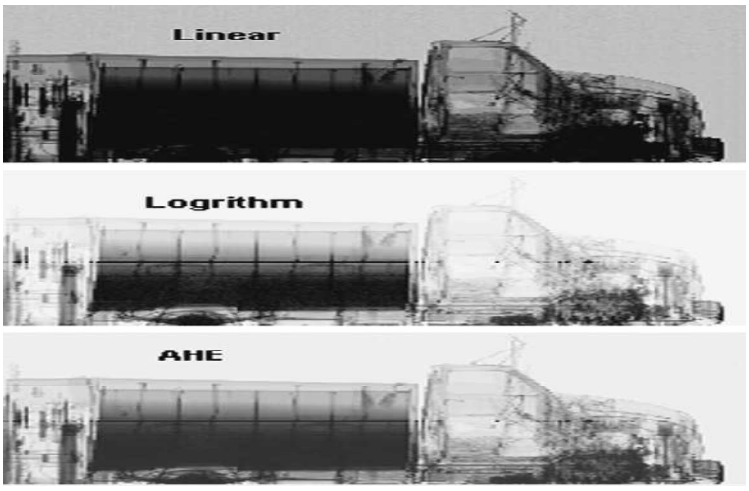


Figure 2.5 Three different displays of a cargo image (Image obtained from [1]) Top: data are linearly mapped to 256 grayscale levels. Middle: the logarithm of data is mapped to 256 grayscales. Bottom: after applying AHE to the data, the signal is mapped to grayscale.

2.2.4 Source of Noise

As mentioned earlier, sources of pixel noise can be physical such as statistical errors, scattering, and electronic noise, or can be digital—artifacts of signal processing.

Noise from an X-ray imaging system is mostly X-ray quantum noise, which is mainly shot noise. Shot noise is created by the fluctuation of detected photons. Shot noise is governed by the Poisson distribution. A major feature of the Poisson noise is that the standard deviation is proportional to the square root of number of events. In the low X-ray penetration situation, the pixel noise, or standard deviation of pixel signal, is slightly larger than the square root of the average number of photons per pixel. In this case, quantum noise is not a dominant noise. Whether the standard deviation of photon counts equals the square root of photons can be used as an indicator of the penetration quality.

Both maximum penetration and contrast sensitivity will suffer with reduced photons per pixel. Wire resolution, or the ability to detect a single wire smaller than the detector size, is not affected by detector size. When detector size increases, signal contrast caused by the wire is diluted, but pixel noise is reduced, and the two effects compensate each other. Low energy systems usually have detectors of high efficiency. For high energy systems, the detector thickness and therefore detection efficiency is usually limited by cost considerations. Typical CWO detector thickness in cargo imaging is 20 or 30 mm. Gas detectors are theoretically more radiation resistant but their low detection efficiency is a major disadvantage and are rejected by most of the industry.

Electronic noise, including that from photodiodes, amplifiers, and other circuits, plays a major role in determining the maximum penetration. It is the primary contributor to pixel noise in very low signal conditions, as in the case for testing maximum penetration. Everything else being equal, doubling electronic noise means twice as much remaining signal is needed to detect an item.

Crosstalk (XT) refers to a leaking or blending where a photon aiming for one detector is received by another detector. When this happens, the calculated attenuation

for a pixel will be smaller than it supposed to be. Crosstalk among detectors tends to blur images and needs to be controlled to within a few percent. Scintillation detectors naturally achieve this level as long as detector units are optically isolated. In tomography, where further reconstruction is needed, X-ray separators are essential. Crosstalk can also be corrected through software.

Pixel noise also comes from image processing artifacts. One commonly seen artifact is the edge effect--a “ghost” edge caused by various factors. The X-ray sources usually produce a fan-shape beam, meaning pixels receive X-rays from different angles. Figure 2.6 illustrates the difference that occurs with X-rays coming from different directions. When the incident beam is perpendicular to the surface of a uniformly thick object with thickness *t*, the detectors will measure the same thickness from all three X-rays. When the beams are coming at a different angle, different thickness for different rays will be read as a result. The first X-ray from the top passing through the shortest path will have the smallest attenuation among the three X-rays. If an object receives X-rays from different angles, there will be a penumbral effect around the edges of an object. The size of the detectors (resolution) will also affect the degree of edge effect. Because pixels are formed by averaging the pulse sample within the width of a detector, object edges falling within a detector will be smeared. This effect is more prominent when detectors are bigger (lower resolution.) As illustrated by Figure 2.7, an object with thickness *t* and length equivalent to three-detector width will not have much misreading when the signal from the object is evenly distributed to all three detectors as shown in (a). It will be read as 4 pixels wide when the object is not properly aligned with the detectors as shown in (b). Detector 3 and detector 4 will still have the reading *t*, while detector 2 and 4 will have readings from the object and the air. The pixel value of detector 2 and detector 5 will neither be *t* nor zero, but some value in between. The output image will be a 4-pixel-wide object with 2 pixels in the center darker than the edges. The problem can be improved by increasing the resolution. When the detector size becomes half of its original size as shown in (c), 6 pixels will have the same reading. Because cargo X-ray beams are fan-shaped, X-rays come in at different angles, and because of the artifacts associated

with digital X-ray imaging, the edge effect usually contributes to a source of misclassification in material discrimination systems.

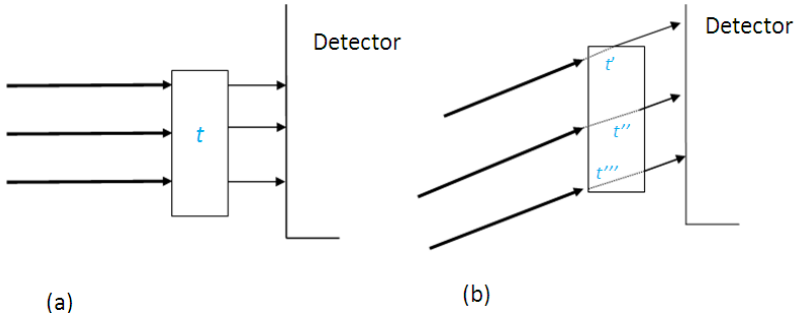


Figure 2.6 Edge effect illustrates (a) three identical X-rays projected on an object with thickness *t* perpendicularly. Without noise, the detectors will have the same reading for all three. (b)When the beam comes at a different angle, different detectors will have different readings since different rays go through different thicknesses.

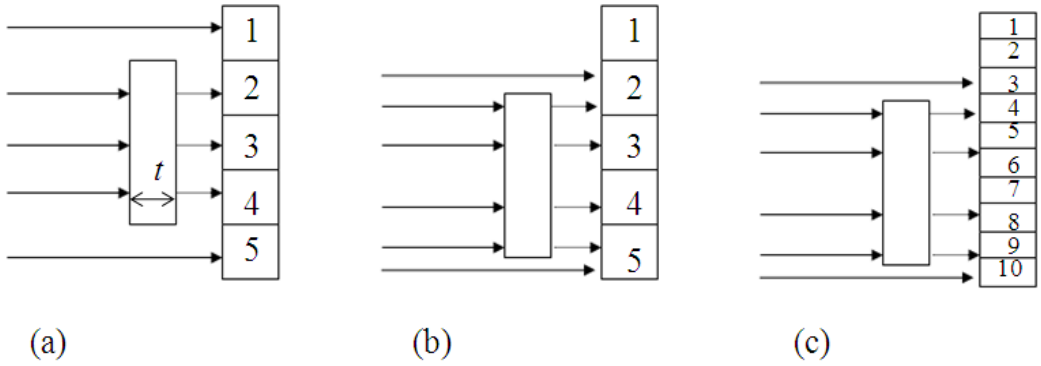


Figure 2.7 Edge artifact and resolution illustrate (a) an object with thickness *t* and length equivalent to three detector's size is placed aligned with detector 2, 3, and 4. They will have the same readings. (b) The edge of the object falls within detector 2 and 5; detector 3 and 4 will have different readings than detector 2 and 5. (c) When detector size becomes half of the original one, most parts of the object fall within detector 5, 6, 7, 8, and 9. Ideally, detectors 1, 2, 3 and 4 will not have a reading of thickness from the object.

2.3 Dual Energy X-ray Imaging

Recently, some cargo imaging systems offer material discrimination [87-91], usually by exploring the difference between the photoelectric effect and Compton scattering, and are common for 160, 320 and 450 keV systems. Mega-voltage imaging systems work by exploring the difference between Compton scattering and pair production. Higher energy X-rays will trigger more Compton scattering events, causeing a greater attenuation. As higher atomic number materials have higher photoelectric cross sections, the ratio of low energy photon attenuation to high energy photon attenuation tends to be larger. The technique that uses two levels of X-ray energy sources to exploit the difference in attenuation for material discrimination is called dual energy (DE) imaging. The details of the physics of such systems will be given in the next section.

Material discrimination systems usually classify objects into three groups: organic, inorganic and metal. More recently, scanners started to use a dual high energy X-ray sources at mega-volt levels to identify very high atomic number. The effectiveness of material discrimination is usually very limited in cargo imaging mainly due to dilution (caused by random scattering or absorption) along the beam path – the items of interest are often on a very complex background of other material.

Dual energy imaging uses two energy levels of X-rays for scanning an object to extract atomic information about the scanned object. This technique has been applied to the medical field since the 1960's to separate the tissue and bones in the X-ray images. Until recently, methods were proposed to use this technique in the security industry for threat target identification. Although the mechanism of the imaging system for airport security systems and cargo screening systems are not exactly the same, because the energies involved, they are very similar in general: both of them use the ratio of the measured attenuation from the two X-ray images to explore different mechanisms.

2.3.1 Physical Background

When the energy of the X-ray is below 200KeV, the interactions between the photons and the atoms of the object being scanned are dominated by two major effects: Compton scattering and the photoelectric effect. Each interaction is independent of the other, and is a function of photon energy and atomic number (Z). Incident photon energies can be absorbed by the scanned object. As a result, electrons are emitted from the atoms. This phenomenon is called the photoelectric effect. With photons interacting with objects, the collisions with the electrons of the material atoms will cause a photon to lose its energy. The photon will scatter at an angle to conserver the total momentum. It is quite possible that while it still has enough energy, the scattered photon will collide with another electron and lose energy again. This process is the known as the Compton scattering or the Compton Effect. The input ray gets attenuated by such a process. The mass attenuation in Equation (2.3) can be rewritten as:

$$\tau(E,Z) = \frac{\mu(E,Z)}{\rho} = a_c f_c(E) + a_p f_p(E) \tag{2.8}$$

where the coefficients $a_c$and a_p are constants each related to the atomic number, Z of the object, and correspond to the Compton effect and the photoelectric effect respectively. On the other hand, assume for a material ξ that its mass attenuation is a linear combination of two other materials α and β. Then we have

$$\tau_\xi(E) = a_\alpha \tau_\alpha(E) + b_\beta \tau_\beta(E) \tag{2.9}$$

That is

$$\frac{\mu_\xi(E)}{\rho_\xi} = a_\alpha \frac{\mu_\alpha(E)}{\rho_\alpha} + b_\beta \frac{\mu_\beta(E)}{\rho_\beta} \tag{2.10}$$

Then, we have

$$\mu_\xi(E) = a_\alpha \frac{\rho_\xi}{\rho_\alpha}\, \mu_\alpha(E) + b_\beta \frac{\rho_\xi}{\rho_\beta} \mu_\beta(E) \tag{2.11}$$

On the other hand, if an object is composed of several materials, its logarithmic attenuation is the effects of the summation of all thicknesses of the materials being penetrated by X-ray. Equation (2.6) is then

$$m(E,Z) = \sum_{material\ i} \mu_i(E,Z)\ t_i. \tag{2.12}$$

For material ξ composed of two materials α and β with thicknesses t_α and t_β respectively, the logarithmic attenuation of material ξ is

$$m_\xi(E) = m_\alpha(E) + m_\beta(E) = \mu_\alpha(E)t_\alpha + \mu_\beta(E)t_\beta. \tag{2.13}$$

Assuming we have the X-ray image at two different energy levels, one is high energy denoted *HE* and other one is at lower energy denoted *LE*. Substituting into Equation (2.12), we have

$$m_\xi(HE) = \mu_\alpha(HE)t_\alpha + \mu_\beta(HE)t_\beta$$

$$m_\xi(LE) = \mu_\alpha(LE)t_\alpha + \mu_\beta(LE)t_\beta\ , \tag{2.13}$$

We can express Equation (2.13) in matrix form

$$\begin{bmatrix} m_\xi(HE) \\ m_\xi(LE) \end{bmatrix} = \begin{bmatrix} \mu_\alpha(HE) & \mu_\beta(HE) \\ \mu_\alpha(LE) & \mu_\beta(LE) \end{bmatrix} \begin{bmatrix} t_\alpha \\ t_\beta \end{bmatrix}\ . \tag{2.14}$$

The thickness of the two materials can be easily obtained as long as the determinant $\Delta = \mu_\alpha(HE)\mu_\beta(LE) - \mu_\alpha(LE)\mu_\beta(HE)$ is not zero.

For simplicity, we rewrite Equation (2.14)

$$\begin{bmatrix} m_H \\ m_L \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = A \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}\ , \tag{2.15}$$

where m_H and $m_L$are the attenuation of a material obtained from high and low energy X-rays respectively. The thickness of the two objects can be obtained by

$$\begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = A^{-1} \begin{bmatrix} m_H \\ m_L \end{bmatrix}. \tag{2.16}$$

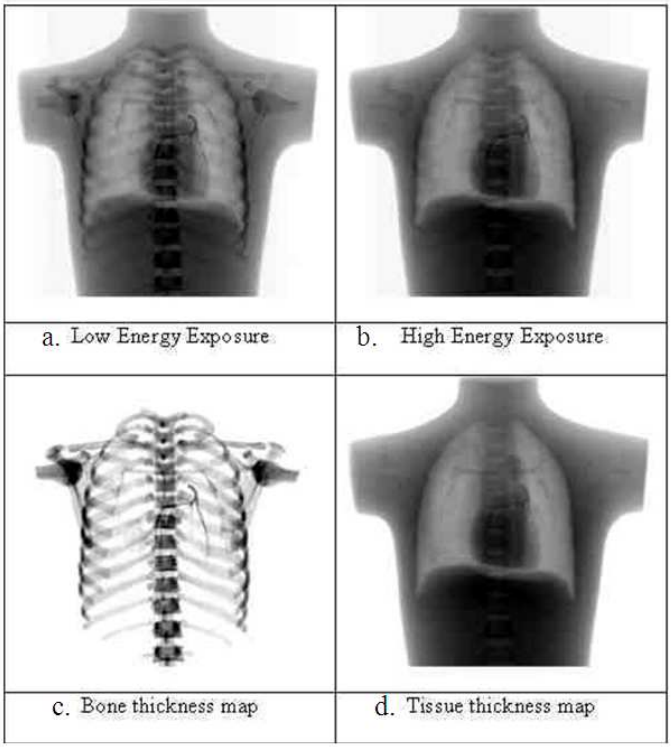


Figure 2.8 (a) a low energy X-ray chest image (b) the high energy X-ray chest image (b) bone thickness map calculated by Eq.(2.16).(d) tissue thickness map calculated by Eq.(2.16). (Image obtained form http://saturn.radsci.uci.edu)

Let human bone be the material α and tissue be the material β in Equation (2.14), the thickness of bones and tissues can be obtained from Equation (2.15) and Equation (2.16). The mass attenuations of the two materials can be obtained through samples obtained from cadavers. Figure 2.8 shows the two energy images and the thickness map of bones and tissues of a chest X-ray image from a medical DE system.

Besides separating known materials from X-ray images as shown in Figure 2.8, dual energy imaging can also be used to determine the composition of unknown materials. Assume an object has layers of two known materials: material 1 and material 2, with thickness t_1 and t_2 respectively.

If the object is solely composed of material 1, that is t_2 is zero, then, from Equation (2.16),

$$m_H = \frac{a}{c} m_L, \tag{2.17}$$

Now, recall from Equation (2.6) that attenuation is a function of energy, atomic number and thickness

$$m(E, Z) = \mu(E, Z)t. \tag{2.6}$$

Substitute the two energies into Equation (2.6) for a material with atomic number Z with a given thickness,t, gives

$$\frac{m(HE, Z)}{m(LE, Z)} = \frac{\mu(HE, Z)t}{\mu(LE, Z)t} = \frac{\mu(HE, Z)}{\mu(LE, Z)}. \tag{2.18}$$

If the material in Equation (2.18) is material 1 with $Z = Z_1$, we have

$$\frac{m(HE, Z_1)}{m(LE, Z_1)} = \frac{\mu(HE, Z)}{\mu(LE, Z)} = \frac{a}{c} = r_1. \tag{2.19}$$

Equation (2.18) and (2.19) can be interpreted that for a given type of material, when probed with two levels of X-ray energies, the ratio of the two measured attenuations is a constant regardless of the thickness. However, the statement is true only when the sources are monochromatic. That constant gives information on the atomic number. The reason for this is that the attenuation coefficient is highly Z-dependent in a certain energy and Z range; when energy is about 500KeV, the photoelectric effect dominates. Attenuation coefficients are approximately a function of Z^3.

Similarly, in Equation (2.16), if the material is solely composed of material 2 alone, we have

$$m_H = \frac{b}{d} m_L \tag{2.20}$$

and

$$\frac{m(HE, Z_2)}{m(LE, Z_2)} = \frac{b}{d} = r_2. \tag{2.21}$$

The constant $\frac{b}{d}$ gives information about the atomic number of material 2. Using the two material as base materials, we know that materials with atomic number between Z_1andZ_2will have a ratio between r_1 and r_1. Now, consider the ratio of t_1 to t_2,

$$r_z = \frac{t_1}{t_2} = \frac{m_H b - m_L d}{m_L c - m_H a}. \tag{2.22}$$

The ratio $r_z$can be seen as a relative Z number or effective Z number based on material 1 and material 2.The thicknesses t_1 and t_2 obtained here from solving Equation (2.15) are effective thicknesses since they are based on the assumption that the measured attenuation is caused by purely material 1 or purely material 2. Equation (2.22) can be stated as following:

When a tested object has a measured effective thickness t_1 of material 1 and a measured effective thickness t_2 of material 2, the ratio of the two effective thicknesses can be used as an indicator of its effective atomic number. The ratio value indicates how close the tested material is to base material 1 and base material 2. The higher the value, the greater portion of the tested object is composed from material similar to base material 1.

In practice, the two base materials are plastic and aluminum respectively. The four constants ***a, b, c***, and ***d*** can be obtained through calibration and experiments. This method is used to discriminate materials into three groups: organic, inorganic, and mixed. Airport security systems use such methods to detect plastic explosives in baggage.

Recently, there is an ongoing interest in applying dual energy imaging to cargo container inspection for customs to detect very high atomic number materials (high Z materials). As discussed in Section 2.2, when X-ray images are formed by a single X-ray energy source, the image is formed by displaying the relative attenuation, or conversely, transmission, of all objects being scanned. The thicker, denser objects will have relatively high X-ray attenuation (low transmission) compared to the thinner or

low density objects. Therefore, the best that customs inspectors can do from the traditional single source X-ray image is to identify suspect objects from their shapes. This created an opportunity for people to sneak in contraband materials by shaping them into non-suspicious looking objects. In that case, a box of uranium can avoid detection by masquerading as a box of china. Also, a single X-ray energy image cannot really provide the information of the thickness of an object. When an area appears to be dark, we cannot really tell whether it is caused by a very thick material or a thin good absorber. Dual energy imaging, by providing the atomic information of the scanned object, is a solution to that problem.

To discriminate high Z materials such as uranium in a cargo container, the X-ray energy has to be much higher than what is used for airport security systems. One reason is that high Z materials are harder to penetrate. It is also because the dual energy imaging mechanism operates differently. When an object is scanned by photons with energy up to several mega electron-volts (~MeV), there are two dominant reactions: Compton scattering, and pair production. With photons having higher energy (~MeV), pair production can occur: higher energy photons are able to bombard the nucleus and cause an electron and a positron to be formed. Pair production will be more salient in materials with greater atomic number. This is because greater atomic number materials have greater packed nuclei that increase the chance of collision. Also, the higher the photon energies, the greater chance pair production will occur.

The most commonly used dual energy pair is a 10-4 MeV pair or a 9-6MeV pair. The higher energy source is able to trigger more pair productions than its low energy counterpart. When scanning a high Z material with high energy- and low-energy X-rays respectively, it is anticipated that fewer photons will be detected from the higher energy detector than from a lower energy detector because pair production causes photons to lose energy on their path to the detectors, and causes greater attenuation as a result. Therefore, object will appear to be thicker in a high energy scan than in its low energy counterpart. Taking the ratio of the attenuations from high

energy and low energy images, a value greater than unity is expected in the regions where more pair production takes place in the high energy scan. Those regions are the high atomic number material regions. Theoretically, the higher the atomic number is, the greater the ratio will be.

For low atomic number materials, the less dense nuclei give less chance of collisions between atoms and photons; the pair production is not obvious. In this case, the lower energy photons will be more easily absorbed than high energy photons. Therefore, low Z material will appear to be thicker on low energy images. As a result, the ratio of attenuation for high- and low-energy scans is less than unity.

Table 1 lists the attenuation coefficients of several materials scanned by two different energy X-rays. A 9-MeV X-ray beam is sufficient to trigger enough pair production in contrast to a 6-Mev X-ray. The third row of Table 1 shows the ratio between the two high and low energy attenuation coefficients. As we can see that the ratio increases with the Z value. This is because low Z value materials do not have a very salient pair production effect. When the nucleus is dense enough, for example, Z=82, the pair production starts to be prominent. Fewer photons will pass through the material as a result; and the attenuation coefficient becomes greater than that of lower energy X-rays.

Table 2.1 Attenuation coefficients and ratios of several materials (Data obtained from NIST website)

	H_2O (Z=10)	Steel (Z=26)	Cu (Z=29)	W (Z=74)	Pb (Z=82)	U (Z=92)
Density(g/cm^3)	1	7.86	8.94	19.3	11.35	18.68
μ_{9MeV}/ρ	2.31E-02	2.97E-02	3.08E-02	4.61E-02	4.82E-02	4.583E-02
μ_{6MeV}/ρ	2.77E-02	3.05E-02	3.11E-02	4.21E-02	4.39E-02	5.195E-02
$r = \frac{\mu_{9MeV}}{\mu_{6MeV}}$	0.8339	0.9738	0.9904	1.0950	1.0979	1.1335

2.3.2 Dual Energy Imaging for Material Discrimination in Cargo Screening

So far, several methods have been proposed to use a dual energy imaging system for material discrimination. In the pioneering work by Ogorodnikov and Petrunin (Ogorodnikov et al., 2002), the main idea is to use the ratio of the attenuation at two energy levels HE and LE for determination of the atomic numbers, Z.

$$\frac{\ln Tr(HE,Z)}{\ln Tr(LE,Z)} = \frac{\mu(HE,Z)t}{\mu(HE,Z)t} = \frac{\mu(HE,Z)}{\mu(HE,Z)} = \delta(HE,LE,Z) \qquad (2.23)$$

The ratio δ is a function of the two energies and atomic number. If the system is monochromatic, for the same material scanned in the same system, the ratio is a constant regardless of the thickness of the material. For a polychromatic imaging system, the ratio δ will vary with material thickness. The reason is that the spectrum of photons will be affected by the depth of the beam path. Thicker materials will filter out more lower-energy photons in the spectrum. This effect is called hardening. When the spectrum is hardened, the attenuation coefficient becomes larger. Therefore, for a practical DE imaging system, the ratio is a function of measured attenuation, which is a function of object thickness.

A similar idea was proposed by Zhang et al (Zhang et al., 2005). They proposed an H-L (high and low energy transparencies) curve method for material recognition in dual energy X-ray inspection systems. The two transparencies are defined

$$H = e^{-\mu_H t} \qquad (2.24a)$$

$$L = e^{-\mu_L t}, \qquad (2.24b)$$

given

$$\frac{\ln H}{\ln L} = \frac{\mu_H}{\mu_L}, \qquad (2.25)$$

therefore,

$$H = L^{\frac{\mu_H}{\mu_L}}. \tag{2.26}$$

Different materials will form different curves on the *H-L* plane in the form of $y = x^R$. The key factor that determines the shape of the *H-L* curves is the power term, $\frac{\mu_H}{\mu_L}$, which is the ratio of two attenuation coefficients. One implication of this method is that the ratio for a specified material is a constant. In reality, this assumption is applicable for certain materials within a certain range of thicknesses. Our experimental results show that the ratio of steel remains a constant until it becomes thicker than 11 inches.

The measured attenuation can be expressed in term of effective thickness or equivalent thickness, which is the thickness of the selected calibration material that causes that amount of attenuation. In this way, user can have a better understanding of how dense or thick the material is when compared with a known reference.

Let us assume that copper is used as the calibration material for a system. After scanning this calibration material with attenuation coefficient, μ_{Cu} and thickness, t_{Cu}, according to Equation (2.1), we have

$$N_1 = N_0 e^{-\mu_{Cu} t_{Cu}} \tag{2.27}$$

where N_0 and N_1 are the number of incoming photons and remaining photons, respectively. Assuming that the remaining number of photon after probing an unknown material is N_2, then, the attenuation of that material can be obtained by

$$N_2 = N_0 e^{-\mu_2 t_2} \tag{2.28}$$

Although we do not know what μ_2 and t_2 are, we can express them in terms of μ_{Cu} and t_{Cu}. If $N_1 = N_2$, it means that

$$\frac{N_1}{N_2} = 1 = \frac{N_0 e^{-\mu_{Cu} t_{Cu}}}{N_0 e^{-\mu_2 t_2}} = e^{-(\mu_{Cu} t_{Cu} - \mu_2 t_2)} \tag{2.29}$$

And therefore,

$$\ln\frac{N_1}{N_2} = \ln 1 = 0 = \ln e^{-(\mu_{Cu}t_{Cu}-\mu_2 t_2)} = -(\mu_{Cu}t_{Cu} - \mu_2 t_2) \tag{2.30}$$

and

$$\mu_{Cu}t_{Cu} = \mu_2 t_2. \tag{2.31}$$

Equations (2.27) to (2.31) tell us that the thickness and attenuation coefficient of an unknown material is causing the same attenuation of photons as copper with attenuation coefficient μ_{Cu} and thickness t_{Cu}. In other words, we can say that the object with thickness t_2 has an equivalent thickness t_{Cu}, in terms of copper, the calibration material.

Therefore, the equivalent thickness of an object with attenuation coefficient μ and thickness *t* can be computed using the equation below:

$$t_{eq} = \frac{\ln\frac{N_0}{N}}{\mu_{Cal}} = \frac{\mu t}{\mu_{Cal}} \quad , \tag{2.32}$$

where t_{eq} is the equivalent thickness in terms of the chosen calibration material and μ_{Cal} is the attenuation coefficient of the calibration material.

Now, consider $\frac{\mu_H}{\mu_L}$, the ratio of attenuation coefficients from the two energy scans. From Equation (2.31) and Equation (2.32) we know the equivalent thickness of high energy scan can be expressed as

$${t_H}^{eq} = \frac{t_H \mu_H}{\mu_H^{cal}}\,, \tag{2.33}$$

where $\mu_H^{cal}$and t_H^{eq} are the attenuation coefficient and equivalent thickness of the calibration material respectively, and t_H is the thickness of the material. Similarly, for the low energy scan, we have

$${t_L}^{eq} = \frac{t_L \mu_L}{\mu_L^{cal}}\,, \tag{2.34}$$

Therefore, the ratio of two attenuation coefficients becomes

$$\frac{\mu_H}{\mu_L} = \frac{\mu_H^{cal}\, {t_H}^{eq} \;/\; t_H}{\mu_L^{cal}\, {t_L}^{eq} \;/\; t_L} = \frac{\mu_H^{cal} \;/\; t_H}{\mu_L^{cal} \;/\; t_L}\frac{{t_H}^{eq}}{{t_L}^{eq}} = \frac{{t_H}^{eq}}{{t_L}^{eq}}. \tag{2.35}$$

The equivalent thickness is the result of two scans of the same object. The thickness $t_H = t_L$. Therefore, the constant $c = \frac{\mu_H^{cal}}{\mu_L^{cal}}$.

What we can see from Equation (2.34) is that the ratio of two equivalent thicknesses gives the same information as the ratio of two attenuation coefficients. Therefore, the ratio of equivalent thicknesses serves the same function as the ratio of attenuation coefficients for the purpose of material discrimination. In this thesis, we will use the ratio of equivalent thicknesses as our ratio for Z determination.

Theoretically, the ratio serves as an indicator for discriminating between materials. The effective atomic number can be identified by taking the ratio of two images obtained from the DE system. By properly selecting the base materials to create a lookup table, the image can be classified into different groups of materials. However, in practice, it is somewhat more complicated. First of all, the pixel noise caused by statistical noise, scattering or artifacts will cause errors in the measurement of equivalent thickness and thus the ratio. The measurement of ratio is very sensitive to any perturbation. Secondly, due to machine error, the measurements are not consistent every time. These factors provide the possibility of misclassification.

2.3.3. Statistical Errors

Let the ratio of measured equivalent thickness from high and low energy scans ${t_H}^{eq}$ and ${t_L}^{eq}$ be

$$R = \frac{{t_H}^{eq}}{{t_L}^{eq}}. \tag{2.36}$$

Since Poisson noise is the dominant sources of error for thin objects in X-ray imaging system, the measured attenuations and consequently, the ratio, will be affected by it. If the signal to noise ratio is low, contrast sensitivity is also low, and the ratio will be an

unreliable indication of atomic number. The model of noise can be obtained from experiments. By scanning calibration material of the same thickness many times, a distribution of measured thickness is obtained. The standard deviation of the measured thickness is then defined as the error. When penetration is good, the major noise contributed to the system will be from the Poisson behavior of X-ray photons. We also know that the Poisson distribution is close to the normal distribution if the number of events, in this case, the number of photons, is large. It is therefore reasonable to use Gaussian distribution to model the noise.

Recall that the probability distribution function (PDF) of the Poisson distribution is such that for a temporal or spatial interval, if the expected number of occurrences in this interval is λ, then the probability that there are exactly k occurrences (k being a non-negative integer, k = 0, 1, 2, ...) is equal to

$$f(k,\lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \tag{2.37}$$

For sufficiently large values of λ, (say λ>1000), the normal distribution, with mean λ and variance λ, is an excellent approximation to the Poisson distribution. That is

$$f_{Poisson}(x;\lambda) \approx f_{normal(x;\mu=\lambda,\sigma^2=\lambda)} \tag{2.38}$$

If λ is greater than about 10, then the normal distribution is a good approximation if an appropriate continuity correction is performed. Equation (2.38) can be interpreted in this way: if the random variable x is the number of photons, then, we know that if N photons are detected, the standard deviation is about$\sqrt{N}$. The error is then approximately $\frac{1}{\sqrt{N}} * 100\%$. The greater N is, the smaller the error will be. Greater N is achieved by greater penetration, meaning the object is thin. Also, since the equivalent thickness is proportional to $\ln N$, it means when the object being measured has a thickness T, the measured thickness will have an error in the range of $\frac{1}{2}T$ That means the thicker the object is, the greater the error will be. When X-ray penetration is poor, fewer photons will be detected and the SNR will be low. As a result, the calculated equivalent thickness will be less reliable. Therefore, it is hard to obtain the atomic

information using the ratio in Equation (2.36) when objects are insufficiently penetrated.

The probability density function of equivalent thickness, which is a function of photon count, can be seen as a Gaussian distribution with mean ${t_H}^{eq}$ and standard deviation σ_H (${t_L}^{eq}$ and σ_L for low energy thickness). The standard deviations of high and low energy thickness ($\sigma_H, \sigma_{\mathrm{L}}$) are derived from the calibration data, taking into account the number of pulses used to form each pixel. It is reasonable to assume that the distribution function for the ratio is also a Gaussian. In practice, the standard deviations of the equivalent thicknesses, σ_H and σ_L, is obtained by repeatedly measuring the same object to obtain the statistics.

From the standard deviation of the high and low energy thickness, we then can calculate the standard deviation of the ratio:

$$\sigma_{\mathrm{R}} = \mathrm{R}\sqrt{\left(\frac{\sigma_H}{{t_H}^{eq}}\right)^2 + \left(\frac{\sigma_{\mathrm{L}}}{{t_L}^{eq}}\right)^2}\,. \tag{2.39}$$

What Equation (2.39) means is that when a ratio is obtained from calculating Equation (2.36), because of the errors in measured thicknesses, the ratio will also have an error of about σ_{R}. The value in Equation (2.39) is the minimum ratio error in the system, since it only considers the Poisson noise and disregards other sources of noise. When an object is thick, the Poisson noise is no longer dominant. However, the electronic noise and pixel artifacts will cause a smaller signal to noise ratio than what Poisson noise can produce.

For high Z material discrimination, to screen out the material above certain atomic number from other materials, one can choose a “minimum high Z ratio” obtained from experimental data such that the ratio is a threshold value one can expect to detect a minimum size (usually 100cc) of high Z object of interest. The system’s maximum penetration determines the value of the minimum high Z ratio. The lowest

minimum ratio can be the ratio of the effective Z of steel as thick as the maximum penetration plus a 5cm thick lead plate.

Taking statistics errors into consideration, the probability of an object with a calculated ratio R being a high Z material is

$$P_{highZ} = 1 - 0.5\left[1 + \mathrm{erf}\left(\frac{R_{min} - R}{\sqrt{2}\,\sigma_R}\right)\right] . \qquad (2.40)$$

The probability is obtained by integrating the normal distribution from -∞ to $R_{min} - R$. As shown in Figure 2.9, the shaded area represents the probability an object with a measured ratio R=1.09 being greater than the minimum high Z ratio(which is set to be 1.02).

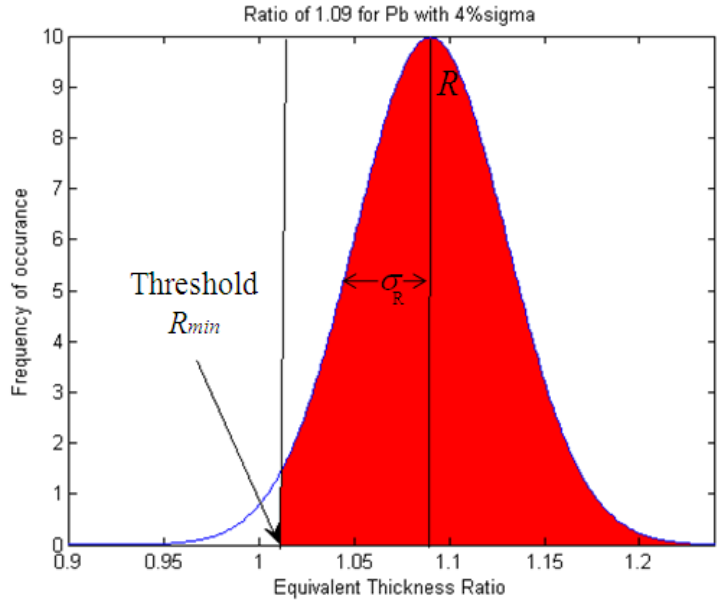


Figure 2.9 Probability of being high Z is the integral above the test ratio (the shaded area).

The meaning of Equation (2.40) is, if a ratio R is obtained from the measurement, that value of R is then seen as the expected value of a normal distribution, with standard deviation that can be obtained from Equation (2.39). Therefore, the actual value has the highest probability to be R and has 96% chance being within the range $R \pm \sigma_R$. If the minimum ratio is less than $R - \sigma_R$, i.e. the

measured ratio is much higher than the threshold, the probability obtained from Equation (2.40) will be high. Conversely, if the minimum ratio is between $R - \sigma_R$ and R, the probability will be smaller. When R is equal to the threshold, the probability of having high Z is fifty percent.

As discussed before, the standard deviation of thickness and, consequently, R increases as the object thickness increases. As a result, for a chosen minimum ratio, if R is obtained from scanning on a thick object, the bell-shaped distribution function in Figure 2.6 will expand in the horizontal directions. That will make the total area above the threshold ratio line decrease. The uncertainty of the tested object being high Z will then increase. This again explains why it is hard to detect a heavily shielded high Z object. In reality, high Z contraband materials are radioactive and are heavily shielded to prevent the radiation from leaking.

2.4 Challenges of a Dual Energy Cargo Imaging System

The idea of using dual energy imaging for material discrimination in cargo security is a new technology and contains a lot of challenges. An ideal system will carry out the task of threat material detection as automatic as possible. Identifying the effective atomic number of all objects inside the container requires a very accurate measurement of attenuations and ratios. From our previous discussions, we already know that it is not possible due to various sources of noise.

At the same time, once a threat target is identified, the operator should be able to locate the location of that threat target by inspecting the X-ray image. In that case, good image quality is required. What is needed is a sophisticated image processing algorithm for visualization involving contrast enhancement and noise reduction, while at the same time, the computation for that image processing algorithm has to be done in real time to accommodate the commerce traffic flow.

Among the factors that cause low sensitivity of the system is the influence of scattered radiation, which is usually controlled by careful collimation. Due to the geometry of the objects and the complexity of the object arrangements, photons will be randomly scattering all over the place. Using a beam controller does not guarantee the elimination of such a problem. The accuracy of the ratio calculation is impeded by unpredictable scattering. Detecting small objects requires a high resolution detector. The tradeoff is that low counts for a single detector, which might not provide enough statistics.

The cluttering nature of the container makes it harder to develop a pattern recognition algorithm. Besides, the contraband materials can be formed into any shape to deceive the inspector. A pattern recognition algorithm will not be fully helpful in high Z material detection. As discussed before, high Z objects are often heavily and shrewdly shielded to avoid detection. The shielding will make the effective Z equivalent to a mid-Z material. How to distinguish the shielded high Z object from other objects having the same ratio is one of the biggest challenges.

2.5 Summary

In this chapter, we've given a comprehensive discussion of the background of X-ray imaging system and dual-energy X-ray imaging. Dual energy techniques can be very useful for industrial national security technology applications. The complexity of the problem depends on the number of constituent materials, prior knowledge about them, and the information required (thickness or other). Constraints should be considered, especially acquisition geometry and time limits. Although several challenges still remain to be overcome, this newly developed technology for cargo material discrimination has a promising future. In later chapters, we will address proposed solutions to solve the problems from several different perspectives.

3

Review of Digital Image Processing

3.1 Introduction

In this chapter, we will review several ideas of digital image processing including basic definitions, the most commonly used approaches, and some applications. From the discussions in this chapter, we will obtain an understanding of several image processing techniques and how they can be applied to an X-ray imaging system.

An image can be defined as a two dimensional function, $f(x, y)$, where x and y are spatial coordinates, and the amplitude of f at any pair of coordinates (x, y) is called the intensity at that point. When all values x and y are finite and discrete quantities, i.e. finite integers, and all values of f are quantized, we call the image a *digital image*. The images referred in this thesis are all digital images. Digital image processing refers to processing digital images by computers. The scope of our discussion in this chapter will focus on the several basic definitions and the means of image enhancement and image segmentation.

3.2 Fundamental Concepts

A digital image is a two dimensional array, each element of that array corresponds to the location in the image, and the value of that element is the intensity, or gray level, of that location in the image. Each element is the smallest unit of an

image. *Pixel*, obtained from combining “picture” and “element”, is the most used name for the smallest unit of a digital image. One useful way to define image processing is to categorize three types of computerized processes in this continuum: low-, mid-, and high-level processes.

3.2.1 Hierarchy of Image Processing

A low-level process has both the input and output of the processing system as images. Contrast enhancement and noise reduction are examples of low-level image processing. A mid-level process involves tasks such as segmentation (partitioning images into regions or objects), description of those objects to reduce them to a form suitable for computer processing, and classification or recognition of individual objects. The output of mid-level processing are attributes extracted from the input images (e.g. contours, edges, and identity of individual objects). High-level processing involves machine intelligence. Its tasks are to make sense of the recognized objects, ranging from image analysis to cognitive functions associated with vision. Currently, some auto teller machines (ATM) are able to read handwritten numbers on checks, converting the image to text information, which is a good example of high-level image processing. A cargo imaging system, which will be discussed throughout this thesis involves all three levels of digital image processing. The digital X-ray images will be filtered to eliminate noise and processed to enhance the contrast during the first stage of processing. Image analysis in the second stage of processing includes object edge and boundary detection, and segmentation of objects. The last stage will be threat identification, reading the image to extract information to determine whether there is a threat that should trigger an alarm.

3.2.2 Digital Image Concepts

Image Formation

An image can be seen as a two dimensional function $f(x, y)$. The value or amplitude of f at spatial coordinates (x, y) is a positive scalar quantity whose physical meaning is determined by the source of the image. The source of an image is a

physical object that radiates light (electromagnetic wave). The values of the pixels in an image are proportional to the intensities of the lights radiated from the source.

The function *f* (*x*, *y)* can be characterized by two components: (1) the amount of source illumination incident on the scene, and (2) the amount reflected (for X-ray images, transmitted) by the objects in the scene. The first is called the illumination component, and the second is called the reflectance component, and they are denoted by *i*(*x*, *y*) and *r*(*x*, *y*) respectively. The two functions combine as a product

$$f\ (x,y) = i(x,y)r(x,y) \tag{3.1}$$

where

$$0 \leq i(x,y) < \infty \tag{3.2}$$

and

$$0 \leq r(x,y) \leq 1 \tag{3.3}$$

Equation (3.3) sets the boundaries of reflectance. When $r(x,y) = 0,$ it means a total absorption of light. That kind of material is called a blackbody. In that case, none of the signal can be received by the image receptor. When $r(x,y) = 1$, it is a total reflectance. This means the object reflects all the EM waves projected on the object. For X-ray images, the reflectivity represented by the reflectance function, $r(x,y)$, should be replaced by the *transmissivity*, which is a function of thickness and attenuation coefficient as shown in Equation (2.4) and Equation (2.5) in Chapter 2. However, the model of Equation (3.1) is still applicable.

3.2.3 Image sampling and Quantization

We already established that an image is a two dimensional function as shown in Equation (3.1), where the variables *x* and *y,* and the values of the function *f*(*x*, *y*) can be either continuous or discrete. Digitization converts continuous values into discrete quantities. A digital image is a discrete two-dimensional function (i.e. both variables *x*

and *y* are digitized to integers) and the amplitudes of the function are represented by finite levels of quantities.

To create digital images, we need to convert the image function from continuous into digital form. That involves two processes, sampling and quantization. Sampling is to digitize spatial coordinates into discrete values; and quantization is to digitize amplitude, *f*(*x*, *y*), into different levels.

Sampling refers to taking samples in spatial domain. Assuming the sampling step size to be Δx, then, a sampled one-dimensional function will be

$$f(n) = f(n\Delta x) \qquad \text{where } n = \text{positive integers} \qquad (3.4)$$

The values of Equation (3.4) corresponding to evenly spaced locations of the continuous (analog) image. The output of Equation (3.4) spans from a continuous space which is composed of an infinite number of quantities. The quantization of amplitude is to round each $f(n)$ to the closest values at discrete levels. In practice, the number of quantization levels is a power of 2.

The size of sampling, Δx, determines the *resolution* of an image. Resolution is the measurement of the fineness of an image. The quality of an image is affected by its resolution. A high definition image is a high resolution digital image, meaning each portion of a scene is represented by more pixels. The resolution of a digital image is somehow determined by the nature of the image acquisition sensors. For digital *X-ray* images, the resolution is determine by spacing of the sensor array.

Digital Image Representation

An *M* by *N* digital image is a two dimensional discrete function with digitized amplitudes. In this thesis, we use the convention *x* and *y* to represent the discrete coordinates. Some authors prefer use *m* and *n* instead of *x* and *y* to represent integers. In a digital image, the vertical direction is *x* and the horizontal direction is *y* axis. The values of the origin are $(x, y) = (0, 0)$. The last pixel is located at (*M*-1, *N*-1).

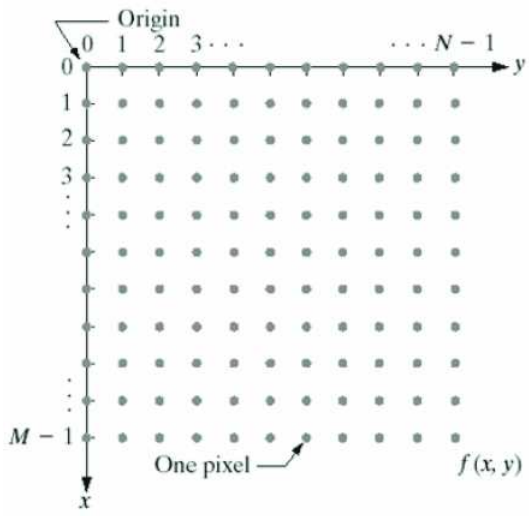


Figure 3.1 Coordinate convention used to represent digital images. (Images from Gonzalez, Digital Image Processing, 2nd edition)

In Matlab Image Processing Tool Box, the coordinates are all positive integers. The values of origin are $(x, y) = (1, 1)$. And the last pixel is located at (*M*, *N*)

3.2.4 Basic Relationships between Pixels

In this subsection, some basic terminologies of digital images are defined. The basic concepts of pixel operations come from set theory. Most of the definitions provided here are defined by the concepts of set theory.

Neighbors of a Pixel

As defined, a digital image is a 2-D array of pixels. The pixels surrounding a specific pixel (*x, y*) are called the neighbors. The nearest four pixels are the ones directly adjacent to the center pixel (*x, y*). They are pixels (*x*-1, *y*), (*x*+1, *y*), (*x*, *y*-1), and (*x*, *y*+1). The nearest 8 neighbors are the nearest 4 plus the other four pixels in two diagonal directions including (*x*-1, *y*-1), (*x*+1, *y*+1), (*x*-1, *y*-1), and (*x*-1, *y*+1). As expected, the nearest 24 neighbors are the pixels of a 5 by 5 sub-image except the center pixel. Figure 3.2 illustrates the location of nearest 4 and nearest 8 to a center pixel.

The nearest four neighbor pixels of a point p=(*x, y*) are denoted N_4 (*p*). The nearest 8 neighbors are denoted N_8 (*p*). And similarly, the nearest *D* pixels are denoted

N_D (p). If a pixel falls on the border of an image, several of its neighbors will fall outside the image.

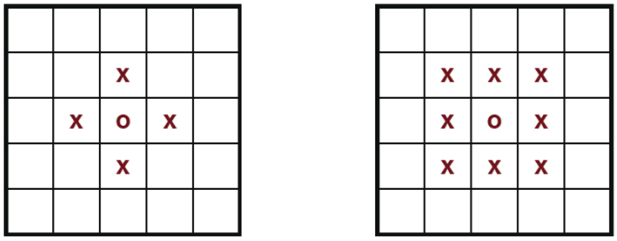


Figure 3.2 Illustrations of (a) nearest 4 neighbors and (b) nearest 8 neighbors

Adjacency, Connectivity, Regions, and Boundaries

To say two pixels p and q are connected, two conditions have to be satisfied: the two pixels have to be adjacent, and the gray level values of the two pixels satisfy the criterion of similarity. The similarity between two pixels is determined by a predefined set, V. All elements in set V belong to a group of similarity. If the gray values of two pixels both fall inside the set V, the two pixels are said to be similar in gray values. The simplest case is binary images which only have 0 and 1 as pixel values. In the binary case, $V = \{1\}$. For gray level images, the set can contain as many elements as necessary.

The adjacency of two pixels means the gray values of the two pixels belong to the same set of similarity and each pixel is a neighbor pixel to the other. Three types of adjacency are commonly used:

(1) *4-adjacency*: two pixels p and q with values from V, if q belongs to the set $N_4(p)$

(2) *8-adjacency*: two pixels p and q with values from V, if q belongs to the set $N_8(p)$

(3) *m-adjacency*(mixed adjacency): two pixels p and q with values from V, if

(i) q belongs to the set $N_4(p)$

and

(ii) q is in $N_D(p)$ and the set $N4(p) \cap N4(q)$ has no pixels whose values are from V

Figure 3.3 gives examples of the cases of different adjacencies. Figure 3.3(a) is a 3x3 binary image comprised of only 1's and 0's. In Figure 3.3 (b) the dashed lines connect the adjacent pixels together. All those adjacent pixels belong to the center pixels's 8-nearest neighbors, and all of them have the same pixel value i.e. they all are in set V = {1}. Figure 3.3 (c) is an example of m-adjacency. The dashed lines connect the m-adjacent pixels together. Note that the pixel on the top right is not an *m*-adjacent pixel to the center pixel because they belong to the 4-nearest neighbor to each other while their pixel values are both in set V={1}.

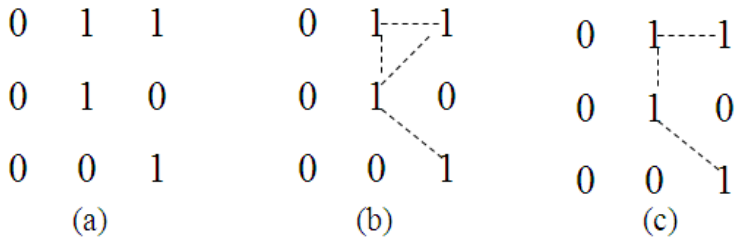


Figure 3.3 (a) a 3x3 binary image (b) pixels that are 8-adjacency to the center pixel (connected by the dashed line) (c) m-adjacency

A path between two pixels *p* and *q* is a sequence of pixels such that each pixel is adjacent to its previous and next pixel. For example, the dashed lines in Figure 3.3 (c) form a path between the top right pixel and the pixel on the bottom right.

Assuming ***S*** is a subset of pixels in an image, two pixels *p* and *q* are said to be connected in ***S*** if there is a path between them and all the pixels on that path consist entirely of the pixels in ***S***. The set of connected pixels in ***S*** are called the connected component of ***S***. If there is only one connected component, meaning the pixels in the whole set S are connected, S is then called a connected set.

As discussed, the connectivity between two pixels requires the two pixels meet the criterion of similarity and adjacency. A region is defined by the connectivity of a

set of pixels. In a digital image, a subset of pixels *R* is called a region if *R* is a connected set. For example, the 1's in Figure 3.3 form a region. The boundary of a region *R* is defined as a set of pixels in R that have at least one of its neighbors not belonging to *R*.

The edges in an image are the regions of pixels that have a discontinuity of gray level values. The boundary of a region is then certainly an edge. However, within a region, there could be other edges as well since it is possible that the values of pixels in a region are not uniformly or smoothly distributed.

3.3 Image Enhancement in the Spatial Domain

In the previous section, we have defined what a digital image is and several conventions used to describe properties of a digital image. The term spatial domain refers to the two dimensional image plane itself. Some people also refer it as the pixel domain. There are techniques that process image in a transformed domain such as the frequency domain or wavelet domain. The term image enhancement refers to processing an image to meet the specific objective of a problem. Therefore, different methods are applied to meet different needs. For example, the approaches that work best for enhancing the contrast of an X-ray image might not work best for a human photo. In this section, we will discuss some of the most used image enhancement techniques that can be applicable to most kinds of images, including X-ray images.

The most common purposes of image enhancement are contrast stretching, image smoothing, and image sharpening. In this section, we will briefly discuss some general image enhancement approaches in the pixel domain.

3.3.1 Contrast Enhancement

Contrast is the degree of difference in pixel values between an object and its background. When the difference is large, the object stands out more from the background. Contrast enhancement or contrast stretching is a very useful technique for low contrast images such as X-ray images. Several techniques are popular for contrast enhancement such as histogram transformation and gray level transformation. Histogram transformation adjusts the histogram of pixel values and maps the new histogram back to the corresponding pixels. Gray level transformation refers to manipulation of grayscales of pixels through a certain equation. For X-ray image a commonly used contrast enhancement approach is the negative transform. Negative transform can be expressed as

$$s = T(r) = L - 1 - r\,, \quad 0 \le r < L \tag{3.6}$$

where L is the upper bound of input gray level. Negative transform is a commonly used method for contrast enhancement of X-ray images. Figure 3.4 is an example of negative transformation applied to a mammogram.

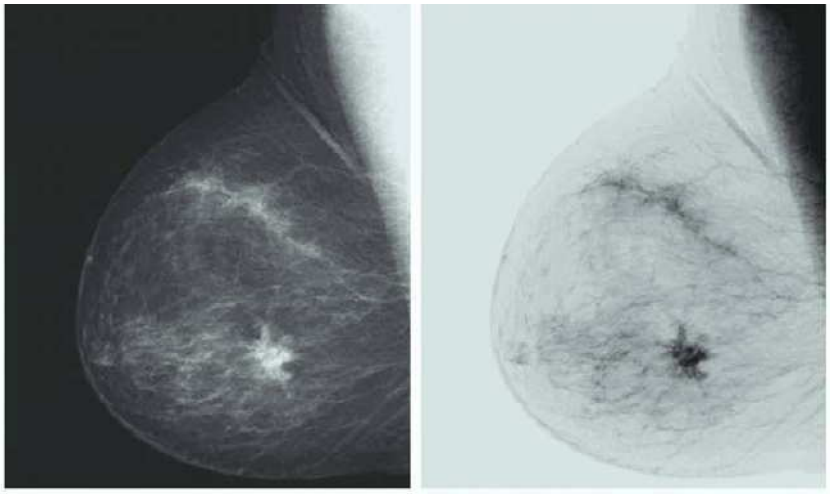


Figure 3.4 (a) The original mammogram image (b) the negative transformed image (Images from Gonzalez, Digital Image Processing, 2nd edition)

3.3.2 Spatial Domain Filtering

Image Noise

Digital images are prone to a variety of types of noise. Noise is the result of errors in the image acquisition process that result in pixel values that do not reflect the true intensities of the real scene. There are several ways that noise can be introduced into an image, depending on how the image is created. It can occur during image capture, transmission or processing, and may be dependent on or independent of image content.

Noise is usually described by its probabilistic characteristics. For example, white noise, which has a constant power spectrum (its intensity does not decrease with increasing frequency), can be modeled by Gaussian noise which is a very good approximation of noise that occurs in many practical cases. Noise can be additive, which has a general form

$$g(x,y) = f(x,y) + \eta(x,y) \tag{3.9}$$

where $g(x,y)$ is the image $f(x,y)$ degraded by noise $\eta(x,y)$. In Equation (3.9), noise $\eta(x,y)$ and image signal $f(x,y)$ are independent.

Multiplicative noise is the case where noise is a function of signal magnitude, which can be expressed as

$$g(x,y) = f(x,y) + f(x,y)\eta(x,y) = f(x,y)(1 + \eta(x,y)) \tag{3.10}$$

This kind of noise is rarely seen.

Impulse noise, or called salt and pepper noise, is a random impulse that appears to be white and black dots superimposed on an image.

Linear Filtering

Filtering refers to operations that eliminate the power of the signal within a certain frequency bandwidth. If the signal is a two dimensional image, the frequency here refers to the number of changes of gray-level values within a spatial frame. The

higher the number, the greater the frequency will be. Therefore, in a smooth region of an image, that region contains low frequency. The regions that have high contrast of gray-level values are seen as high frequency regions. For example, in an image of a zebra on grassland, the regions of the zebra will be high frequency regions since there are constantly changing gray-level values, due to the stripes. On the other hand, the background regions, such as the grassland or the sky, are the low frequency regions since they are relatively smooth in gray-level values.

Noise, since it introduces a sudden change in gray-scale values, makes the noise-degraded regions high frequency regions. A filter that smoothes that region will reduce the high contrast in gray-scale values. That filter is a low pass filter since it eliminates the power in high frequency components. One result of low pass filtering is blurriness of edges. The edges of an image are the pixels that have an abrupt change in value relative to their neighbor pixels. Because the high frequency portions of the image are removed, the contrast in pixel values is also reduced in the spatial domain. As a result, the image becomes blurred. If after filtering, the contrast of the image is enhanced, for instance, the edges of the objects in the image become sharper, that filter is then a high pass filter since it removes the blurriness in an image.

In spatial domain filtering, a filter is a sub-image, which can be called a mask, a kernel, a template, or a window. The elements of the sub-image are called the coefficients. Linear filtering an image is to convolve the mask with the region on the image being masked. The mask shifts along the image in two directions and compute the value of the pixel located at the center of the filter mask.

The general form of linear filtering an M by N image f with a filter mask of size m by n is given as

$$g(x,y) = \sum_{s=-(m-1)/2}^{(m-1)/2} \sum_{t=-\frac{n-1}{2}}^{\frac{n-1}{2}} w(s,t)\, f(x+s, y+t) \qquad (3.10)$$

Figure 3.5 illustrates how the convolution in Equation (3.10) is done. The mask shifts along in x and y directions until every pixel in f is covered as the center

pixel under the mask. Each center pixel of value $f(x, y)$ is then replaced with value $g(x, y)$after the operation.

Smoothing Spatial Filters

Two examples of smoothing filter are given in Figure 3.6. The first kernel is a general form of a 3 x 3 mask, the second and third kernels are the coefficients that function as smoothing filters.

For an image *f*, output image *g'* s value at location (*x*, *y*) will be

$$g(x,y) = \frac{1}{(\sum_{i=0}^{8} w_i)}[f(x,y)w_0 + f(x-1,y-1)w_1 + f(x,y-1)w_2 + f(x+1,y-1)w_3 + f(x-1,y)w_4 + f(x+1,y)\ w_5 + f(x-1,y)+1)w_6 + f(x,y+1)w_7 + f(x+,y+1)w_8] \quad (3.11)$$

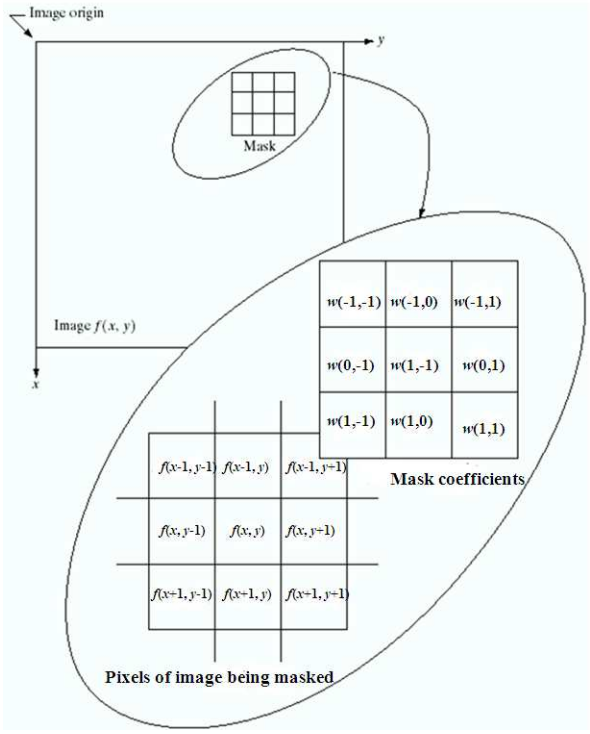


Figure 3.5 Illustration of spatial filtering. A 3 by 3 mask operates on an image *f* (*x*, *y*). The region under the mask is shown on the second image. (Images adopted from Gonzalez, Digital Image Processing, 2nd edition)

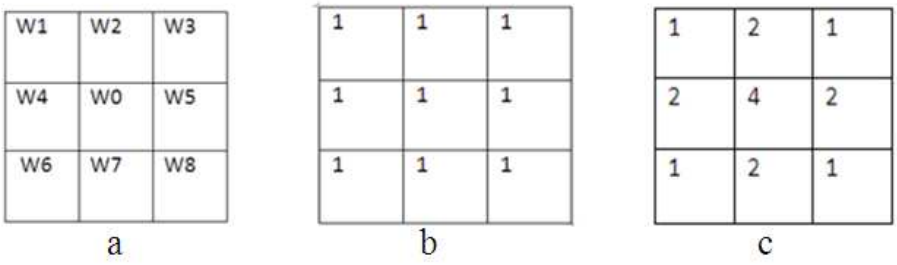


Figure 3.6 A 3 by 3 filter mask (a) and two examples of low lass filter coefficients (b & c)**.**

The second mask in Figure 3.6 is called the mean filter since it simply takes the mean of the values of the region being masked. This operation smoothes the contrast of pixel values between the center pixel and its neighboring pixels. If the center pixel has a much greater value than its 8 neighbors, it will be reduced to average of all 9 pixels. The third mask in Figure 3.6 is another averaging filter with different weights. Here the center pixel has the greatest weight in averaging, while its diagonal neighbors have the least influence. This smoothing filter therefore will not result in an image as blurry as the result of the previous one. The size of filter depends on the objective of the processing and the resolution of the image. A 3 by 3 or a 4 by 4 mask are most often used.

Sharpening Linear Filters

Sharpening filters are masks that after convolving with an image, they enhance the fine details in the image that have been blurred. Sharpening is usually done by a linear combination of an all pass filter and high pass filter. The output of an all pass filter will be the original image. As explained before, a low pass filter smoothes an image by lowering the contrast between neighboring pixels. We can expect that a high pass filtering is an operation that will enhance such contrast.

One of the most commonly seen high pass filters is the Laplacian mask. The Laplacian of a continuous function *f* is defined as

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \tag{3.12}$$

For f (x), a discrete function, the definition of the first-order derivative of a one dimensional function $f(x)$ is

$$\frac{\partial f}{\partial x} = \frac{[f(x + \Delta x) - f(x)]}{\Delta x} \tag{3.13}$$

Since we are dealing with a discrete function, the smallest Δx we can get is $\Delta x = 1$. This definition tells us that in a digital image the minimum difference we can use for taking a one dimensional derivative is one, which is the distance between two adjacent pixels. That is

$$\frac{\partial f}{\partial x} = \frac{[f(x + 1) - f(x)]}{1} \quad = \quad f(x + 1) - f(x) \tag{3.14}$$

For the same reason, we can easily see that the second derivative in one dimension is then defined as

$$\frac{\partial^2 f}{\partial x^2} = [f(x + 1) - f(x)] - [f(x) - f(x - 1)] = f(x + 1) + f(x - 1) - 2f(x) \tag{3.15}$$

For f (x, y), a 2 dimensional digital image, the Laplacian is the sum of the partial second derivatives in both x- and y- directions

$$\begin{aligned} \nabla^2 f &= [f(x + 1, y) + f(x - 1, y) - 2f(x, y)] \\ &\quad + [f(x, y + 1) - f(x, y - 1) - 2f(x, y)] \\ &= f(x + 1, y) + f(x - 1, y) + f(x, y + 1) - f(x, y - 1) \\ &\quad - 4f(x, y) \end{aligned} \tag{3.16}$$

An intuitive way to look at the Equation (3.16) is that the differences between a pixel and its four neighbors are summed. Therefore, the outcome of the Laplacian operation can be viewed as a map that records the magnitude of difference between a

pixel and its neighbors. So, it is expected that pixels belonging to the edges in an image (defined as a region having a discontinuous distribution of grey-scale values) would be non-zero value after filtering. In a smooth region, if a pixel has a value equal to all its neighbors, the output of the operation is zero. This means that the low frequency component is filtered out.

Figure 3.7 gives examples of 3x3 high pass Laplacian filters. These high pass filters are all Laplacian since all of them implement the Laplacian equation above. The two on the top are direct implementations of the Equation (3.16). The two on the bottom are the Laplacians that take the two diagonal directions into consideration. Each direction adds another $-2f(x,y)$ to the sum. It becomes

$$\nabla^2 f = [f(x+1,y) + f(x-1,y) + f(x,y+1)f(x,y-1)] - 8f(x,y). \quad (3.17)$$

It is not important whether it should be the center or the neighboring pixel to be negative. What matters here is that after convolution, the contrast will be computed and recorded.

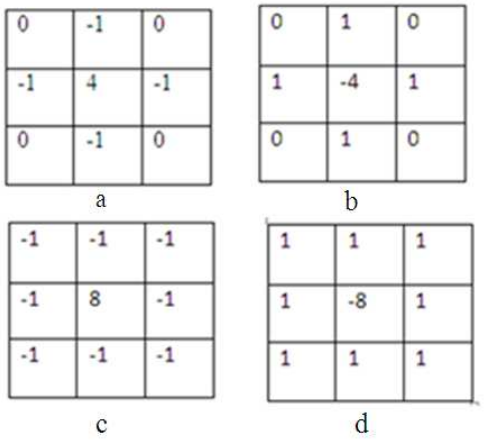


Figure 3.7 (a) and (b) are Laplacians in x- and y- directions, (c) and (d) are Laplacians in x- and y- plus two diagonal directions.

The output of high pass filtering can be very useful for image enhancement. The process called unsharp masking is expressed as

$$f_s(x,y) = f(x,y) - \overline{f}(x,y) \quad (3.18)$$

where $f_s(x,y)$ denotes the sharpened image obtained from unsharp masking, and $\overline{f}\,(x,y)$ is the blurred version(which could be obtained from linear low pass filtering) of $f(x,y)$. The origin of unsharp masking comes from developing film in the dark room. People found that a combination of the image itself and the negative of its blurred version will give a sharper image. A more general form of unsharp masking is called the high-boost filtering, which is defined as

$$f_{hb}(x,y) = Af(x,y) - \overline{f}\,(x,y) \tag{3.19}$$

where $A \geq 1$. We can then rewrite the equation to

$$f_{hb}(x,y) = (A-1)f(x,y) + f(x,y) - \overline{f}\,(x,y)\,, \tag{3.20}$$

Therefore, we obtain

$$f_{hb}(x,y) = (A-1)f(x,y) + f_s(x,y), \tag{3.21}$$

The high-boost filter given above has a very good ability to sharpen images. Figure 3.8 gives two examples of the high-boost sharpening filter by choosing $\overline{f}\,(x,y)$ to be the Laplacian filtered image. The high boost filter then can be expressed as

$$f_{hb}(x,y) = Af(x,y) \pm \nabla^2 f \tag{3.22}$$

It will be plus in the equation if the center of the Laplacian is positive; and the sign will be a minus if otherwise.

0	-1	0
-1	A+4	-1
0	-1	0

-1	-1	-1
-1	A+8	-1
-1	-1	-1

Figure 3.8 Four-neighbor and 8-neighbor Laplacian filters.

Figure 3.9 gives examples of how the spatial domain linear filters discussed so far work. Figure 3.9 (a) is the original image of coins. Figure 3.9 (b) is the result of

low pass filtering with a 3 by 3 average filter. We can see that the details of heads and tails on the coins are smeared by such an operation. It tells us that while a linear low pass filter is able to smooth out the noise, it will also cause blurriness and loss of details. Figure 3.9 (c) is the result of filtering the coin image with a high pass filter. As a result, the high contrast portions on the image, especially the edges of the coins, are emphasized while low contrast regions, the background and the regions within the rims are all filtered out. Figure 3.9 (d) is the result of processing Figure 3.9(a) with a high boost filter when A=1 and $\overline{f}(x,y)$ is the Laplacian. The visual quality has improved by enhancing the details.

So far, we have discussed several spatial linear filters used for noise reduction and quality enhancement in image processing. Low pass filters can smooth the image, and therefore, reduce the noise. The tradeoff for using a linear low pass filter is that we can lose the details that might be crucial. On the other hand, a high pass filter can help sharpening an image by enhancing the contrast. However, the noises in the image will be enhanced as a result. To obtain a desirable result, we often will need to apply a proper combination of those filters. In addition to the linear filers, other filters, such as the order-statistics filter and adaptive filter, are often used, which will be discussed in the following sections.

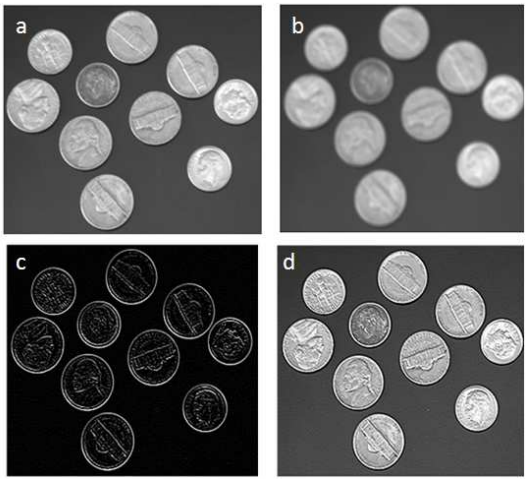


Figure 3.9 (a) The original image (b) result of smoothing (c) result of Laplacian (d) unsharp masking. (Image obtained from Matlab database)

Order-Statistics Filters

Order-statistics filters refer to the operations that use the statistics in masked sub-images. The response of those filters is based on ordering the pixels contained in the filter masked region. The response of the filter at any point is determined by the ranking result. In this section, we discuss one kind of order-statistic that deals with impulse (salt-and-pepper) noise—the median filter—at length since it is one of the most commonly used filters.

Median Filter

As discussed in the previous section, using a low pass filter will help us reduce noise and gives a smoother image. It is also true that low pass filtering will blur the edges of objects, leading to loss of important information. Also, not all kinds of noise can be eliminated by a smoothing filter. Figure 3.10 (a) shows an image of four coins. If we add one type of noise caused by random impulses called “salt and pepper” noise (because it looks like sprinkling salt and pepper on the image) to Figure 3.10 (a), we obtain the image as shown in Figure 3.10 (b). After filtering the noise-degraded image with a linear low pass filter, the noise is still not smoothed out. Because impulse noise has extreme values relative to neighboring pixels, it is not eliminated by averaging with its neighbors. Figure 3.10 (d) is the result of applying a 3x3 mean smoothing filter. Even though the magnitude of the noise is lowered, the noise is still there. One type of filter that is good at dealing with impulse noises is the median filter.

Different from the linear filter that performs convolution, median filtering takes statistical information of the masked region and finds the median values of the pixels within the mask. The output image f (x, y) of median filtering at any point (x, y) can be expressed as

$$f(x,y) = \underset{(s,t)\in S_{x,y}}{\text{median}}\{g(s,t)\}, \tag{3.23}$$

In the equation above, *f(x, y)* is the value of the pixel, g(s, t) represents values in the region being masked. S is the set of coordinates within g(s, t). A median filter works by replacing the center value of masked region with the median value of that masked

region. Figure 3.10 (c) is the result of applying a 3 by 3 median filter to the degraded image in Figure 3.10(b). The correct image is almost totally reconstructed. A contrast of performance between the median filter and the mean filter is illustrated by comparing Figure 3.10(c) and Figure 3.10(d). It can be seen that the edges are blurred in Figure 3.10 (d) but are preserved in Figure 3.10 (c).

Median filtering works very well in eliminating noise in a cargo X-ray image, especially for low-dose X-ray systems. When penetration quality is poor, the X-ray image will be noisy. The extreme values in a region can be eliminated by a median filter.

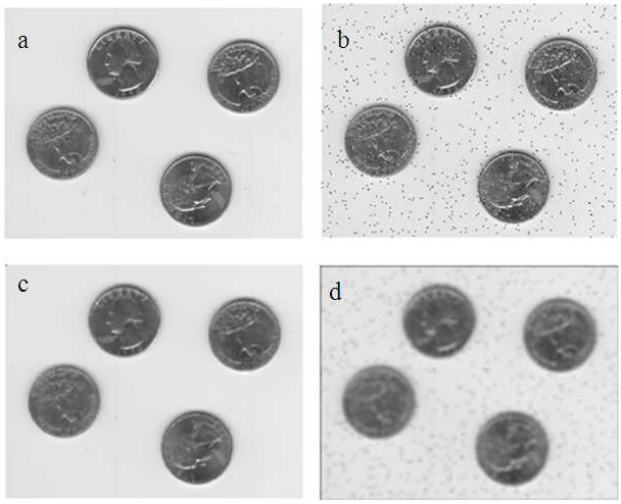


Figure 3.10 Demonstration of median filter (a) original image (b) image degraded by impulse noise (c) result of median filtering (d) result of smoothing filtering. (Image obtained from Matlab database)

Minimum and Maximum Filters

Other order-statistics filters include minimum and maximum filters. The minimum filter replaces the center pixel with the minimum value in the filter masked region, while the maximum filter replaces with the maximum. In fact, the minimum filter has exactly the same effect as grayscale morphological erosion and the maximum filter is just the grayscale morphological dilation when the structuring element is a 2-dimensional square. Morphology will be discussed in later sections.

Removing Noise by Adaptive Filtering: Wiener Filter

The concept of adaptive filtering is that for a signal, the nature of the noise might vary with time or spatial location. For example, a 10-minute speech signal could encounter different sorts of noise during the speech. Similarly, an image can have different noises in different regions. Take *X-ray* images for example, the thicker the object, the noisier the image of that region will be. Filtering the signal adaptively means dividing the signal to different blocks, within each block or window, the nature of the noise assumed to be the same or, at least, approximately the same. Based on the property of each region, we adjust the coefficients of filter adaptively. The coefficients of the adaptive filter are obtained by estimating the function of local or regional noise, and as a result, the adaptive filter will give an outcome that has minimum mean square error (MSE).

Let *M*, *N* denote the dimension of our adaptive filter window and S_{xy} denote the region being masked by the *M* by *N* window at a given time. Let $g(x,y)$ be the noise degraded image, and $f(x,y)$ be the output of the adaptive filter. Only four values are needed to obtain the output: (a) the pixel values of $g(x,y)$, (b) the local mean μ_L, which is the average of the pixel values of $g(x,y)$within S_{xy}, (c) the variance of the noise that caused $g(x,y)$, σ_η^2 and (d) the local variance σ_L^2 within the region $S_{xy.}$ Here, we have

$$\mu_L = \frac{1}{MN} \sum_{x,y \subseteq Sxy}^{M,N} g(x,y) \tag{3.24}$$

$$\sigma_L^2 = \frac{1}{MN} \sum_{x,y \subseteq Sxy}^{M,N} g(x,y)^2 - \mu_L{}^2 \tag{3.25}$$

$$f(x,y) = g(x,y) - \frac{\sigma_\eta^2}{\sigma_L^2} \left[g(x,y) - \mu_L \right] \tag{3.26}$$

The only thing that must be estimated is the variance of the noise of the whole image, σ_η^2, which is usually unknown. It can be obtained by

$$\sigma_{\eta}^{2} = \frac{1}{MN}\sum_{x,y}^{M,N}\sigma_{\mathrm{L}}(\mathrm{x},\mathrm{y})^{2} \tag{3.27}$$

where $\sigma_{\mathrm{L}}(\mathrm{x},\mathrm{y})^{2}$ is the local variance of each pixel (x, y) within the window masked region. The assumption is that the noise of the whole image is the approximate average of the local variance of all the subimages.

The adaptive filter discussed above is called a Wiener filter because it is based on the idea proposed by Norbert Wiener during the 1940s and published in 1949. In Matlab, the "wiener2" function applies a Wiener filter to an image adaptively, based on the equations (3.24) to (3.27) above, tailoring itself to the local image variance. When the local variance is much greater than the overall variance, the term $\frac{\sigma_{\eta}^{2}}{\sigma_{L}^{2}}$ becomes insignificant and the filtered image *f*(x, y) will be almost the same as the original image *g*(x, y); wiener2 performs little smoothing. Where the local variance is small, $\frac{\sigma_{\eta}^{2}}{\sigma_{L}^{2}}$ becomes large and wiener2 performs more smoothing. This approach often produces better results than linear filtering methods discussed before. The adaptive filter is more selective than a comparable linear filter, preserving edges and other high-frequency parts of an image. In addition, there are no design tasks; the wiener2 function handles all preliminary computations and implements the filter for an input image. However, the wiener2 function requires more computation time than linear filtering.

Figure 3.11 (a) is an image degraded by Gaussian noise. Figure 3.11 (b) is the result of applying a 5 by 5 Wiener filter. Figure 3.11(c) is the result of applying a 5x5 median filter. We can see that both the Wiener filter and median filter did a good job filtering out the noise. But, the Wiener filter gives a smoother background than the median filter.

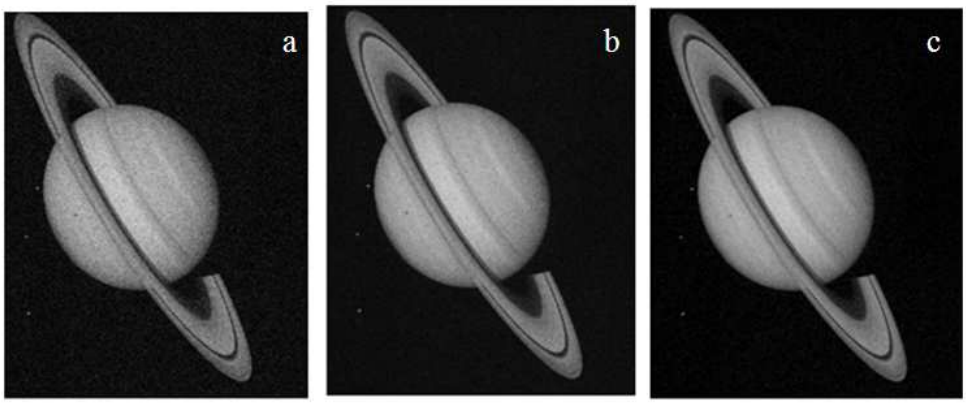


Figure 3.11 (a) A noise-degraded image of Saturn (b) result of applying a 5 x5 Wiener filter. (c) Result of applying a 3x3 median filter. (Image obtained from Matlab database)

3.4 Morphological Filtering

Morphological Operations

Morphological operations refer to procedures that deal with the shape and structure of objects. The mathematical morphological operation is a tool to extract useful information for representation and/or description of region shape such as boundary, skeletons and convex hull. Morphological operations were originally applied to binary images to extract information of an object's structure or to modify the shape of an object. The concept was then extended to gray-scale images. In this section, we will only discuss four morphological operations that are most commonly used as noise reduction filters: erosion, dilation, close, and open.

Each morphological operation is defined by a foreground and a structuring element. Foreground is a set of connected pixels of value 1 in a binary image. The structuring element is a set of connected pixels of a certain size and shape such as square or circle that will operate on the foreground based on the defined operations. All the morphological operations come from the linear combination of two operations: dilation and erosion.

Dilation and Erosion

Dilation expands the foreground by a structure element, which is a mask of zeros and ones. Dilation slides the mask center through every point of the foreground (the center of the structuring element mask must be within the foreground all the time), then mark the regions covered by the mask as foreground in the output image.

Mathematically, dilating a foreground ***A*** with a mask ***B*** can be expressed as

$$A \oplus \mathrm{B} = \{ \mathrm{z} \mid \left(\widehat{\mathrm{B}} \right)_{\mathrm{z}} \cap A \neq \emptyset \} \tag{3.28}$$

where $\oplus$ is the notion for dilation, z is all the non-zero pixels of the output, and $\left(\widehat{\mathrm{B}}\right)_{\mathrm{z}}$ is the translation of the reflection of B. A translation $(\mathrm{B})_{\mathrm{z}}$ is the displaced version of an object centered at point *z*; a reflection of B denoted $\widehat{\mathrm{B}}$ is the flipped version of B. If the structuring element is symmetric, using B instead of $\widehat{\mathrm{B}}$ does not make any difference in the result.

What Equation (3.28) says is that the output of dilation is a set of pixels ***z*** that is the region covered by all the translations of $\widehat{\mathrm{B}}$ while $\widehat{\mathrm{B}}$ has at least one element overlapped with A. In other words, the center of $\widehat{\mathrm{B}}$ has to be within A. Figure 3.12 (a) demonstrates how a square foreground is dilated by structuring elements.

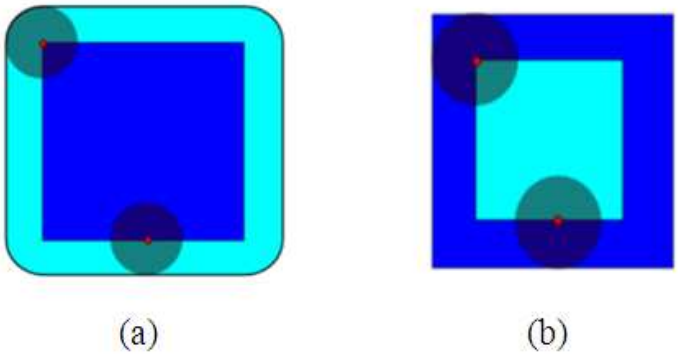


Figure 3.12 (a) The dilation of the dark-blue square by a disk structuring element, resulting in the extra region of the light-blue square with rounded corners. (b) A square is eroded by a disc structuring element. The outcome of this operation is the light-blue region. (Images obtained from Wikipedia, under Dilation and Erosion)

Conversely, erosion is an operation to expand the background, contract the foreground. It works by placing the mask so that whole of it lies inside the foreground. The mathematical expression for erosion is

$$A \ominus \mathrm{B} = \{ \mathrm{z} \mid (\mathrm{B})_{\mathrm{z}} \cap A \subseteq A \} \tag{3.29}$$

Again, *z* is the element belonging to the set of output pixels; $(\mathrm{B})_{\mathrm{z}}$ is the translation of structuring element B centered at point z. What Equation 3.29 says is that the output of eroding A by B is a set of pixels ***z*** such that all pixels of any translation of B centered at ***z*** will fall entirely within A. Figure 3.12 (b) demonstrates how erosion is performed. The set *A* (the bigger square) is eroded by a structuring element *B*, a disc.

Opening and Closing

The opening of A by B is obtained by the erosion of A by B, followed by dilation of the resulting image by B:

$$A \circ B = (A \ominus B) \oplus B, \tag{3.30}$$

Opening is able to eliminate small artifacts by eroding the objects smaller than the structuring element in the image and later using dilation to reconstruct the remaining eroded objects. Opening can also smooth the boundary of an object. Opening can also be expressed as

$$A \circ B = \cup \{ (\mathrm{B})_{\mathrm{z}} | (\mathrm{B})_{\mathrm{z}} \subseteq A \} \tag{3.31}$$

which means that it is the locus of translations of the structuring element B inside the image A. Figure 3.13 (a) shows a dark-blue square being opened by a disc structuring element. The resulting light-blue area is the total area covered by the disc moving inside the square when the whole disc is inside the square. The four corners are smoothed by the opening operation.

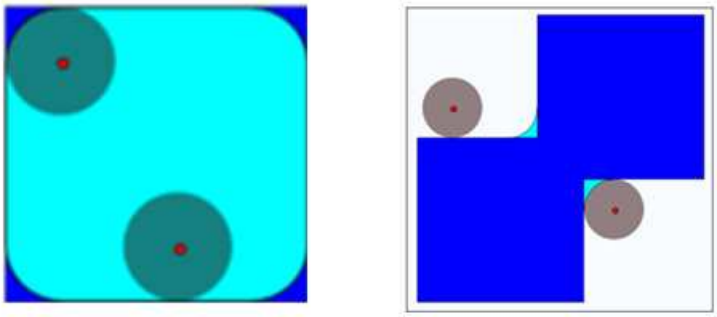


Figure 3.13 (a) Opening a dark-blue square by a disc structuring element. The light-blue region is the result of opening. (Images obtained from Wikipedia, under Opening and Closing)

The closing operation, denoted by •, of A by B is obtained by the dilation of A by B, followed by an erosion of the result of A by B:

$$A \bullet B = (A \oplus B) \ominus B \tag{3.32}$$

The closing can also be expressed as

$$A \bullet B = (A^c \circ \hat{B})^c \tag{3.33}$$

where X^c denotes the complement of X. For binary images, it means $X^c = 1 - X$. Again, $\hat{B}$ is the flipped version of B. A region of two connected squares is denoted as the dark blue region in Figure 3.13 (b). The result of closing it by a disc structuring element is shown as the total region of light and dark blue. As we can see, the dark and light regions are the complement of what is left after applying opening to the non-dark region. Closing can also eliminate the small objects and smooth the boundary of an image.

Gray Scale Morphology

Morphology of gray-scale images is based on the same concepts as those for binary images. Grayscale structuring elements are functions of the same format, called "structuring functions", but are now a function, not a set. The structuring function is operating on a 2 dimensional function f (x, y).

The grayscale dilation of f by b is given by

$$(f \oplus b)(s,t) =$$

$$\max\{f(s-x,t-y)+b(x,y) \mid (s-x),(t-y) \in D_f; (x,y) \in D_b\} \quad (3.34)$$

where D_f and D_b are the domains of f and b respectively.

What Equation (3.34) says is that each pixel (s,t) of f is used as the center of the structuring function b. Within the mask-covered region, for each point, the value of b corresponding to f is added. Then, the maximum of all the values after the addition is the output of dilation of pixel (s,t). A flat function is most commonly used for structuring function b. That is

$$b = \begin{cases} 0, & (x,y) \in D_b \\ -\infty, & otherwise \end{cases} \quad (3.35)$$

If b is a flat function, the output of dilation is equivalent to the local maximum of the mask covered region.

Similarly, the erosion of f by b is given by

$$(f \ominus b)(s,t) =$$

$$\min\{f(s-x,t-y)+b(x,y) \mid (s-x),(t-y) \in D_f; (x,y) \in D_b\} \quad (3.36)$$

Here, erosion is defined as the minimum of the region masked by the structuring function.

Just as in binary morphology, the opening and closing of f by b are given respectively by

$$f \circ b = (f \ominus b) \oplus b \quad (3.37)$$

and

$$f \bullet b = (f \oplus b) \ominus b \quad (3.38)$$

An opening followed by a closing is often used to eliminate all the dark and bright artifacts in an image. This operation is called morphological smoothing.

3.5 Image Segmentation

Image segmentation refers to techniques that separate the region of interest (ROI) or object of interest (OOI) in an image from the background. It is an essential preliminary step in most automatic pictorial pattern recognition and scene analysis problems. Depending on the objective and characteristics of the problem, one technique is chosen over another. Some popular methods include boundary (edge) detection, thresholding, and k-mean.

Since the goal of image segmentation is to partition an image into meaningful sub-images (regions), we must define what forms a region. One necessary condition to define a region is connectivity. A region must be formed by connected pixels. As defined in Section 3.2.4, the connectivity is determined by similarity. To identify a tumor in a chest X-ray image, the region of the tumor is a group of connected darker pixels, the similarity here is defined by the intensity. For different applications, other criteria for similarity such as color, size or texture can also be used.

Image segmentation is also an important technique for a cargo material discrimination X-ray imaging system. Several systems use this technique to segment the regions of possible threats. This technique is also used as a filter to get a smooth image. In this section, we will briefly discuss several popular methods.

3.5.1 Segmentation by Thresholding

A thresholding technique is preferred because of its intuitive properties and simplicity of implementation. The idea of thresholding is to select a grey scale value T so that, in an image $I(x, y)$, any point (x, y) for which $I(x, y) > T$ will be called an object point, and otherwise, the point will be called a background point. The object points will then be assigned a value 1, while the background pixels will be set to zero.

The main issue with this segmentation approach is the choice of threshold. A user can manually try different values and then select the best one according to his need. However, for most intelligent systems, an automated threshold selection method is required. There are two kinds of threshold selection depending on the user's objective and the nature of the images. The first is global thresholding, meaning the whole image will be segmented using a single threshold. When the image is relatively simple and the contrast between the background and objects is good, using a single threshold is usually sufficient. However, if an image is affected by other factors, for example, the illumination, a single object or objects of interest might have different grey scales over different regions. In that case, using a single threshold will not be able to separate the objects of interest

Global Thresholding

Global thresholding uses the same threshold over the whole image grid. The assumption of this method is that there is sufficient contrast between the object and background. That is, the histogram of grey scale values of the image is nice and clean; object pixels and background pixels pile up two hills in the histogram and the distance between the two peaks is wide. In that case, any value in the valley, the regions between the two hills, will be a good candidate for the threshold. The algorithm below [37] can determine the global threshold automatically:

1. Select an initial threshold T

2. Segment the image using this T. Two groups of pixels will be formed: $G_1(x, y) >= T$ and $G_2(x, y) < T$.

3. Get the average of grey scale values u_1 and u_2 of regions G_1 and G_2 respectively.

4. Set new threshold to be $T= (u_1+u_2)/2$

5. Get the difference between the new threshold and the previous one

6. Repeat step 2 to 5 until the difference is less than a preset value, i.e. T is converged.

The initial value of T can be obtained by either computing the average of the grey level values of the image, in which case, the size of object and that of the background should be comparable, or by computing the average of the maximum and minimum of the grey level values.

Adaptive Thresholding

Using a single threshold for the whole image in some cases will fail to segment the whole object because pixels in different regions are significantly different in grey level value due to the uneven illumination. The idea of adaptive thresholding is to divide the whole image into subimages and select different thresholds according to the grey scale value distributions of each of the subimages. Choosing threshold(s) adaptively will solve the problem addressed at the beginning of this paragraph. To minimize the error in segmentation, an approach shown in [37] is to get a good estimate of the probability density functions (PDF) of the grey level values of the background pixels and the object pixels in each subimage. If both are Gaussian-like, set the threshold to

$$T = \frac{u_1+u_2}{2} + \frac{\sigma^2}{u_1-u_2}\ln\left(\frac{P_1}{P_2}\right) \tag{3.39}$$

Where σ^2 is the variance of both PDFs and u_1 and u_2 are the means, and P_1 and P_2 are the probability of object and background pixels respectively.

The most difficult part of this method is the estimation of the PDFs. A good estimate is sometimes impossible to obtain or even if it is possible, very time consuming. That makes this method impractical in applications that require fast processing time.

3.5.2 Segmentation by Clustering

Clustering refers to a technique that classifies data into different groups. A cluster of data is a group of similar elements in the data set. Each cluster has a center called the “centroid”. Figure 3.14 shows two groups of 2-dimensional vectors, each group forms a cluster centered about a centroid. The centroid can be seen as the center of mass of each cluster, representing a feature that is distinct from other centroids. A vector is classified to a cluster if its distance to the centroid of that cluster is smallest compared to its distances to other centroids.

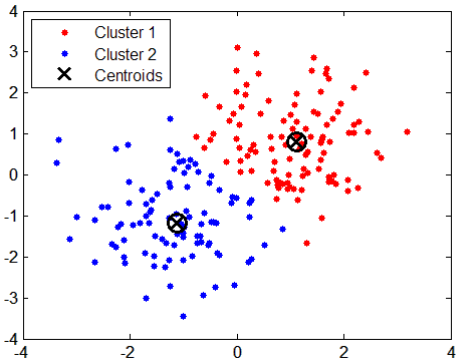


Figure 3.14 Two clusters of a data set.

When clustering is applied to an image, pixels are partitioned into different groups by defined properties of similarity. Each cluster can be similar in intensity, texture, or other features. Among all clustering methods, K-means is the most popular one.

K-means clustering partitions the data into k classes. It was invented in 1956. The most common form of the algorithm uses an iterative refinement heuristically known as Lloyd's algorithm. Lloyd's algorithm starts by partitioning the input points into *k* initial sets, either at random or using some heuristic data. It then calculates the mean point, or centroid, of each set. It then constructs a new partition by associating each point with the closest centroid. Then the centroids are recalculated for the new clusters, and the algorithm is repeated by alternate application of these two steps until

convergence occurs, which is obtained when points no longer switch clusters (or alternatively centroids are no longer changed).

In terms of performance, the algorithm is not guaranteed to return a global optimum. The quality of the final solution depends largely on the initial set of clusters. Since the algorithm is extremely fast, a common method is to run the algorithm several times and return the best clustering found. A drawback of the k-means algorithm is that the number of clusters k is an input parameter. An inappropriate choice of k may yield poor results.

K-means clustering is widely used in computer vision. The initial *k* set can be obtained by equally dividing the whole image into *k* subimages or by randomly assigning *k* pixels in the image as initial centroids. The pixels that meet the similarity requirements, for example similar gray-level value, will be grouped to its closest centroid.

3.5.3 Segmentation by Region Growing

As discussed before, a region is defined by similarity and proximity. That is, a region is a group of connected similar pixels in an image. Region growing is a segmentation technique based on the region directly.

Let the whole image be a region *R*. The result of segmentation then divides the whole region into *n* subregions, each of them is denoted R_i. They will have the properties:

(a) $\bigcup_{i=1}^{n} R_i = R$

(b) R_i is a connected region, i=1, 2, 3, …,n

(c) $R_i \cap R_j = \emptyset$ for all $i \neq j$

Region growing is a procedure that groups pixels or subimages into larger regions based on the predefined criteria. The idea is to start with a set of ‘seed’ points and from these points grow regions by including neighbor pixels to the seed that have

properties similar to the seeds by the predefined criteria of similarity. A region stops growing when it can no longer find any similar neighbor pixels.

It is possible that a pixel meets the criteria for two or more regions. In that case, it means that pixel is on the boundary of adjacent regions. One can classify that pixel to the cluster that the pixel is most similar to.

Two subregions can also be merged by some predefined criteria. That procedure is called “region merging”. If two adjacent subregions are similar in some sense, for example, if the difference in their regional averaged gray-scale is less than a preset value, or the color of the two regions are approximate the same, the two subregions are then merged into a new region. The new region then continues to check its adjacent subregions and merge subregions until no other subregion meets the criteria to merge.

3.6 Summary

In this chapter, we have discussed the general ideas of digital image processing including the hierarchy of image processing, terminologies frequently referred to in pixel domain image processing, and several useful filters for noise reduction. The concepts of morphological operations and image segmentation have also been covered in this chapter.

A cargo security imaging system is comprised of three major tasks: image enhancement, noise reduction, and segmentation. Image enhancement is to enhance the visual quality of X-ray image, letting small objects be easily identified by inspectors. For the purpose of material discrimination, the task of noise reduction was not to smooth the image, but more importantly, to reconstruct the signal so that accurate information about the attenuation ratio can be obtained.

Image segmentation in a cargo X-ray image refers to identification of the regions of interest. The regions of interest are possible regions where contraband materials could be hidden. Since an X-ray image is a map of relative thickness that is

obtained after photons penetrate through layers of materials, the measured thickness of an object can be affected by the objects put before or behind it. It is a very challenging task to identify a single object in an X-ray image of a cluttered container.

In the following chapters, we will discuss combining image processing techniques for material discrimination in cargo X-ray imaging systems.

4

Adaptive Masking for Cargo Material Discrimination

4.1 Introduction

In Chapter 2, we introduced the concept of dual energy X-ray imaging system for material discrimination. Although the technique has been widely used in medical applications, it has only recently been applied to the detection of high-Z materials for cargo inspection. For cargo image systems, a mega-volt level of X-ray energy is used to ensure the penetration quality and to trigger pair production as well. The ratio of the high energy scan to the low energy scan attenuation (or transmission) can be an indicator of the scanned object's effective atomic number (Z). However, for multiple-layered objects, e.g. a blob of uranium in a steel box, the effective Z will be dominated by the material with greatest attenuation. To detect a heavily shielded high Z material, it is hard to determine what the threshold ratio should be for classifying that object as high Z. As the thickness of the shield becomes greater, the calculated effective Z will get closer to that of the shield. This implies that obtaining the effective Z value of each pixel is not sufficient for the purpose of detection since the effective atomic number can be manipulated by the selection of shielding material and its thickness.

In this Chapter, we will discuss an approach we firstly developed for high Z material discrimination using a 9-MeV to 6-MeV dual energy system. We designed

experiments to test if such system can perform high Z detection at a reasonable speed. All image processing operations are done in the spatial domain. A global minimum ratio threshold is used in this model. Our approach is designed to find not just unshielded but more importantly, hidden high Z materials. We have named our algorithm “adaptive masking”. This method solves the problem associated with using a single threshold ratio for decision making and can effectively eliminate many false alarms. Test results show an increased accuracy of high-Z detection using this approach. Experimental results illustrating the effectiveness of the method will be shown at the end of this chapter. We will start our discussion by introducing some related works.

4.2 Related Works

Research into material discrimination for cargo X-ray imaging systems is still in its early stages; there has not been a long history. There exist some, though not many, publications related to image processing for this specific application. In this section, we will briefly discuss several recent publications in the history of this specific field.

The pioneering work of Ogorodnikov and Petrunin in the early 2000s [3-7] proposed a method using the ratio of the measured “transparencies,” which is equivalent to the ratio of attenuation coefficients, of two levels of X-ray energy (4MeV/10MeV) scans. Four basic groups of materials are classified: Organic, Organic-inorganic, Inorganic, and Heavy metals. Hydrocarbon (CH2, Z~=5), silicon (Si, Z=14), iron (Fe, Z=26), and lead (Pb, Z=82) were chosen to be the base material for each group. For the purpose of visualization, a TZ plane is used. T is the transparency, which is the same thing as transmission, and Z is the equivalent atomic number for each corresponding pixel on the transmission image. The TZ plane is obtained from the lookup table created by experiments. For a specified material, the plot of ratio versus transmission is obtained. Figure 4.1 is an example of the visualization of such system. As we can see, water is an organic material therefore the color of water is close to red. The six purple bricks are made of heavy metals.

To obtain an image with the quality as shown in Figure 4.1, as the authors pointed out, the source needs to be moving at a very low speed to have sufficient samples per pixel to suppress pixel noise. Also, multiple scans might be taken and then averaged to remove the noise and boost SNR. Both requirements limited the feasibility of a practical system. The main point of their work was the possibility of such technology.

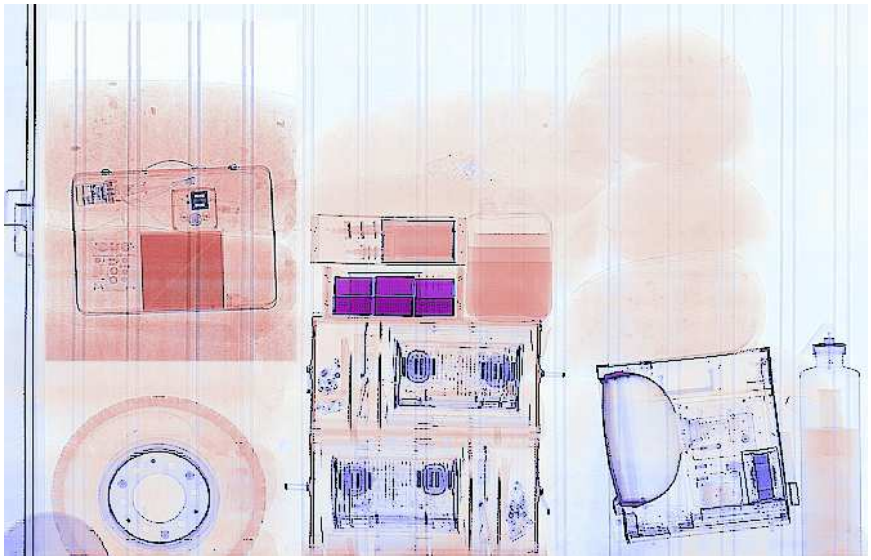


Figure 4.1 The visualization of system proposed by Ogorodnikov and Petrunin. Image of container in TZ-palette: organic–red, inorganic–blue, heavy metals–lilac. (Image obtained from [6])

In 2002, Ogorodnikov et al [5], applied some image processing methods to their earlier work. A bilateral filter [7-8] was used to reduce the pixel noise on the image. A bilateral filter works like the adaptive filter mentioned in Chapter 3, which is an edge-preserving filter. It works by averaging the pixels with a mask if all pixels satisfy the preset requirement of similarity, otherwise, it keeps the value of the center pixel. Leader clustering [9-10], a region growing algorithm for image segmentation was also used in the paper. The reason for doing that is to obtain a smoother filtering result.

Based on Ogorodnikov‘s theory, Zhang et al. [26] proposed an H-L (high and low energy transparencies) curve method for material recognition in dual energy X-ray inspection systems. The H-L curve they used is a plot of the ratio of high and low

energy attenuation coefficient ratio versus atomic number. Their method is reported to demonstrate good performance in discrimination of practical objects. The concept of intrinsic difference between materials is introduced to quantitatively investigate the possibility of material recognition under a given noise level. As discussed in Chapter 2, the H-L plan method presumes that the same material will have the same ratio regardless of the hardening effect. This assumption is true only when the source is purely monochromatic or penetration is good, so that the spectrum is not affected by the thickness of the objects.

For detecting low Z plastic explosives in luggage, Liang et al. [30] proposed a method to segment X-ray images in carry-on luggage for airport security systems, which is based on the Radon Transform to determine the optimal number of clusters and to evaluate the segmented images. For the same purpose of explosive detection in baggage, Ying et al. [35] used a CT scanner for the dual energy system. The resulting Z images display three-dimensional X-ray images that make it easier for inspectors to identify the objects. Taking a different image processing approach, Chen et al. [25] proposed wavelet-based image fusion that applies background subtraction-based noise reduction and an enhancement technique to reconstruct the final image for airport luggage scan.

The dual energy material discrimination methods discussed above only work either under ideally controlled situations or the in luggage inspection where penetration quality is easier to achieve than its cargo container counterpart. The inevitable statistical errors are either ignored or treated as a challenge for future work in most related papers. A system with reasonable scanning speed for use in the flow of commerce will inevitably have statistical errors that lead to misidentification of materials, since the differences between ratios of different materials can be very small.

Bjorkholm [13] proposed a method to deal with the statistical errors and simplified the task of a DE cargo inspection system. The goal is just to discriminate high atomic number metals from other materials. One important assumption in his model is made: general goods cargo containers won’t contain a lot of metal. If metals

are detected inside the container, most of the time they are iron or copper. Therefore, for a measured transmission (the inverse concept of attenuation), if the ratio of high- and low-energy transmission is greater than a threshold (e.g., that of iron), a pixel can be eliminated as representing a high-Z metal. It is a concept equivalent to eliminating pixels with a ratio of attenuations less than a threshold. For consistency, we will use ratio of attenuations to describe this method. The heart of the method is to identify all pixels of passing the threshold first and then eliminate the false alarms. The first step is identifying all pixels having ratio greater than a threshold; the second part uses statistics. To characterize the system's noise, the average ratio $\bar{R}$ and the standard deviation σ associated with a measured attenuation value for iron is obtained through series of experiments. Then, for a measured attenuation, a threshold T for a ratio is assigned to be equal to or a little bit greater than $\bar{R}$. The probability of a pixel having a ratio R which is greater than the threshold T is given by integrating the normal distribution from $(T - R)/\sigma$ to ∞. This value is called p_i, which is the probability of false alarm of a single pixel. If a cube of iron with thickness t causes attenuation m in high energy image, and the measured ratio for that is R, the same m caused by other materials with higher atomic numbers will yield greater ratio R'; therefore, the probability of R' greater than T will be greater than that of R. All pixels above the threshold will be marked. False alarms in the image are filtered out by using a 9 by 9 examination window. Only when the number of marked pixels in that window exceeds some predefined threshold can those pixels be considered a threat. The probability of detecting n pixels above T within an 81-pixel examination window will be given by binomial distribution:

$$P(n|81) = \frac{81!}{n!\,(81-n)!} p_i (1-p_i)^{81-n}. \tag{4.1}$$

While this method provides a good direction, several questions remain. How do we determine the threshold number of pixels passing ratio threshold in an examination window? And, if the false alarm object is significantly larger than the size of the examination window, which is usually the case, the false pixels won't be removed. In addition, shielding effects can lower the ratio in the region of a shielded object,

making it hard to detect. In this case, a real detection can be easily eliminated by such a method.

4.3 Adaptive Masking Model

To resolve the issues with Bjorkholm's method, we modified his method and came up with our adaptive masking model. The goal for this method is simple: to locate all possible high Z objects while eliminating all false alarms. Our approach has proven to be very effective in signal enhancement, false alarm elimination, and visual quality improvement. The whole task of material screening is comprised of three major parts: noise-reduction, z-image creation, and output visualization. As discussed before, there has not been work that covers all three parts of the task. It is reported that to get good results, a lot of conditions need to be well-controlled, such as X-ray images have to be noise-free. That can be achieved by using a very slow scanning speed, or taking a sequence of images and then averaging those images to reduce the Gaussian noise. For the signal to be strong, a high-dose X-ray system is preferred. However, doing so would violate safety regulations.

Unlike the methods from other publications in which X-ray images are obtained under highly ideal situations, we are trying to develop models that work in more realistic situations. That is, we are dealing with a system where data is obtained from a realistic speed, low-dose X-ray source. The consequence of that is very high level of noise that causes very weakly discriminating ratio images.

4.3.1 Z-image Creation

The goal ambitious to identify the atomic number of all objects inside the cargo container is in fact nearly impossible to achieve and unnecessary. Since most threat materials either have very high atomic number or are radioactive and need to be shielded by heavy metal, it would be reasonable to design a system that detects materials with atomic numbers higher than lead (Z=82), a high Z material often used for shielding radioactive materials. For that reason, our method looks for materials

with ratio of attenuations higher than a threshold (minimum high Z ratio) in the images of a dual energy imaging system. That is, we look for the lilac or darker regions in Figure 4.1, and screen those regions out. A high Z image, which is a binary image that flags pixels belonging to high Z materials, is the final output we obtain. Alarmed pixels in the high Z image are the regions where the effective atomic number is higher than that of common metals.

As discussed in Chapter 2, due to the high uncertainty associated with thickness, the calculated ratios of pixels in thicker or denser regions are relatively not reliable. Therefore, instead of using the ratio as the indicator of effective atomic number for each pixel, we use Equation (2.40) to calculate the probability of each pixel being high Z. A probability image is then used to indicate each pixel's probability of having a ratio of attenuations greater than that of a predefined threshold. However, the system operator may not have the knowledge to interpret the meaning of those probabilities; he is not able to make a decision based on that information. An algorithm needs to be intelligent enough to make decision for its users.

The problem remaining is what is the cutoff probability we need to call a pixel high Z and trigger the threat alarm. What qualifies as a "high" probability? What is the meaning of "low probability"? Also, taking the shielding effect into consideration, what should be the threshold for the minimum high Z ratio? How do we segment the high probability regions from the image? Those are all questions that need to be answered. Determining these thresholds (for ratio and for probability) is not an easy task. As we discussed in Chapter 2, the probability is a function of thickness. A very high Z material can be heavily shielded and give a low calculated probability. Table 4.1 provides the measured ratios of several steel-shielded high Z materials from one experiment. The atomic numbers of the objects are given. Each column is the ratio using different thicknesses of shielding. Employing Bjorkholm's presumption that it is not usual a cargo container has a high percentage of metal, and Liang's assumption that the ratio of a specified material is constant, we decided the minimum ratio be set to the ratio measure when a high Z object is shielded behind 10" of steel. This

thickness is almost equivalent to the maximum penetration. From data provided by Table 4.1, the minimum ratio can be a number between 1.01 and 1.02.

Table 4.2 shows the corresponding probabilities calculated by Equation (2.40) for those objects. Recall that

$$P_{highZ} = 1 - 0.5\left[1 + \mathrm{erf}\left(\frac{R_{min} - R}{\sqrt{2}\,\sigma_R}\right)\right] . \qquad (2.40)$$

A global minimum ratio being used here is 1.02, which is approximately the ratio of 100cc of lead shielded by 10 inches of steel from experimental data. The objects being tested here are all cubic, so that a 100cc object will have a thickness about 4.64cm or 1.84 inch. The ratios are taken from the values of the center pixels of the objects. Table 4.1 shows that for shielding steel thicker than 8 inches, the errors make the ratio unreliable. Due to noise, thicker shielding can end up having greater probability than thinner shielding, since the accuracy of probability would also be affected by pixel noise.

Table 4.1 Ratios of equivalent thicknesses for different shielded objects. (N/A=not available)

Shielding\Material	100ccW (Z=74)	100ccPb (Z=82)	200ccW	200ccPb	150ccDu (Z=92)
5″	1.047	N/A	N/A	N/A	1.07
7″	1.054	1.036	N/A	N/A	1.06
8″	1.044	1.05	N/A	N/A	1.052
9″	1.031	1.037	1.04	1.046	
10″	1.026	1.021	1.026	1.012	
11″	1.024	1.015	1.016	1.013	

Table 4.2 The corresponding probabilities for Table 4.1 (N/A=not available)

Shielding\Material	100cc W	100ccPb	200ccW	200ccPb	150ccDu
5″	.97	N/A	N/A	N/A	.998
7″	.975	.868	N/A	N/A	.972
8″	.873	.898	N/A	N/A	.964
9″	.648	.729	.719	.781	
10″	.591	.576	.564	.428	
11″	.557	.487	.511	.397	

Table 4.3 Probability of copper (Z=29) under different steel shields

Thickness of shield(inch)	5	7	8	9	10	11
Probability	0	.005	.004	.26	.33	.3

Regardless of the effect of pixel noise, we can still see a trend in the ratio behavior. For the same object, as the thickness of shielding increases, the ratio decreases. This coincides with the theoretical prediction that the effective attenuation coefficient of an object is a linear combination of all materials of the composition, and the ratio of steel starts to dominate as it gets thicker.

Recall that when X-rays pass through multiple layers of objects, the total attenuation can be expressed as

$$\mu(E,Z)t = \sum_{material\ i} \mu_i(E,Z_i)\ t_i\ . \tag{4.2}$$

For a steel-shielded high Z object, the observed ratio is then

$$r = \frac{t_Z^H + t_{steel}^H}{t_Z^L + t_{steel}^L} = \frac{a\mu_Z^H + b\mu_{steel}^H}{a\mu_Z^L + b\mu_{steel}^L} \tag{4.3}$$

where t_Z^H and t_Z^L are the equivalent thickness of high Z material from high- and low-energy scans respectively, and t_{steel}^H and t_{steel}^L are the equivalent thickness of steel from high and low energy scans respectively, and the two constants ***a*** and ***b*** are the thickness of the materials. The thicker the steel, the more weight it has when computing the ratio. When *b>>a,* the ratio will be dominated by the steel term, which on average has a ratio of attenuation coefficients about 0.99.

Table 4.3 is the calculated high Z probability of a 200-cc copper bar under different shielding conditions. The atomic number of copper is 29, which is not high Z. Putting a copper bar next to a high Z object behind a shield can test how well the system works. Ideally, their values should all be close to zero since their theoretical ratio is less than 1.02. Without the effect of noise, the calculated probability should be

zero. As we can see from Table 4.3, a 9-inch-steel shield caused probability of about 0.3. This is due to the pixel noise that makes the ratio higher than that obtained when the shield is only 5-inch thick. Comparing Tables 4.2 and 4.3, we see that the probability of high Z and copper are close for thick shielding. An indication of that result is that the *SNR* in those scans are not good enough for high Z discrimination.

The result shown in Table 4.2 tells us that using a fixed threshold of probability for segmentation might cause two kinds of problems: if the threshold is set low to identify heavily shielded objects, it might cause a lot of false alarms. On the other hand, if the threshold is set high, then heavily shielded objects will not be detected. In fact, due to the unpredictable nature of cargo containers, neither of the two choices seems to be a good one.

When the images of high and low energy scans are properly aligned, and the frequency of X-ray can be controlled to be band limited, and when scattering is not a big issue, a good way to determine the probability threshold is to create a lookup table through a series of experiments, finding the threshold probability for different thicknesses. This has similar effects to creating a lookup table for minimum ratios.

We here provide a solution that can avoid the need for making a lookup table; it is the *adaptive masking method*. Adaptive masking is based on the fact that we can adaptively choose our probability threshold in regions of different thicknesses. If a definite value of the threshold for different thickness is not available, we need to use a method that does not require that kind of information. Recall that in Chapter 3 we discussed a thresholding segmentation method using an automatic threshold finding algorithm. If a group of pixels has a greater value than its background, the minimum value of that group will be selected as the threshold for segmentation.

If we look at the last row of Table 4.2, we will see that all the high Z targets shielded by 11 inches of steel have a probability greater than 0.4. Ideally, if a non-high Z object is also behind the 11-inch shield, its calculated probability should be less than 0.4. Therefore, applying the automatic thresholding algorithm to the 11-inch- shielding

region probability image, we can segment the high Z region with a threshold of 0.4. By that principle, we can assume that under that same background, the high Z object will have higher probability than other objects and therefore can be segmented by thresholding. The segmented image is then called the Z-image or Z-map, which we will use for visualization. The overlay image of Z-map and thickness map is the output displayed to users.

4.3.2 Image Types

Four kinds of images used in our model will be referred to throughout the rest of this chapter: equivalent thickness image which presents the attenuation, ratio image, probability image, and Z image.

As discussed in Chapter 2, the attenuation can be expressed in equivalent thickness. There are two equivalent thickness images: the high energy equivalent thickness map and the low energy equivalent thickness map. For convenience, we just call them the thickness maps or T-maps. The former, denoted $H(x, y)$ is obtained via the X-ray scans with a 9 MeV source, while the latter, denoted $L(x, y)$ is the result of 6-MeV X-ray scans. Each pixel in $H(x, y)$ and $L(x, y)$ represents the equivalent thickness of the X-rayed object in terms of the thickness of calibration material.

The ratio map or the R map is obtained by taking the ratio of high and low energy thickness maps:

$$R(x, y) = \frac{H(x, y)}{L(x, y)} \tag{4.4}$$

The probability map $P(x, y)$, is the probability of a pixel having ratio greater than the preset minimum ratio, and is obtained by applying equation 2.40 to each pixel:

$$P_{highZ} = \quad 1 - 0.5 * (1 + \text{erf}\left(\frac{R_{min} - R}{\sqrt{2}\sigma_R} \right) \tag{2.40}$$

The range of the P map will be [0 1]. If a pixel (x, y) with $R(x, y) \gg R_{min}$, we can expect that $P(x, y) \cong 1$. For pixels with $R(x, y) \ll R_{min}$, $P(x, y) \cong 0$. If $R(x, y) \cong R_{min}$, $P(x, y) \cong 0.5$.

Finally, the Z-map is the output of our material screening algorithm, which indicates pixels that have high probabilities of being high atomic number material and the system makes a decision to identify them as high Z pixels. The details of how Z map is formed and processed will be discussed in later sections.

Pseudocoloring is applied to those data images for the purpose of easy analysis. The standard colormap, Jet map, is used here. A colormap is an m-by-3 matrix of real numbers between 0.0 and 1.0. Each row is interpreted as a color, with the first element specifying the intensity of red light, the second green, and the third blue. That is, each row is an RGB vector that defines one color. The *k*th row of the colormap defines the kth color, where map(k,:) = [r(k) g(k) b(k)]) specifies the intensity of red, green, and blue. Let Cmax denote the maximum and Cmin denote the minimum value of intensity in an image, the range of that image [Cmin, Cmax] will be rescaled to [0, 1] for the purpose of colormapping. As a result, the minimum values in the image will be represented in blue and the maximum will be shown in red. If an image is a binary image, then red will be 1 and blue will represent 0. Figure 4.2 shows the jet colormap and the corresponding grayscale.

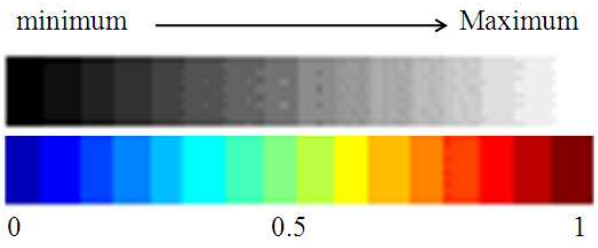


Figure 4.2 Jet colormap (bottom) and its corresponding grayscale(top). Jet ranges from blue to red, and passes through the colors cyan, yellow, and orange. The jet colormap is associated with an astrophysical fluid jet simulation from the National Center for Supercomputer Applications.

4.3.3 Image De-noising

Noise reduction is probably the most important step in the algorithm for material discrimination. As we have discussed, noise can cause erroneous measurements of equivalent thickness. As a result, the measured ratio and the calculated probability will not be reliable high Z indicators. To obtain a desirable R-map, two noise-suppressed T-maps are needed.

There are several useful filters that can help to reduce the level of noise. Low pass filters such as Gaussian filter can efficiently suppress the noise by averaging neighboring pixels. For random noise such as salt-and-pepper noise, a median filter is a good choice. Due to the fact that the noise we are dealing with here is dependent on thickness, we suggest using a Wiener filter since it is an adaptive filter that deals with noise on a region by region basis. Another advantage of choosing the Wiener filter over others is that the Wiener filter, unlike other spatial averaging filters, is an edge-preserving filter. Edge-preserving is crucial when objects are small.

Figure 4.3 (a) and (b) show two T- maps of a setup with a 100cc cube of tungsten (right) and a 2-inch thick copper bar (left) behind 8 inches of steel using high- and low-energy X-rays respectively. Visually the two images (a) and (b) are not that different. The ratio of the two images in Figure 4.3 (a) and Figure 4.3(b) is shown in Figure 4.3(c). Figure 4.3 (d) is the probability map obtained from the ratio map in Figure 4.3(c), using the threshold ratio = 1.02 and standard deviations from calibration data. All images are shown in pseudo-color and a color bar is provided for each image to see the range of the image data.

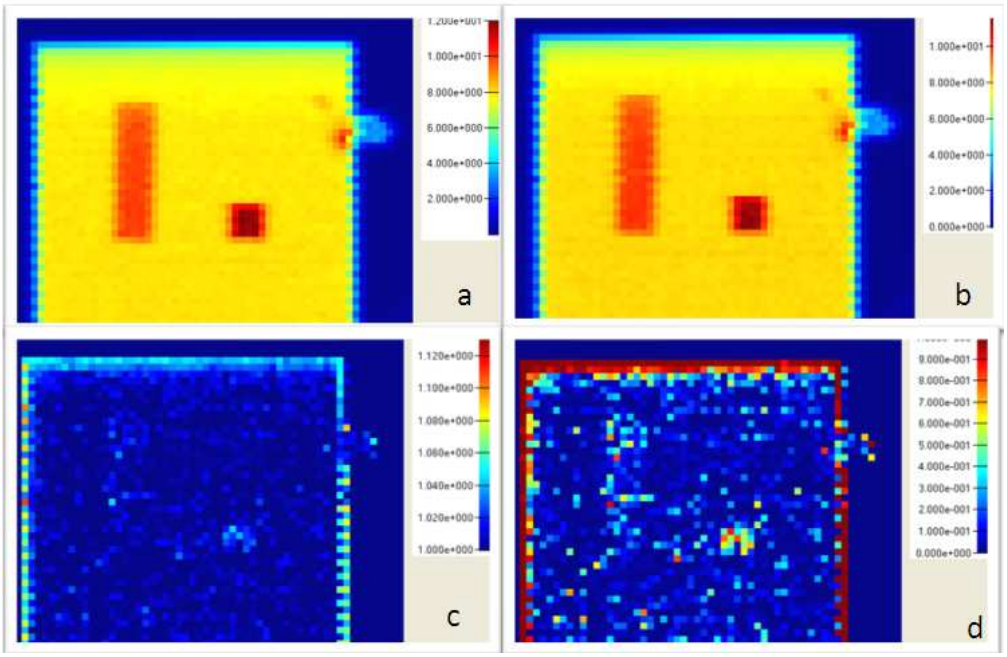


Figure 4.3 (a) The high energy scan thickness map of a setup with 100cc W and 2"-thick Cu bar behind 8 inches of steel (b) the low energy scan thickness map (c) the ratio map (d) the probability map

Since we only are interested in finding high atomic number material with a ratio greater than 1, we only show the range on the ratio map from 1 to the maximum value. The bright spots on the map are pixels with higher ratios than their neighbors. As shown in the image, the pixels corresponding to the tungsten cube are relatively brighter, conversely, there are not many bright spots in the copper bar region. Figure 4.3(d) illustrates what was discussed in Section 4.3.1. There is a blob of high probability pixels clustered in the tungsten cube region, with probability ranging from 0.5 to 0.9 that is higher than in the other regions. Therefore, as we mentioned in the previous section, that group of pixels can be segmented out by using the automatic thresholding algorithm discussed in Chapter 3.

The probability map in Figure 4.3 (d) shows that the edges of the steel shield have high probabilities when in fact those pixels do not belong to a high Z material. The false high probability is due to the fact that photons scatter in a more unpredictable way at the boundary of probed objects since it is the interface of two

materials. If more photons are scattered on the edges when scanned by the higher energy source, the equivalent thickness tends to be higher on the edges of high energy scans and the ratio of the edge will be greater than that on the inner region. The other possible reason for the edge effect is imperfect registration of two energy thickness images. However, this is not the case here.

Figure 4.4 shows the same images as Figure 4.3; the difference in this set of images is that the thickness maps are filtered with a 3x3 Wiener filter before computing the ratio. A noticeable improvement can be found in the R-map and P-map. Figure 4.4(a) and (b) are obviously smoother than Figure 4.3(a) and (b). The two images Figure 4.4 (c) and Figure 4.4(d) are both much smoother and have more prominent targets than their counterparts in Figure 4.3.

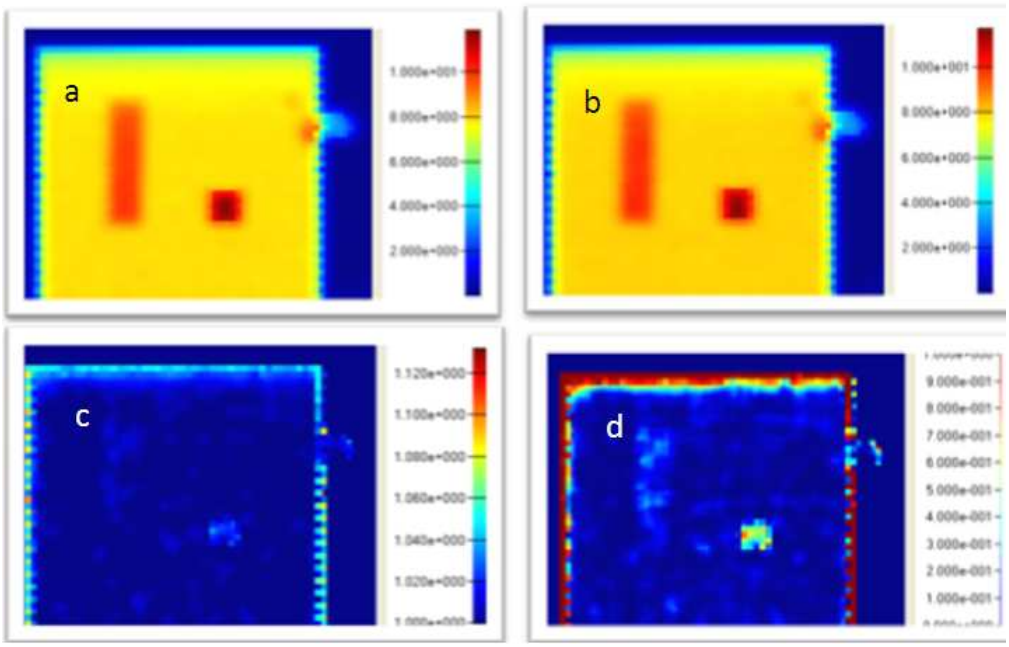


Figure 4.4 (a) The filtered high energy scan thickness map of a setup with 100cc W and 2”-thick Cu bar behind 8 inches of steel (b) the filtered low energy scan thickness map (c) the ratio map (d) the probability map

4.3.4 Gray Scale Morphological Filtering

Gray scale morphological filtering is discussed in Chapter 3. The two principal morphological operations are dilation and erosion. For binary images, dilation allows objects to expand, thus potentially filling in small holes and connecting disjoint objects. Erosion shrinks objects by etching away (eroding) their boundaries. These operations can be customized for clearing unwanted small objects in a binary image by choosing a proper shape and size of the structuring element.

For gray scale images, dilation can be seen as replacing a pixel value with its local maximum in the region under the structuring function, if the structuring function is a flat function. On the other hand, erosion works as a minimum function that replaces a pixel's value with its local minimum. Opening is erosion followed by dilation. As a result, it removes high gray-level regions smaller than the structuring function mask (bright artifacts) from an image. Conversely, closing is dilation followed by erosion. The consequence of minimizing the maximized image is that low gray-level regions (dark artifacts) in an image will be eliminated. Opening followed by closing can eliminate all dark and bright artifacts, which is called morphological smoothing. For pseudocolored images using the colormap shown in Figure 4.2, dilation reduces the area of blue pixels and increases the number of warm color (orange to red) pixels; erosion does the opposite of dilation; opening can smooth out the warm-colored artifacts, while closing smooths out cold color artifacts.

Grayscale morphological filtering can be very helpful in enhancing the signal in probability images. Figure 4.5 (a) shows a T map of a copper bar and a 100cc tungsten cube behind 9 inches of steel. The calibration material used here is copper. The image looks like a mosaic because the magnitude of the thickness causes great uncertainty in the measurement. One can also see that the upper region of the steel appears to be thinner than the lower region. This is due to the random scattering of photons that taints the upper detectors.

If the detector size is about half inch wide (1.25cm), then a 100-cc cube will only be about four pixels wide. Taking edge effect into consideration, there could be only four useful pixels left. The problem of low resolution can be reduced by upsampling and interpolation. To upsample is to insert zeros between samples. Interpolation is done by assigning the value of the nearest pixel to the inserted zero. Applying upsampling and interpolation increases the resolution and gives a greater degree of freedom in choosing image filter size.

Figure 4.5 (a) is the upsampled version of a 47 by 74 pixel image. The original image has a resolution 0.5 inch/ pixel. After upsampling by 4, a half-inch-wide region is represented by 4 pixels. We applied a 5x5 Wiener filter to the thickness maps to obtain the probability map shown in Figure 4.5 (b). The pixels belonging to the tungsten region are dominated by low values, except for the pixels in the lower right corner that has relatively high value. Also, it can be seen in the P map that there are red edges around the boundary of the steel plate, which indicates a false high probability in those regions caused by the edge effect. Since grayscale closing can be viewed as smoothed maximization, we can enhance the signal in the tungsten region by applying a closing filter. Figure 4.5 (c) shows a result of applying a 15 by 15 square closing operation. As a result, the signal in the whole tungsten region is enhanced; a bright, yellow square stands out from the background. The consequence of closing is the enhancement of edge artifacts. To eliminate those high value artifacts, we used an 11 x 11 opening filter. The result, shown in Figure 4.5 (d), has no false high probability edges any more. However, some of the edge pixels have value equal or greater to the target value. By looking at the P-maps such as Figure 4.5 (d), one still cannot tell if there are high Z materials located on the edge. The solution to this is adaptive masking which will be discussed later.

The size of the structuring element is determined by trial and error. The key point is that the size of the structuring element should not be as large as the smallest object we want to detect. Otherwise, the object can be eliminated by the erosion operation. Also, since morphological operations are not linear, the order of the

opening and closing filter does matter to the results. A reverse order of closing and opening is so called morphological smoothing, applying that to the P-map will not preserve the weak signals.

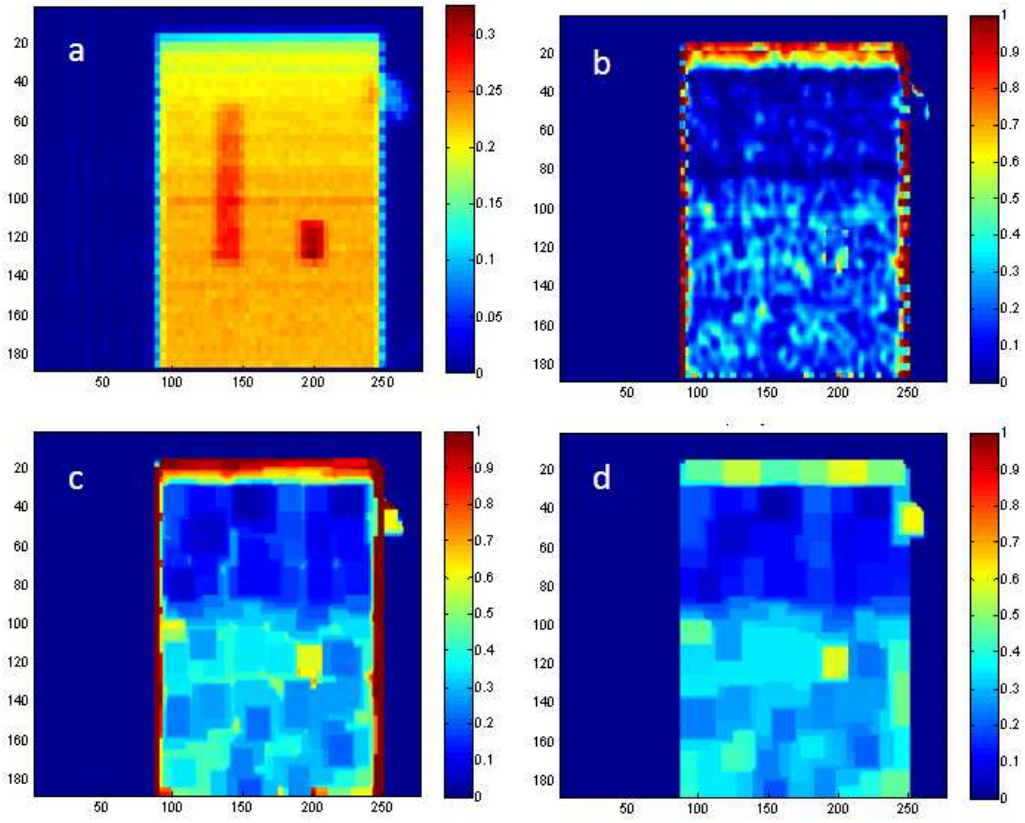


Figure 4.5 (a) T map of a copper bar and a tungsten cube behind 9 inches of steel. The image has not been filtered. The mosaic image is due to noise and low resolution. (b) P map of (a). The region of the copper bar has low probability. The shielding effect makes the tungsten region show only several pixels relative higher to its background. (c) P map of (a) after applying a closing filter with a 15 by 15 square structuring element to (b). The region of the copper bar has low probability. The shielding effect makes the tungsten region show only several pixels relative higher to its background. (d) Applying an opening filter to (c) with an 11x11 square structuring element. The red artifacts around the edges are eliminated after the operation. The tungsten region is now a yellow square in the image.

4.3.5 Adaptive Masking

Adaptive masking is the proposed method to identify high Z materials, especially shielded high Z objects. The heart of this approach is to eliminate false alarms caused by scattering or system-associated noise. The word adaptive means processing the image adaptively on a region by region basis. As mentioned in Section 4.3.2, the probability map can be segmented into high- and low- probability regions by using thresholding segmentation approach. If several objects of interest share the same background, it means that in a region, all targets are shielded by the same level of thickness. Applying the thresholding method to a region on the probability map will segment out the high probability objects. Even if that region is heavily shielded, if the penetration is acceptable, high Z material will still have relative higher probability than other objects in that region. A regional threshold can be determined by the automatic thresholding algorithm described in Section 3.5.1.

As we can see in Figure 4.3 and Figure 4.4, on the probability image there are pixels with high value outside the region of the real target. Even after morphological filters are applied, there are still some false high probability pixels left. Those pixels still will not be eliminated after automatic thresholding. To solve this problem, the idea of masking is introduced. A mask is a logical matrix (zero and ones) that indicates the locations of pixels belonging to a single object. The multiplication of the mask image and the segmented high probability image will make sure that all pixels falling outside the region of an object be eliminated.

Adaptive Masking Algorithm

The algorithm is based on several assumptions:

1. In a region, the area of the background is at least two times of the objects inside it.
2. When shielded by the same material, i.e. sharing the same background, the objects to segment will still have higher calculated probabilities than other objects.

3. Mid- or low-Z materials will have relatively low calculated probabilities that will be classified to background after segmentation.

The assumptions made above are very reasonable and in general, are applicable in actuality most of the time. If an object is approximately the same size as its shield, then that object combined with the shield is considered an isolated region. In that case, applying the threhsolding segmentation method to that region will not screen out an object. However, if the region itself has a probability greater than 0.6, it is reasonable to assume that whole region is a high Z target.

The adaptive masking segmentation comprises two parts: the first is segmentation of the high energy T-map. Thresholding is applied to each shielded region. The purpose of this is to segment the objects as "masks" for thresholding of the probability map.

The flow chart of T map segmentation is shown in Figure 4.11. First choose an initial threshold T_o to be the average of the maximum and the minimum of the grey level values of the whole image, $I(x,y)$. Then, divide the pixels in the image into two groups: $G_1(x,y)$ and $G_2(x,y)$ such that the former contains regions whose values are greater than the threshold and the latter is the rest of the areas. That is

$$G_1(x,y) = \begin{cases} 0, & \forall\, x,y \in I(x,y) < T \\ I(x,y), & \forall\, x,y \in I(x,y) \geq T \end{cases} \tag{4.5}$$

and

$$G_2(x,y) = \begin{cases} 0, & \forall\, x,y \in I(x,y) \geq T \\ I(x,y), & \forall\, x,y \in I(x,y) < T \end{cases} \tag{4.6}$$

Using the mean of the grey level values of each region, obtain a new threshold and use it to get the new $G_1(x,y)$ and $G_2(x,y)$ until the threshold converges. According to assumption 1, we are interested in looking for objects smaller than the background. Therefore, if the result of the segmentation (the region that $G_1(x,y) > 0$) gives an

object bigger than ½ of area of the background, we need to redo the whole process on $G_1(x, y)$ iteratively until the size reaches this criteria.

The loop ends whenever the new $G_1(x, y)$ is identical to the old $G_1(x, y)$ from the previous iteration. In that case, it means no more pixels can be segmented from that region. If the area of $G_1(x, y) > 0$ is greater than 60% of the background, it is either because there is no object inside that region, or the region itself is an object. In either case, we consider the whole region an object. After segmentation from the thickness map, label all the objects by assigning a number to each isolated, connected region. The result of T map segmentation can be expressed as

$$M(x, y) = \begin{cases} 1, & G_1(x, y) > 0 \\ 0, & G_1(x, y) = 0 \end{cases} \tag{4.7}$$

Figure 4.6 (b) shows the result of T map segmentation of a simple setup. Two objects, a copper bar and a 100cc tungsten cube, are shielded by 8 inches of steel as shown in Figure 4.6 (a). The whole shielding zone is considered a region. The two objects in that region are then segmented by the algorithm described above. Figure 4.6 (b) is a binary image $M(x, y)$ formed of logical 0 and 1 shown by blue and red respectively.

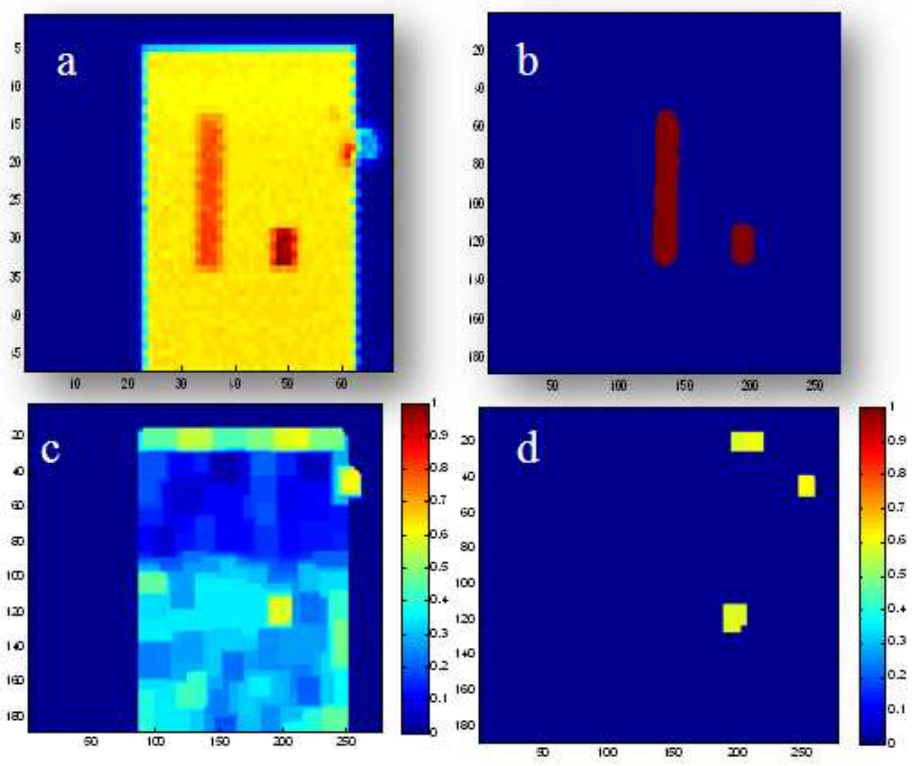


Figure 4.6 (a) Thickness map (b) output of segmentation (c) The filtered P map (d) thresholding segmentation of (c).

The second part of the adaptive masking algorithm is thresholding segmentation on the P map. By applying the same thresholding method as the *T* map, it is expected that the high probability areas in each region will be segmented. Figure 4.6 (c) and (d) show the result of P map thresholding segmentation. However, as shown in Figure 4.6, there may still be many spurious high probability pixels after thresholding segmentation.

Also, the automatic thresholding algorithm will fail to work when the probability map is too noisy. The reason for that is that the contrast between the objects and the background is low. The threshold might converge very fast, classifying a lot of background pixels as object pixels. Figure 4.7(b) is the probability map of a setup with a 100 cc tungsten cube put behind 10 inches of steels. The image is processed by a close followed by an opening morphological filter. Figure 4.8 is the

result of using the automatic thresholding algorithm to separate pixels into two groups that $G_1(x,y) \geq T$ and $G_2(x,y) < T$ where T is the calculated threshold. Apparently, a lot of background pixels are classified into foreground pixels because of the low contrast between the probability of the background pixels and the probability of target pixels.

To fix the problem of erroneous classification caused by fast convergence, we modified the automatic threshold finding algorithm to fit our objective. That is, we will keep doing the threshold finding on the probability image until the output is meaningful; otherwise, we will determine that nothing can be found in the probability map.

As shown in figure 4.6 (d), the result of segmentation can still give spurious pixels, the two yellow blobs on the top in figure 4.6(d) are pixels belonging to the edge of the shield. To eliminate those false regions, we apply a mask *M* to filter out the unwanted regions.

The masking of the thresholded *P* map by mask *M* can be expressed as

$$P_{masked}(x,y) = \begin{cases} 0, & if\ \ M(x,y) = 0 \\ P_{th}(x,y), & if\ \ M(x,y) = 1 \end{cases} \tag{4.8}$$

or simply,

$$P_{masked}(x,y) = P_{th}(x,y)\, M(x,y) \tag{4.9}$$

where $M(x,y)$ is binary image which is the result of *T* map segmentation expressed by Equation (4.4); $P_{th}(x,y)$ is the segmented *P* map resulting from automatic thresholding. One example of the result of masking is shown in Figure 4.9. The segmented *P* map in Figure 4.6 (d) is masked by the 2 regions in Figure 4.6 (b). As a result, the two yellow clusters on the top are gone because there is no corresponding object in M. The flow chart of *P*-map segmentation is shown in Figure 4.12.

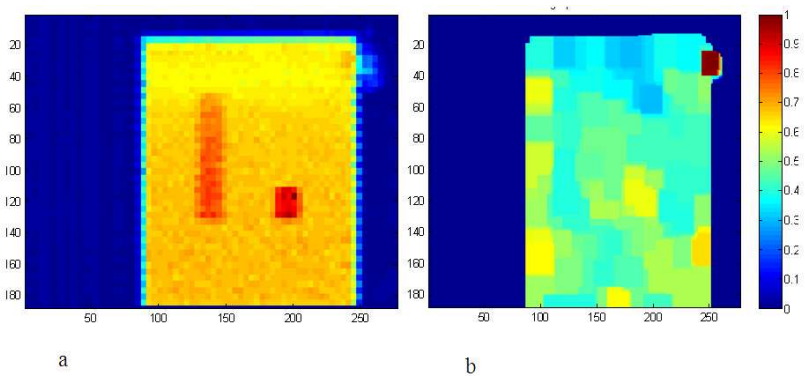


Figure 4.7 (a) The thickness map of a copper bar and a tungsten cube shielded by 10" of steel (b) the probability map of (a)

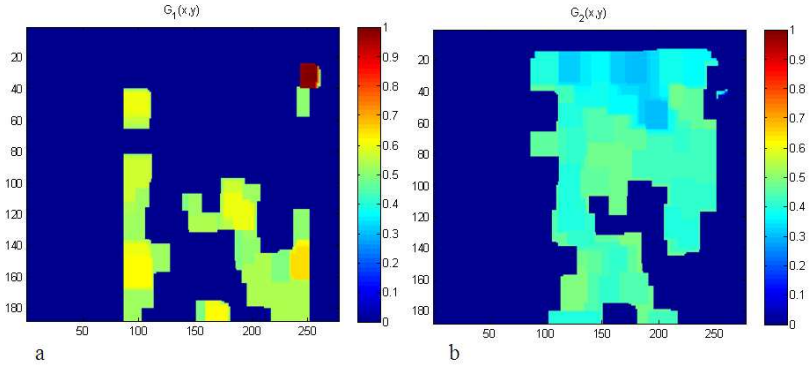


Figure 4.8 $G_1(x, y)$ and $G_2(x, y)$ after applying the automatic thresholding algorithm. There are a significant number of background pixels having values closer to the target pixel value than other background pixels. Therefore, the automatic threshold finding algorithm will converge to a value midway between the minimum of G_1 and maximum of G_2.

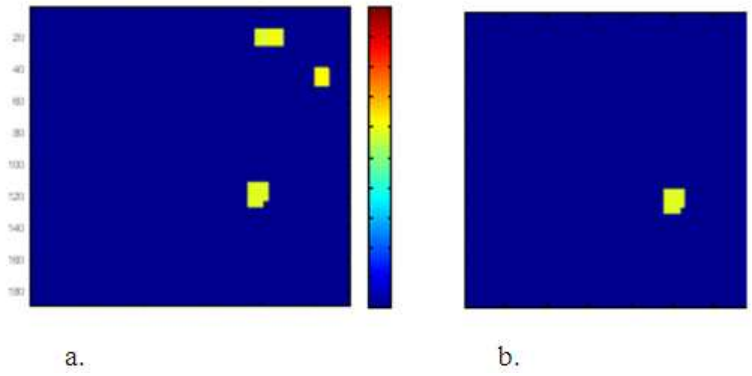


Figure 4.9 (a) The thresholding segmented P map as shown in Figure 4.6 (d). (b) The result of masking. Only the region that corresponds to a red region in Figure 4.6 (b) is preserved.

The problem shown in Figure 4.8 can also be dealt with by employing the "mask" from the segmentation of the T map. Comparing image $P_{th}(x, y)$ with $M(x, y)$, if the outcome of segmentation of the P map dose not significantly overlap with any subregion in $M(x, y)$, then we can say that no high atomic number material is detected. If an object in the segmented T map falls within a segmented high probability region, and the region of segmented P is larger than the region of an object, it is possibly the result of something like Figure 4.8. In that case, keep applying the thresholding algorithm to segment the new $G_1(x, y)$ of Equation (4.2) until either a segmented image that has a region highly overlapped with an object is found and its size is very close to the size of that object, or the loop ends as a result of no high probability object in that region.

Figure 4.10 gives an example to illustrate how the multistage threshold finding algorithm works. The result of P map segmentation shown in figure 4.8 (a) is put back into the thresholding algorithm again for another segmentation. The segmentation result G1 and G2 is shown in the second stage of Figure 4.10. The G1 of the 2nd stage still has area greater than the area of the tungsten, and still covers the region of the tungsten. We therefore still need to put the G1 of the 2nd stage for further segmentation. If we do 4th stage segmentation, the resulting G1 will have no overlap with the mask. So, the loop will stop at the 3rd stage. The output of the algorithm therefore will be the overlapped region of G1 of the 3rd stage masked by the 2 regions of mask M respectively.

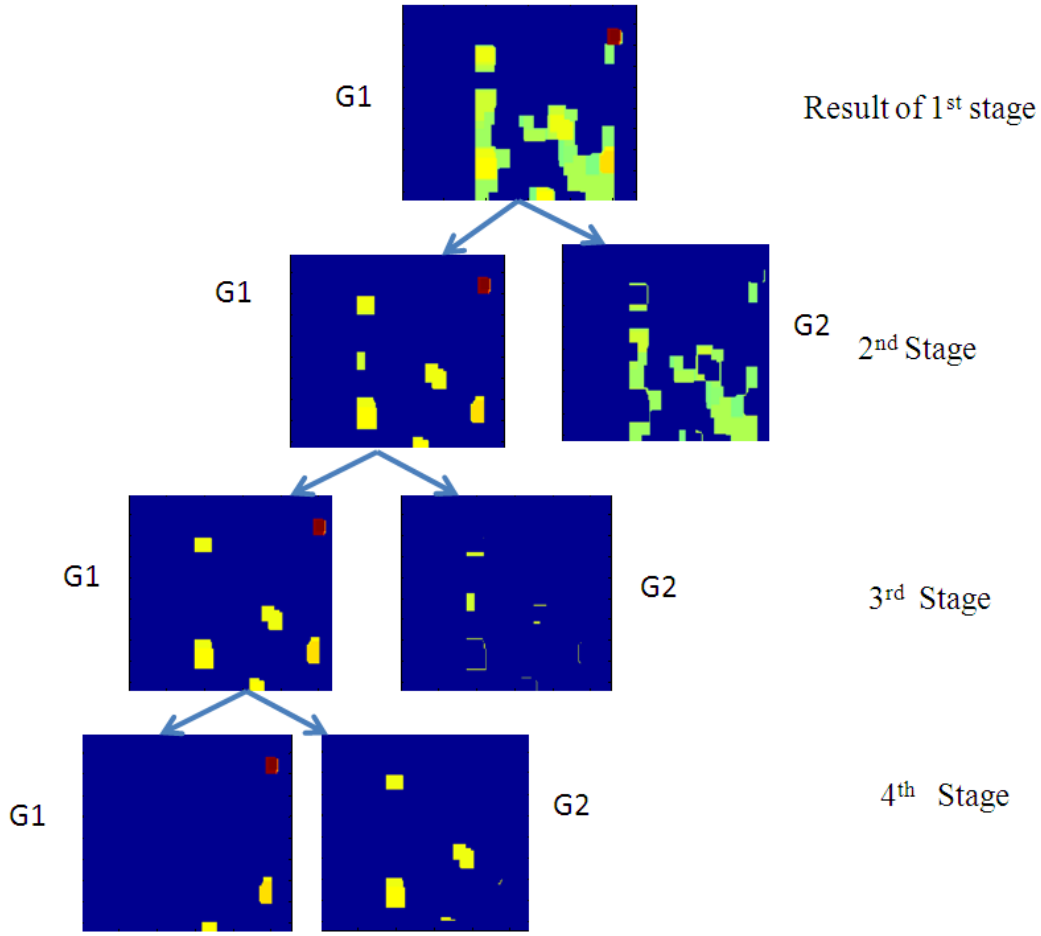


Figure 4.10 Results of each stage of thresholding.

There are two advantages to doing the adaptive masking thresholding: one is that we don't need to select a fixed value for high Z threshold; the second is that we are able to eliminate false alarms. The same object can have different probabilities on a P-map if the thickness of shielding material varies from region to region. Adaptive thresholding can automatically determine the high Z threshold in a region of a given background thickness. A chunk of false alarm pixels can be removed by adaptive masking as long as no corresponding object with comparable size is segmented from T-map.

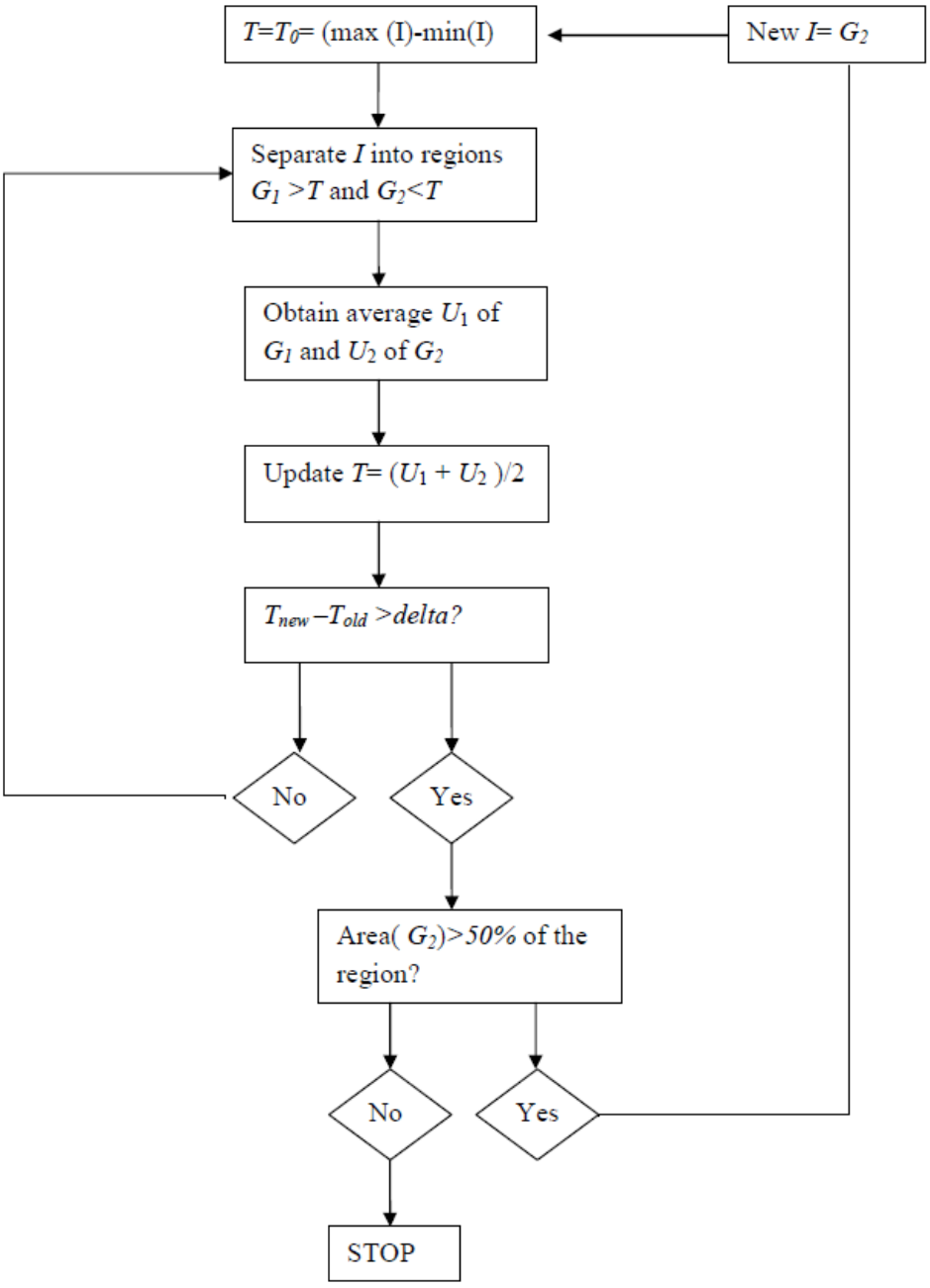


Figure 4.11 Flow Chart of T map thresholding algorithm

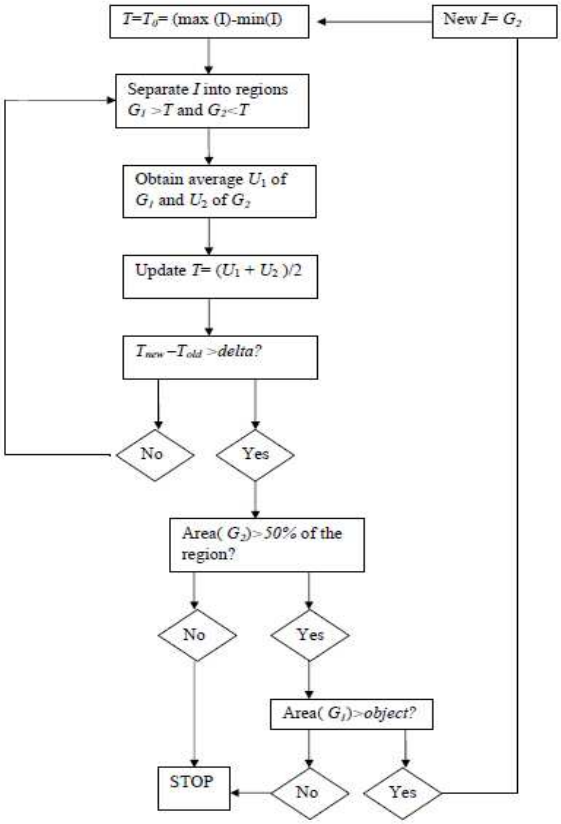


Figure 4.12 Flow Chart of P map adaptive masking algorithm.

4.4 Experiments and Simulations

For testing our method, we set up a simple dual energy test bed system. We used the Linac X-ray system that generates photons at 9MeV and 6MeV maximum energy levels. The maximum penetration of the X-rays is the ability to penetrate up to 16 inches of steel. Our testbed setup has 47 detectors that create 47x70 pixel images. Each pixel represents a half-inch wide region. The moving X-ray source produces 440 pulses per second and scans objects at a speed of 33 inches per second. Each detector is calibrated using copper as the calibration material. No detector collimation is done. Therefore, random scattering can taint the thickness measurement.

Several materials are used for experiments: tungsten, lead, copper and tin. Copper is a low Z material, and tin, which has atomic number 52, is usually a source

of false alarms. Tests of different configurations from simple to more complex are performed: (1) One high Z and one non-high Z object behind the same shield. The ideal is one region has zero probability while the target region has values that stand out from the background (2) Several high Z and low Z objects behind the same shield. The ideal is all high Z detected without any miss or false alarm. (3) Objects behind different shields. Shields can be overlapped with each other or placed at distant locations. A minimum high Z ratio of 1.02 is applied to all pixels for calculating the probability.

For testing our methods, the image processing algorithm is coded in the Matlab 7.6 environment. Matlab is a very powerful language for scientific calculation. The image processing toolbox provides image processing filters such as median filter and the Wiener filter. The image processing toolbox also provides functions for morphological filtering; functions like dilation, erosion, opening and closing are all implemented and all kinds of structuring functions are provided. The downside of Matlab is its processing speed. The real imaging system will be coded in other lower level languages such as C++.

4.5 Results and Discussion

Figure 4.13 (a) shows the high energy scan thickness map of a setup of three separate shields made of 10, 8, and 7 inches of steel from left to right respectively. Each of them has a cube of 75cc of tungsten (W), 75cc of tungsten, and 100cc of lead (Pb) placed at the right of the shield and a copper (Cu) bar on the left respectively. The first region on the left looks noisy because it is the thickest region in the image. Figure 4.13 (b) is the P-map of (a); it's a result of upsampling the T-maps by a factor of 4, then applying a 5x5 Wiener filter, followed by a grayscale morphological closing(11x11) and opening(7x7). As expected, the P-map of the first region is the noisiest one; a lot of background pixels have values equal or greater than the foreground region. Figure 4.13 (c) is the result of adaptive masking. P-map thresholding segmentation is applied to the three regions one at a time. The output of

thresholding is then masked by the regions obtained from T-map segmentation. By employing region thresholding followed by masking, all targets are successfully identified. Segmentation is achieved without using one global threshold for the whole grid.

Figure 4.14 is the result of a more complicated setup. A one-inch steel plate is placed with some overlap with seven other plates of the same kind. The overlapped region has a target behind it. Two free air objects are also put in the scene. The idea here is to test if our algorithm can successfully identify all high Z materials indicated by the red arrows. The tin cubes (pointed to by the orange arrows) are used here to test the discrimination ability of the model. The result of the algorithm is shown in the bottom Z image—a binary image that indicates the pixels classified as high Z materials. In this case, all targets are correctly detected without any miss or false alarm. Tin has an atomic number that produces a ratio higher than common metals. Statistically, when a tin object and a high Z object are both heavily shielded, the difference between their ratios can be as small as the statistical error. It is possible that in some cases, tin objects will appear to be high probability on the *P*-map. Discriminating high Z materials from such metal is one of the major challenges.

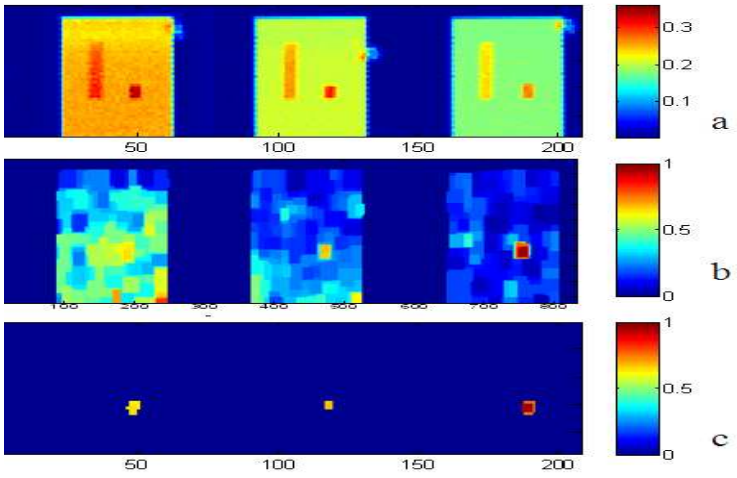


Figure 4.13 (a) The thickness map of 75ccW behind 10” of steel; 75ccW behind 8” of steel, and 100ccPb behind 7” of steel, 100ccPb (b) the image processed P map (c) The regions on P map being detected as high Z materials.

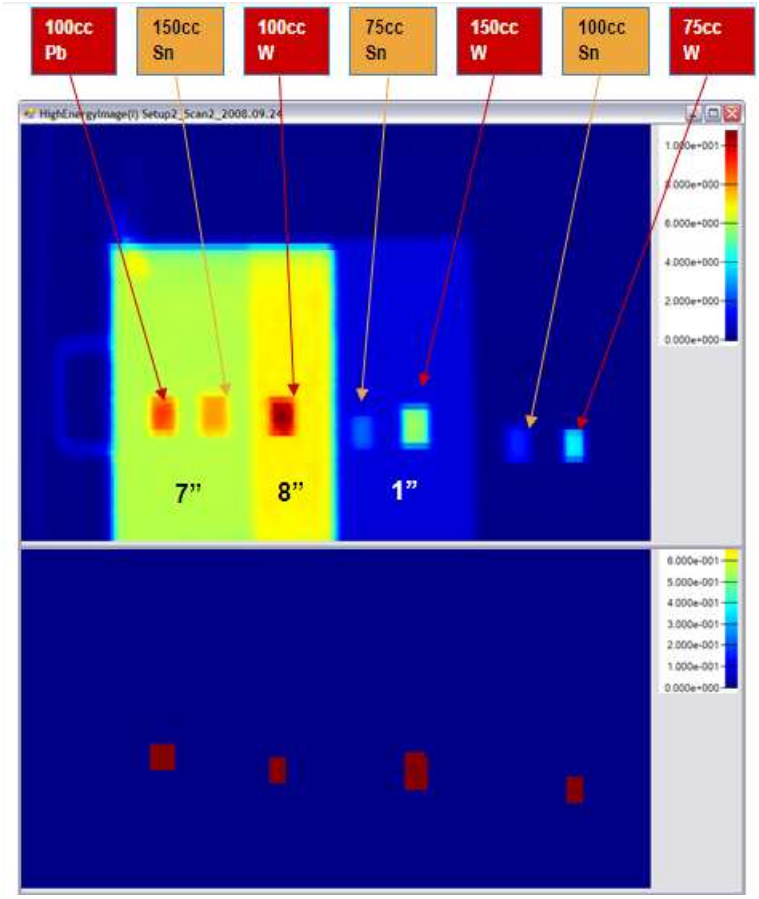


Figure 4.14 Top: the setup of 1 inch steel plate overlapped with 7” of steel plates. A tungsten cube is placed behind the overlapped region. All objects behind the shield are indicated by the arrows. The red arrows point to the target material, the orange arrows point to tin. Two objects are there without shielding. Bottom: The Z image indicating the targets being identified.

Table 4.4 summarizes the performance of the system. In total 300 different combinations of thickness and targets, testing on lead, tungsten, and tin, we find the rate of miss and false positive are less than 1%. A 100cc lead cube was only missed one time when it was shielded by 9 inches of steel. Considering the low resolution, low penetration and noise contamination the system has, the performance of our algorithm is very encouraging. It shows that our method is, in fact, an effective way for high Z object discrimination.

Table 4.4 Summary of test results

Thickness / Object	5”	7”	8”	9”	10”	11”
75cc	ALL PASS	ALL PASS	ALL PASS	No data available	No data available	ALL PASS
100cc	ALL PASS	ALL PASS	ALL PASS	Pb missed once	ALL PASS	ALL PASS
150cc	ALL PASS	ALL PASS	ALL PASS	ALL PASS	ALL PASS	ALL PASS
200cc	ALL PASS	ALL PASS	ALL PASS	ALL PASS	ALL PASS	ALL PASS

4.6 Conclusion

In this chapter, we introduced a procedure shown in Figure 4.15(b) that can be used in dual energy X-ray imaging for detecting high Z materials. Preliminary testing shows that the technology has good potential for finding targets with high atomic numbers. Upcoming experiments will allow examination of more complex scenarios. In contrast to the method shown in Figure 4.15 (a), although the specific effective atomic number of an object cannot be indicated in the output image, suspicious regions that contain high Z materials can be detected using our method. Also, false alarms can be effectively suppressed in a way never reported before.

Denoising of X-ray images can help system operators identify objects more efficiently in enhanced images. When the resolution is low and the signal is weak, upsampling the image can give a greater degree of freedom in choosing filter size. The downside of upsampling will be slower processing speed since the data will be N^2 bigger when we use an upsampling factor N. If the display resolution is improved and/or signal is strong, upsampling is not needed.

Our adaptive masking approach for high Z material discrimination is not only limited to the application of cargo imaging systems, it could also work very well for airport luggage inspection system. In addition to thresholding, other segmentation techniques such as region growing can also be used for finding masks. Because the background of a cargo container image can be very complex, the adaptive thresholding masking technique still requires system operators to manually select region of interest

to separate object from its background in the selected regions. In this case, a system operator will visually determine a region with similar background and use a mouse to select a region for further high Z analysis. A fully automated high Z detection algorithm based on the model in this chapter may require having a shape identifier algorithm that can classify object based on its geometric information.

In later chapters, we will discuss other approaches that require less human involvement in high Z detection. To more accurately define the masks in a cargo image, other segmentation methods are used instead of thresholding. However, the concept of masking remains the same.

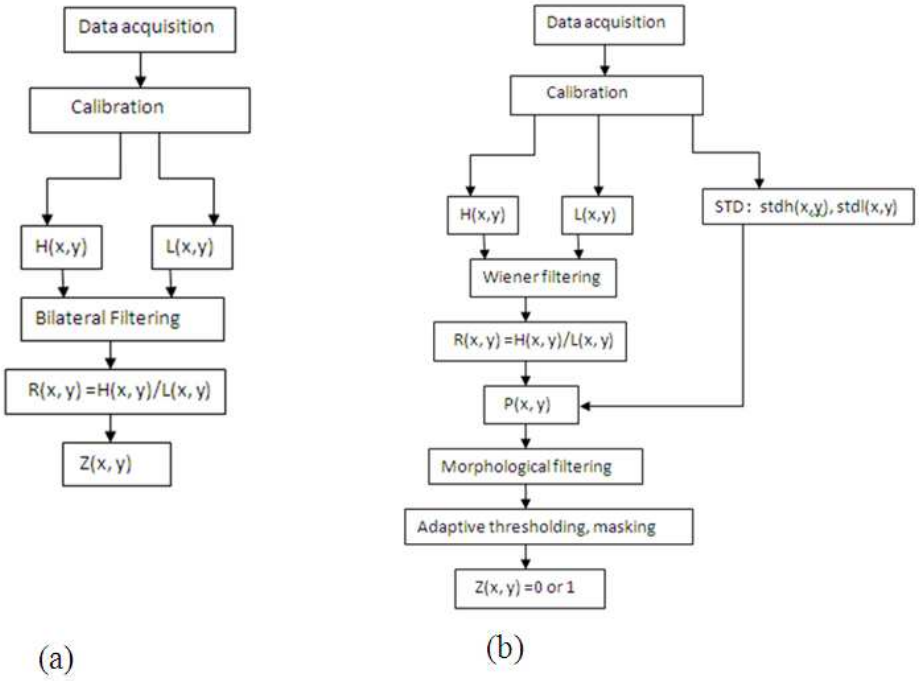


Figure 4.15 (a) the flow chart of a general material discrimination system. (b) The modified high Z detection algorithm.

Chapter 4, in part, is a reprint of the material as it appears in the Proceedings of SPIE (2009), Fu, Kenneth; Ranta, Dale; Guest, Clark; Das, Pankaj "A novel algorithm for material discrimination using a dual energy imaging system." The author of this dissertation is the primary investigator and author of this paper.

5

Segmentation of Objects in Cargo X-ray Images

5.1 Introduction

We have discussed in Chapter 4 using the locations of objects as masks to check if the detected high Z pixels fully or at least mostly cover an object. The method for creating such masks is T map adaptive thresholding. The thresholding is fundamentally fast and straightforward. However, for noisy images or more complex scenarios, thresholding segmentation might introduce spurious regions. Therefore, the segmentation method discussed in Chapter 4 might work better for airport luggage X-ray image scenarios that are not as complicated.

In this chapter, we introduce another segmentation method to screen out the objects of interest. Our new approach for segmentation is a combination of clustering and a seeded region growing method. This segmentation method has proved to be very efficient. To detect hidden objects, we use a region filling technique to find regions whose intensities are greater than that of their surrounding regions. Doing further regional analysis to those regions can extract atomic information about them.

5.2 Segmentation Based on Intensity

Different applications should use different methods of image segmentation. For example, if our purpose is to find a gun in an X-ray luggage image, simple methods such as the thresholding method discussed in Chapters 3 and 4 will definitely not work. A more appropriate way is to develop a pattern recognition system that identifies certain shapes in an image. For simple goals such as separating whatever the object is from the background, we probably will not use pattern recognition techniques since we have no prior knowledge what we are segmenting and what geometric information we need. In that case, segmentation by thresholding might be a better choice if the pixels of the object we are looking for form a somewhat even distribution of intensity. The other method we can use is clustering. Clustering partitions an image into several subimages based on the similarity of selected features. Two clustering methods are often used, which will be briefly discussed in the following sections.

5.2.1 K-means clustering

As discussed in Chapter 3, k-means clustering partitions data into k mutually exclusive clusters, and returns the index of the cluster to which it has assigned each sample or observation.

A sample of the data is treated as an object having a location in space. Or in mathematics, each sample is treated as a vector in vector space. Each element in that vector space is an *N* dimensional vector and each dimension represents a feature of the data. For example, a vector from census data could be a three dimensional vector of (‘gender’, ‘age’, ‘income’). For image data, a vector space could be (x, y, intensity) or simply a one dimensional vector (scalar) of intensity.

The distance between two samples *i* and *j* in an *N* dimensional space is then defined as the distance of two vectors in vector space:

$$D_{ij} = \sqrt{(x_{1i-}x_{1j})^2 + \left(x_{2i-}x_{2j}\right)^2 + (x_{3i-}x_{3j})^2 + \cdots + (x_{Ni-}x_{Nj})^2} \qquad (5.1)$$

where D_{ij} is the distance between two samples and denotes the projection of the *q*th dimension of vector i . The similarity between two samples is then defined by the distance between the two samples. K-means finds a partition in which objects within each cluster are as close to each other as possible, and as far from objects in other clusters as possible.

Each cluster in the partition is defined by its member objects and by its centroid, or center. The centroid for each cluster is the point to which the sum of distances from all objects in that cluster is minimized. For that purpose, it uses an iterative algorithm that minimizes the sum of distances from each object to its cluster centroid, over all clusters. The location of the centroid might be updated as the k-means algorithm moves objects between clusters until the sum cannot be decreased further. The result is a set of clusters that are as compact and well-separated as possible.

The hardest problem for this algorithm will be the selection of initial locations for centroids and the number of centroids, *k.* Especially for the purpose of image segmentation, if no prior knowledge is available, we will not have a good result without several experiments. The selection of different initial points can generate different results.

5.2.2 Leader Clustering

The *leader* cluster algorithm partitions a data set into groups by virtue of a radius distance *T* in feature vector space. A leader object is associated with each group and all other objects in the group lie within distance *T* from that object. Figure 5.1 illustrates how the leader clustering algorithm works. The algorithm starts by selecting the first data point and assigning it as the first leader object, A. Subsequently, the remaining samples are examined and those that are within the distance T are assigned to group one. The first data sample examined that falls outside the radius T is assigned

as the next leader object, B. This procedure is iterated to identify cluster centre C as well as the remaining centroids. This algorithm has the advantage of being fast since it only requires going through the data once. All cluster centroids are at least a distance T from each other, which means that it is dependent on the ordering of the data set and that a distance T is specified rather than the number of clusters.

Like K-means, the result of the leader algorithm is also affected by the selection of the initial point. In Figure 5.1, there is one red point in cluster A that is also within the radius of cluster B. However, since A is the first leader, that red point is classified to A rather than B. However, if B is the first leader, that red point would end up in cluster B.

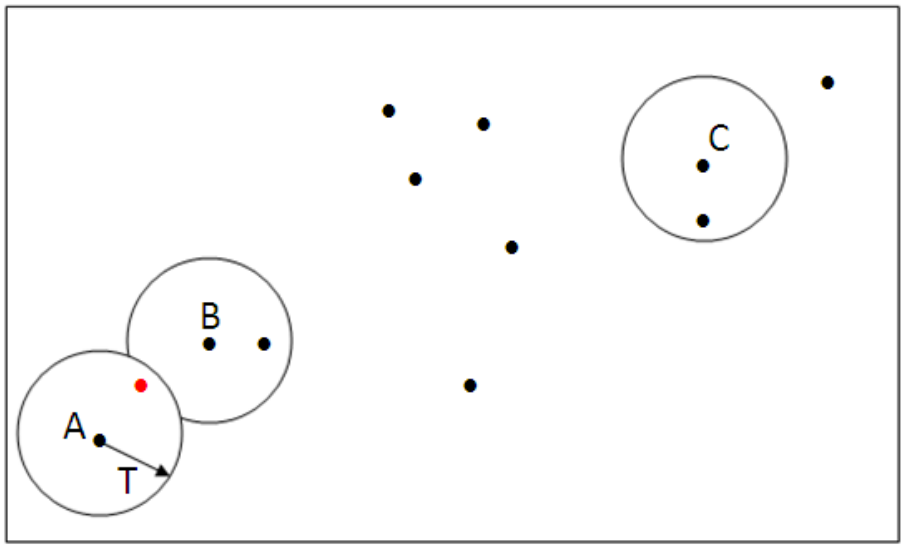


Figure 5.1 Leader Clustering. Each point in the data is a vector in a vector space. Note that there is a red point in cluster A that is also within the radius from leader B.

For clustering a grayscale image, a three dimensional vector space of (x coordinate, y-coordinate, intensity) is used. That means we will group two pixels together if they are close to each other and they have similar intensities. In that case, for a big object with its pixels distributed over a wide range in the image, the pixels of that object could end up in different clusters. Also, there might be a case where a few pixels belonging to a cluster lie within another cluster. On the other hand, if the

constraint of distance is lifted, two distant objects with same intensity will end up being classified into same group. Then the result is of no use if our purpose is to label all isolated regions.

5.2.3 Intensity Grouping: a Hybrid of Leader Clustering and Region Growing

Intensity grouping is an image segmentation approach we developed for the purpose of X-ray image segmentation. However, the application of this segmentation is not limited to just X-ray images. It could also be applied to any images. The clustering method we used is a combination of clustering algorithms and region growing. It has the advantage of leader clustering: fast, straightforward, and only one scan of data is needed. Unlike the leader or k-means clustering methods, our approach does not require a predefined number of clusters or a radius to define the size of a cluster. Also, there is no need to select the initial centroids. As a result, all connected regions in an image will be segmented regardless of their size. If the result of our algorithm ends up with too many tiny regions, we will apply region merging to that result.

The algorithm for pixel classification

The idea of intensity grouping is to classify neighboring pixels with similar intensities together. For a pixel (x, y), if its value $f(x, y)$ is close to the mean value of its neighbor pixels, that pixel is then grouped together with its neighbors. The number of neighbors is at least 8. A greater number can also be used depending on the property of the image we are dealing with.

A step size Δ is selected, each pixel in the image will be assigned to a value quantized by the step size. The total number of levels is determined by the step size. For example, if an equivalent thickness image has a maximum value of 12", there will be totally 6 levels in the image as a result of clustering: pixels thicker than 1" but thinner than 3" will be assigned to level 1; pixels with thicknesses between 3" and 5" will be assigned to level 2, and so on. If a pixel's value is close to the average of its

neighbors, that means that pixel and all or at least most of its neighbors belong to the same region. We then assign the quantized value of its neighbor's mean to that pixel. If the value of a pixel is not similar to its neighbors, that means that pixel is either on the edge (if we check only the 8 neighbors) or its neighbors belong to at least two different regions. In that case, we will use that pixel as an origin, divide its neighbors into four quadrants, and check the average values of each quadrant, and assign the closed quantized quadrant mean to that pixel. This way, the clustering is much faster than the recursive K-means algorithm, and leader clustering methods that requires finding new leaders all the time. More importantly, our method will create continuous regions but not the broken clusters resulting from k-means and leader algorithm.

In Matlab, our method is done using filters. Five filters are defined: one mean filter Q, and four quadrant mean filters Q_1 to Q_4. A radius r is selected to define the size of the filters. The size of the filter is then $2r+1$ by $2r+1$. If r is equal to 1, that means we are looking at the regional average of 8 adjacent neighbors. Figure 5.2 shows five mean filters with $r=1$. The input image is filtered by those five filters. The output $G(x, y)$ of mean filter Q will be the mean of the masked region centered at (x, y). The output of the rest four filters G_1 to G_4 are the mean of each quadrant anchored at (x, y).

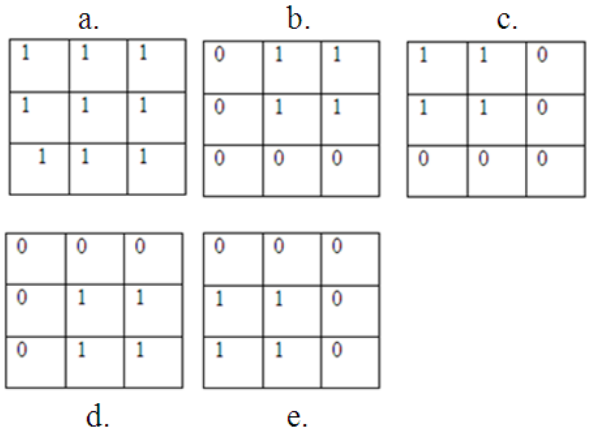


Figure 5.2 The 5 filters used for clustering(a) Q calculating the mean of the masked region (b) Q_1, calculating the mean of the first quadrant centered at the pixel(c) Q_2, calculating the mean of the second quadrant (d) and (e) are Q_3 and Q_4 respectively.

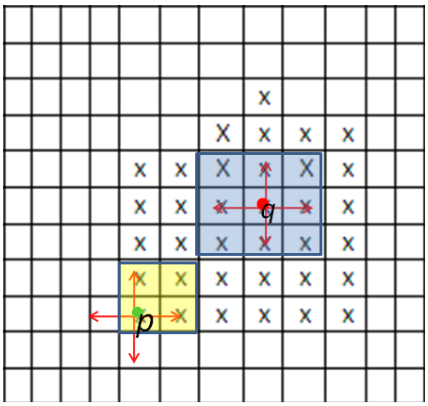


Figure 5.3 Two points *p* and *q* in a binary image. Point *q* lies within a region, its value will be equal to the average of the masked region (blue area). Point *p* is placed on the edge. Its value is equal to the average of the first quadrant (yellow area).

Figure 5.3 illustrates how the clustering works using a binary image for example. Two points p and q are in a binary image. The former is on the edge of a region and the latter is within a region. When a 3 by 3 mask is used on point *p*, we see that the value of *p* is quite different from the mean of the masked region. We then check the mean of the 4 quadrants. The value of the outputs of the 5 filters shown in Figure 5.3 that is closest to the value of *p* will be the output of Q_1, which is the mean of the yellow area. Therefore, the quantized value of the mean value of yellow area is

assigned to point *p*. For point *q*, since it is located within a region, its values will be the same as the average of its neighbors. Also, the five filters will give the same result. We therefore, assign the quantized mean value to that pixel.

Most of the points outside the foreground region will have a value of zero since the outputs of the 5 filters are all zeros. For a background pixel next to *p*, its value is equal to the output of Q_2 and Q_3, which is still zero. Even if it is close to the edge of a foreground, it won't be classified to be an edge pixel. We can then expand this concept to grayscale images. Assuming the foreground and background in Figure 5.3 are both homogeneous regions. The output of clustering will be two regions with a quantized value.

Labeling the region

After classifying all the pixels in the image, the whole image is reduced to several quantized values. For each level, we find the number of regions with pixel value equivalent to that quantization level, and, then, assign a number to each connected region. The total number of regions of that image will be the summation of all regions of all levels. The Matlab function "bwlabel" is used for our region labeling for each level. The function "bwlabel" is used for finding the number of regions in a binary image and assigning a number to each region. We treat all pixels in the clustered image with value equivalent to a quantization level as a binary image and find all the regions within that level. The total number of regions can be obtained after all levels are checked.

Region Merging

The result of our clustering algorithm is affected by the selection of step size. A small step size will create too many unwanted tiny regions, which is not desirable for doing regional analysis. A step size too big can cause two regions belonging to different objects to merge into one region. One solution for that problem will be region merging.

The idea of region merging is to combine regions with similar value into one region. In our region merging algorithm, the largest region is used as the "seed"

region. We check the regions surrounding the seed region: if the difference between the averaged values of the two regions is less than a predefined criterion, we then merge those two regions together. Each time, we grow a region by dilating the region with a 3 by 3 square structuring element. The regions that are overlapped with the dilated region are the surrounding regions of that dilated region. We check all surrounding regions to see if any of them could be merged. After all neighbor regions are checked and merged, a new region is formed. This region continues growing by the same way until no more regions to be merged can be found. We then select the next seed, which is the second biggest value on the cluster map, and start the whole merging routine again. The merging will stop after no more seeds can be found. Region merging is not always necessary for some simple regional analysis. However, it will give a more precise description of the distributions of objects in an image.

5.3 Object Segmentation Using Region Filling

One way to segment an object from its background is region filling. Binary region filling is an algorithm that starts from a point p within a boundary, then replaces all values within that boundary with 1s. Figure 5.4 is an example of binary region filling. Figure 5.4 (a) shows an image composed of white circles with black inner spots. Any black spot can be used as a seed point, the seed will grow to identify all black pixels and then replace them with 1's. Figure 5.4 (b) is the result of filling all circles.

The morphological algorithm for region filling is based on a set of dilations, complementation, and intersections. The filling result X of an image A can be expressed as

$$X_k = (X_{k-1} \oplus B) \cap A^C \qquad k = 1,2,3, \ldots. \tag{5.2}$$

where $X_0 = p$, $X_{k-1} \oplus B$, is a dilation by a symmetric structuring element B, A^C is the complement of the image. The algorithm terminates at the iteration k when $X_k = X_{k-1}$. The whole region inside the boundary definitely belongs to A^C. Starting from a point p, the point is expanded through the dilation operation. The intersection

with A^C of the dilated region therefore will be the region of the dilated region within the boundary. The region continues dilating until the region being dilated is equal to the region inside the boundary.

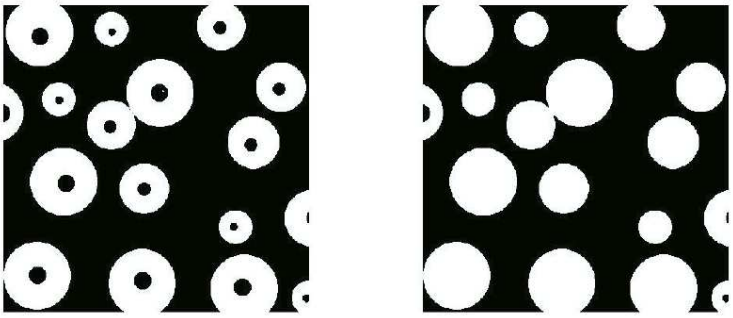


Figure 5.4 (a) The image with holes (b) after region filling, holes are filled (image obtained from Gonzalez)

The biggest problem of this morphological algorithm is that we have to manually select a starting point. For an automated algorithm to work, the computer needs to be able to determine background points from inner sphere points.

Another method for region filling is using the flood fill algorithm on the background. Flood fill, also called seed fill, is an algorithm that determines the area connected to a given node in a multi-dimensional array. When it is applied to the background region in an image, for example, the dark region outside the white circles in Figure 5.3, it will replace all the ones in background with zero. The result of background flood fill will only leave the black dots inside the circles. The remaining dark spots are then complemented (replacing them with 1s) and then added to the original image. This is the method used in Matlab for region filling. This method works very well when we are dealing with a totally enclosed foreground. When any of the white circles in Figure 5.4 have a leak to its inner black spot, that black region will be defined as background by the flood fill algorithm.

Here, we propose a new method for automated region filling. This method is based on finding a pair of indicators for the starting and end point of an interior region.

Gradient of a two-variable function is defined as

$$\nabla f(x,y) = \frac{\partial f}{\partial \mathrm{x}} + \frac{\partial f}{\partial \mathrm{y}} \tag{5.3}$$

As shown in Chapter 3, the gradient in a digital image is

$$\nabla f(x) = \frac{\partial f}{\partial \mathrm{x}} = f(x+1) - f(x) \tag{5.4}$$

in the x-direction and

$$\nabla f(y) = \frac{\partial f}{\partial \mathrm{y}} = f(y+1) - f(y) \tag{5.5}$$

in the y - direction.

In a binary image, if a pixel is inside a region, its value will be the same as its neighbors, we will have $\nabla f(x,y) = 0$. If we scan along the x-axis, when a background pixel is next to the boundary of the foreground, we will have $f(x,y) = 0$ and $f(x+1,y)$=1. There for the gradient in that direction $G_x = f(x+1,y) - f(x,y)$ will be $G_x = 1 - 0 = 1$. Similarly, for a foreground boundary pixel, since its next pixel will be a background pixel with $f(x+1,y) = 0$, the gradient G_x will be $G_x = 0 - 1 = -1$. Therefore, when scanning along the x-direction, if we are going into a foreground region, we will encounter an increase in pixel value. That is, we will encounter a gradient =1. When leaving the foreground region, we will encounter a $G_x = -1$. Equivalently, it is the same for scanning in y direction.

Figure 5.5 shows an object in a binary image. The object has a hole inside its boundary. Two points having $G_x = 1$ are pointed by arrows. The first one is located on the boundary when its next pixel in x-direction enters the boundary. The second location is located inside the hole region. It is on the boundary when the next pixel becomes the foreground pixel.

On the same x-level (horizontal), two location having $G_x = -1$ are also indicated. The first one is located on the inner boundary of the foreground. The second is on the outer boundary of the foreground. We can see that the hole starts from the pixels next to the pixels having $G_x = -1$ and ends at the pixel with $G_x = 1$. Also,

from Figure 5.5, we can see that if a hole region exists, when scanning along x-direction, we will expect to encounter a $G_x = 1$ followed by a $G_x = -1$, followed by another pair of $G_x = 1$ and $G_x = -1$. Between those points, we will have $G_x = 0$. By replacing the zeros between the start and end points shown in Figure 5.5, we are filling the hole.

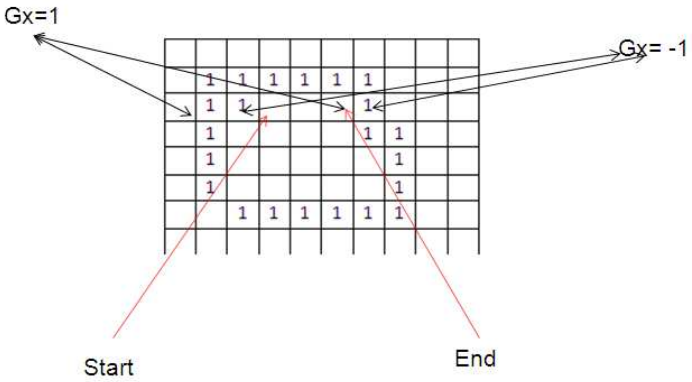


Figure 5.5 An object with a hole inside its boundary. The two locations of Gx=1 are indicated, one is located outside the object, one is on the boundary of the hole. The two locations of Gx=-1 is also pointed, the first one is located in the boundary of the hole, the second one is located on the boundary.

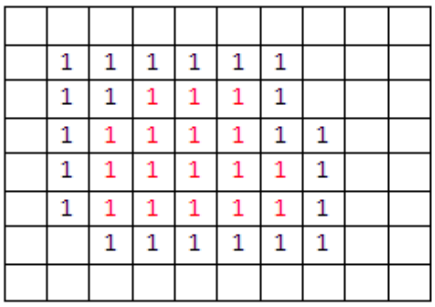


Figure 5.6 The result of region filling. The pixels between start and end points in Figure 5.5 are replaced with 1s.

Unlike the seed growing methods for region filling, our method is straightforward. More importantly, it avoids the problem of defining the interior

region. We simply take the gradient of the whole image, find two pairs of (1, -1) and fill ones to the pixels between the end point and start point of the two pairs.

If all regions that need to be filled are fully located inside an enclosed boundary, then, it does not matter, we can do the gradient scanning and filling in x- or y- direction. They will give us the same result. To guarantee the most accurate result, we apply this procedure to all regions respectively. In that case, the pair finding result will not be affected by a few artifacts.

Filling Gaps

Another advantage of using our methods over the interior seed growing or background flood fill method is that our method allows us to fill in not only enclosed holes, but also regions not fully enclosed. We defined those regions as “gaps”.

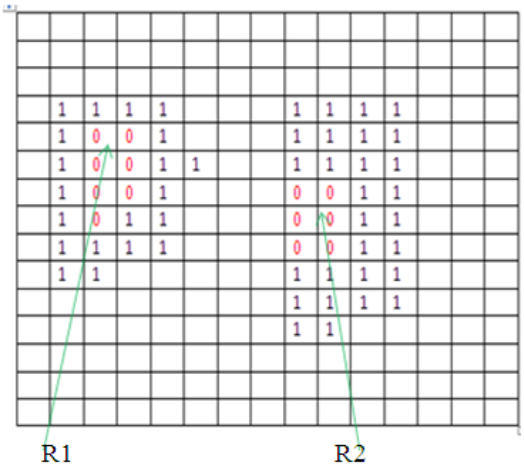


Figure 5.7 R1 is a “hole” region. R2 is what we defined as a gap. R2 can be filled by placing ones between $G_y = -1$ and $G_y = 1$.

Figure 5.7 shows two objects. The first object has a hole inside it. The hole region is denoted R1. The second object has a gap, the gap region is denoted R2. From the discussion above, we know that R1 can be filled by placing 1s between a pair of gradient=-1 and gradient=1 in the x or y direction. However, to fill in R2, we can only use that method in y-direction. If our goal is to fill in both holes and gaps, then, we will take gradients in both x- and y- directions, and take the union of the results in both directions as the final output.

To segment shielded objects in an X-ray image, we can use the region filling method described above. First, we know that the shield object lies within a background region. Also, we know that it must have intensity (thickness) greater than its background region. Therefore, to segment the region of shielded objects, we first cluster the whole thickness map using our hybrid clustering method. After we obtain the cluster map, we look at one level at a time. That is, when dealing with *n*th level, we segment regions that are classified to level *n*, store them in a binary image, and deal only with those regions. Then, we fill in the holes on that binary image. If a filled region has a greater intensity than the current level *n,* that filled region is a region of shielded object. Otherwise, that region is simply just a hole with intensity less than its background. After all the holes of all levels are filled, all objects of interest are identified. Those objects can be used for further analysis.

5.4 Application to Material Discrimination

The result of the hybrid clustering method and the region filling algorithm is very useful for the development of a material discrimination algorithm. The clustering map gives regions of background and object. Therefore, it is easier to do analysis on a region by region basis not on a pixel by pixel basis. The average thickness of a region can be calculated. Therefore, the ratio of high and low energy scan thicknesses can be obtained. With the ratio information, we can obtain information on the materials inside the images.

The concept of adaptive masking discussed in Chapter 4 can also be extended to other segmentation approaches. The objects segmented by the region filling method can be used as masks for filtering out high Z false alarms. Although this method is more complicated than using thresholding, the result of segmentation is more accurate than thresholding segmentation. For complicated scenarios, the method we discussed in this section is more desirable.

5.5 Results and Discussion

Hybrid Clustering

Figure 5.8 shows the result of our hybrid clustering method. Figure 5.8 (a) is the thickness map from a 9MeV energy source. A tungsten cube is placed at the center of 10” thickness of steel plates. A tin cube is placed to the right of the tungsten cube at a slightly higher position. On top of the plates are layers of lead bricks. The bottom layer has more bricks than the top layer. Next to the steel plate is a large lead bar positioned on a stand. Figure 5.8 (b) is the result of the hybrid clustering method. The step size used here is 2”, the radius of the mask is 1. As a result, all pixels are classified to one of 5 intensity levels. Connected pixels belonging to the same level are defined as a region.

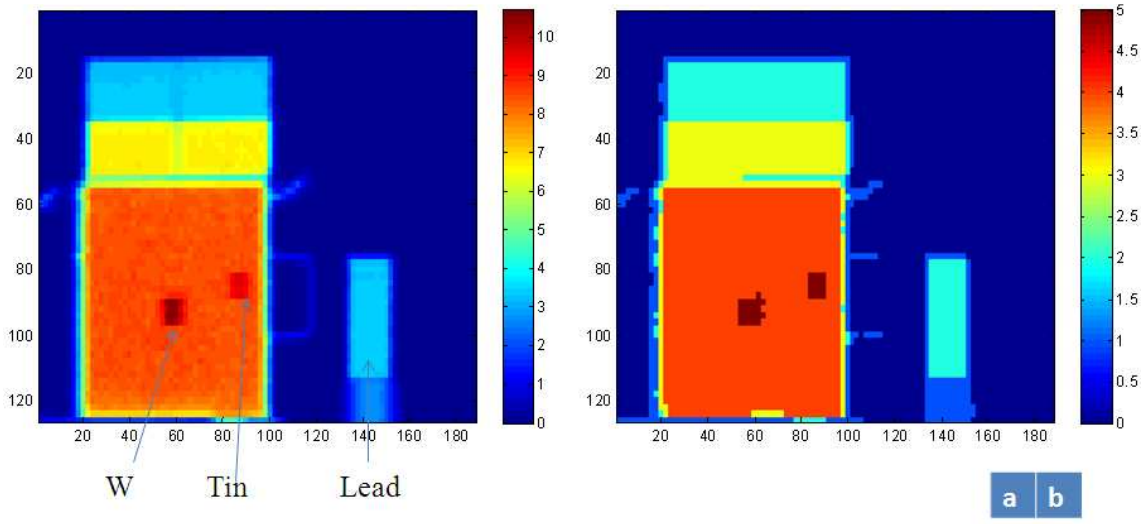


Figure 5.8 (a) The thickness image (b) the cluster map, each color represents a different intensity level.

As we can see in Figure 5.8 (b), a smoother image is the result of clustering. The region of the 100 cc tin cube has become more prominent compared to Figure 5.8(a). Figure 5.9 (a) shows the result of labeling the cluster map. Each region is assigned a color. In this case, we start labeling the regions from the lowest intensity level. Thicker regions in the label map have higher number. As a result, there are 28 regions in the labeled region map. Figure 5.9 (b) is an example of using (a) as a mask for high Z pixel masking. For each of the 28 regions, if the number of pixels identified

as high Z within that region exceeds a threshold, then that region (object) is identified as a high Z region (object). In this case, the threshold used is 50% of the total pixels. Figure 5.9 (b) is an overlay of high Z region and the original high energy thickness map. The high Z regions appear as red in the image. To eliminate some tiny artifacts in small regions, we applied a morphological smoothing (opening followed by a closing) to the masked output.

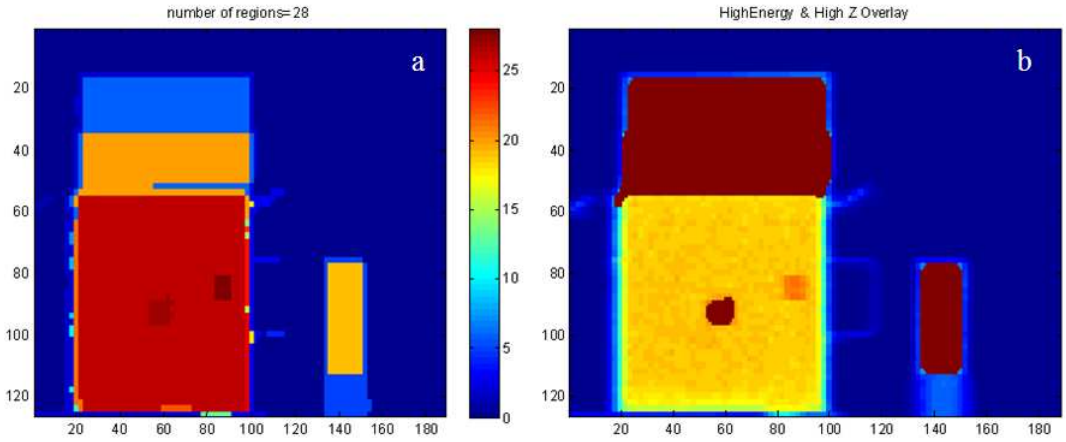
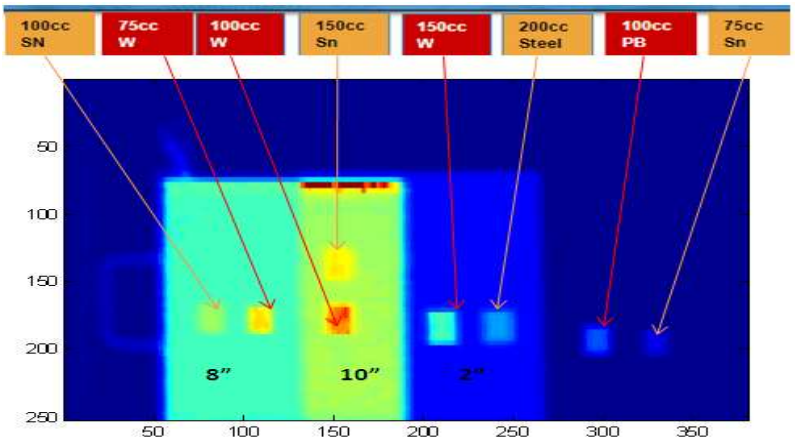


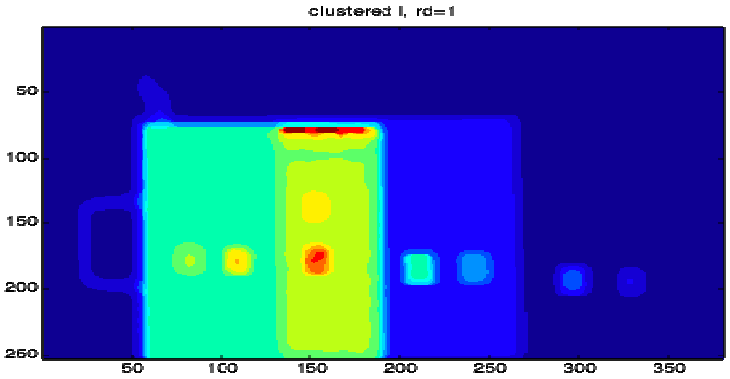
Figure 5.9 (a) The labeled region map. Each number represents a connected region. (b) The overlay of high energy thickness map and segmented high Z regions.

Segmentation using Region Filling

Figure 5.10 shows different objects positioned behind steel plates of different thicknesses: 8", 10" and 2". Each shielding region has one high Z object and one non-high Z object. Figure 5.10 (b) shows the cluster map as a result of hybrid clustering. The quantization step size used here is one inch. We can see that each shielding zone has been classified to one region. Inside each shielded region, there are several sub regions that have greater intensity.



(a)



(b)

Figure 5.10 (a) The thickness map of a setup with different objects behind different shielding plates. Arrows indicate what the corresponding objects are. (b) The cluster map of (a). Pixels with similar thickness are clustered to form a region.

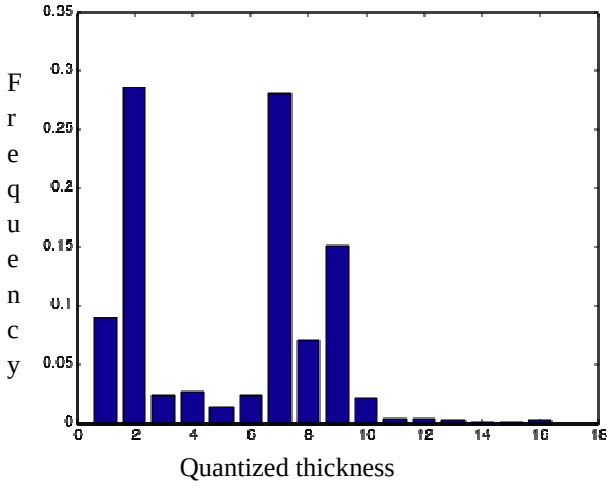


Figure 5.11 Histogram of figure 5.10 (a)

Figure 5.11 shows the histogram of the cluster map shown in Figure 5.10(b). Because the calibration material used here is copper, ten inches of steel is approximately equivalent to 8.7 inches of copper. After quantization, the 10”-steel region will be classified level 9 in the cluster map. From this histogram we know that pixels of level 2, 7 and 9 are dominating. It is therefore reasonable to assume that those are the background pixels. To segment foreground from those background regions, we fill in the holes inside those regions if there are any.

The top row of Figure 5.12 shows three background regions. After histogram analysis, our algorithm will pick the dominant regions (probability greater than 15%). It first finds all regions with intensity equal to 2 in cluster map. Then, it fills all the holes in those regions using the gradient parity check method. Here we take the intersection of the result of x-direction filling and y-direction filling to make sure all the holes and only the holes are filled. A morphological smoothing is applied to the result to eliminate all small artifacts. The final segmentation result is shown in Figure 5.13. Figure 5.13 (a) is a binary image that shows the segmentation of shielded regions. Figure 5.13 (b) is the result of using (a) as mask on the thickness map. The masked region shows that the segmentation method does a fairly good job when the quality of original image as shown in Figure 5.10 (a) is poor in the sense of resolution and noise.

We applied the hybrid clustering and the multistep region filling algorithm in 300 different cases from a simple setup (one object, one shield) to more complicated setups like Figure 5.10 (a). They all give similar results as shown in Figure 5.13 and Figure 5.9. Our method has a very high rate of successful segmentation.

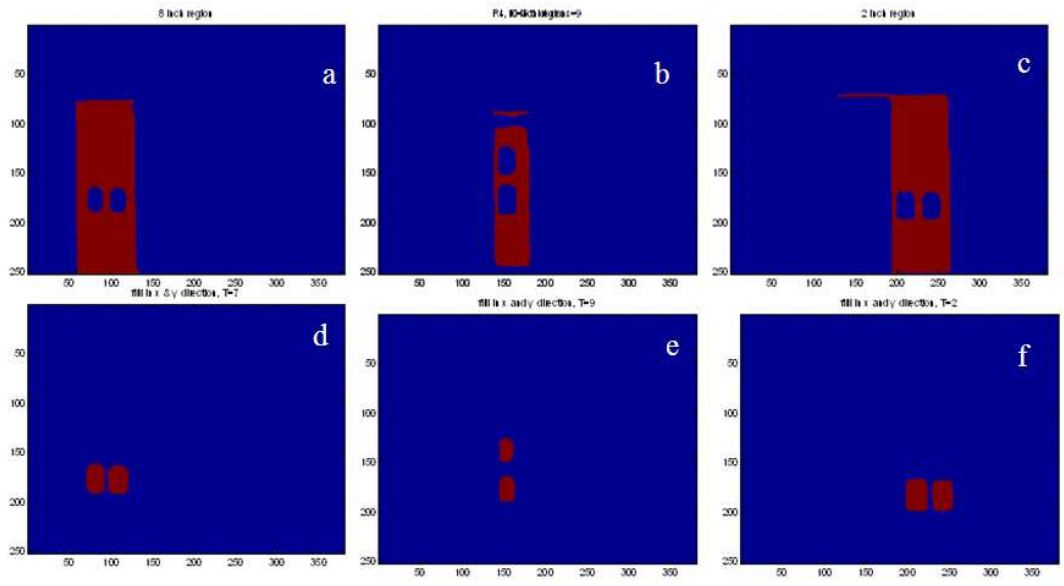


Figure 5.12 Regions of different background and the result of region filling. (a) the 8 in region (b) the 10 inch region (c) the 2 inch region (d) filled region of (a). (e) filled region of (b). (f) Filled region of (c).

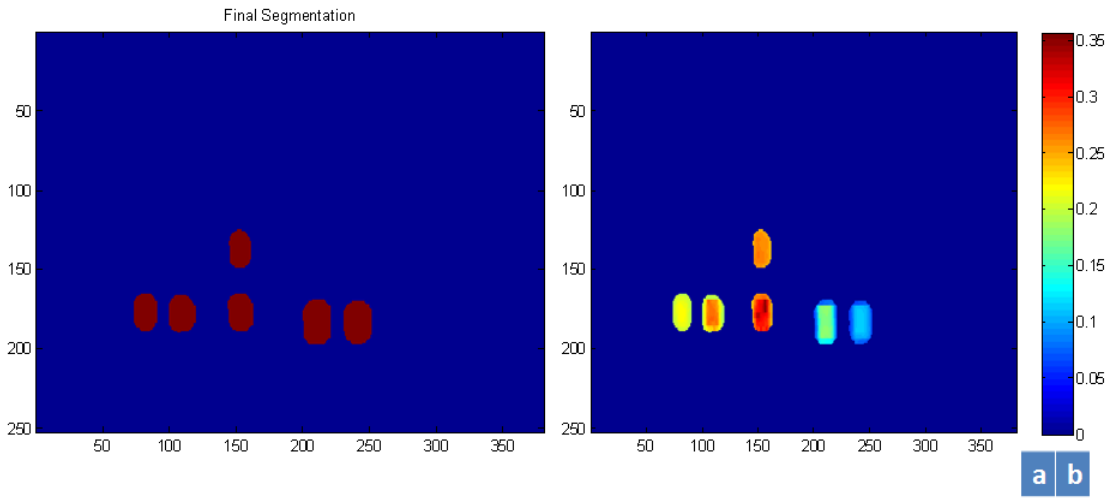


Figure 5.13 (a) result of segmentation, the binary image shows the locations of objects (b) the masked regions of the object on thickness map.

Figure 5.14 shows the result of using our clustering and masking algorithm for material discrimination on a real cargo container. Inside the container, there is an area of bottled waters. Behind boxes of bottled waters, there is a 300cc lead ball, 150cc lead cube, 150cc tin cube, and a 100cc lead cube behind it. Next to the water region,

there are electronics including items such as monitors, computer cases, and wave generators. Three 100cc tungsten cubes are hidden within the electronics and two 100cc tin cubes are placed between the tungsten cubes. Another lead plate is placed on top of a pillar. Outside the container, there is a battery. Next to the battery is a concrete tower with lead bricks on it. Next to the concrete tower, a lead rod is put on a stand. Our goal is to identify materials with atomic number greater than or equal to tungsten (Z=74). Therefore, lead and tungsten should be discriminated from tin. Figure 5.14 (a) shows the cluster map of using a 2-inch step size and 1-pixel radius hybrid clustering. A total of 491 regions are identified. Each region is assigned a number (color). Each region is used as a mask to check if that region has more than 40% of its pixels identified as high atomic number. Figure 5.14 (b) shows the overlay of the segmentation result and the 9-Mev X-ray scan image. The segmented objects are shown as red regions in (b). The detection of objects behind the water region requires another technique called “layer separation” which will be discussed in Chapter 6. The result here shows that the mask filtered out most of the false alarms except a small region on the floor of the container.

Region Merging

Based on the property of the X-ray image, it is sometimes better to use a small step size for clustering and then apply the region merging algorithm to the first stage clustering result. Figure 5.15 shows a result of clustering with a half-inch step size with radius equal to one pixel. The 8-inch steel plate shielding region is classified into many subregions because of noise. A total of 239 regions are the result of clustering.

We then merge any two regions whose difference in region mean is less than 1cm. As a result, about half of the regions are merged. A smoother image with only 121 regions is shown next to the original one. Region merging has a benefit of avoiding improper classifications due to large step size though more processing time is needed.

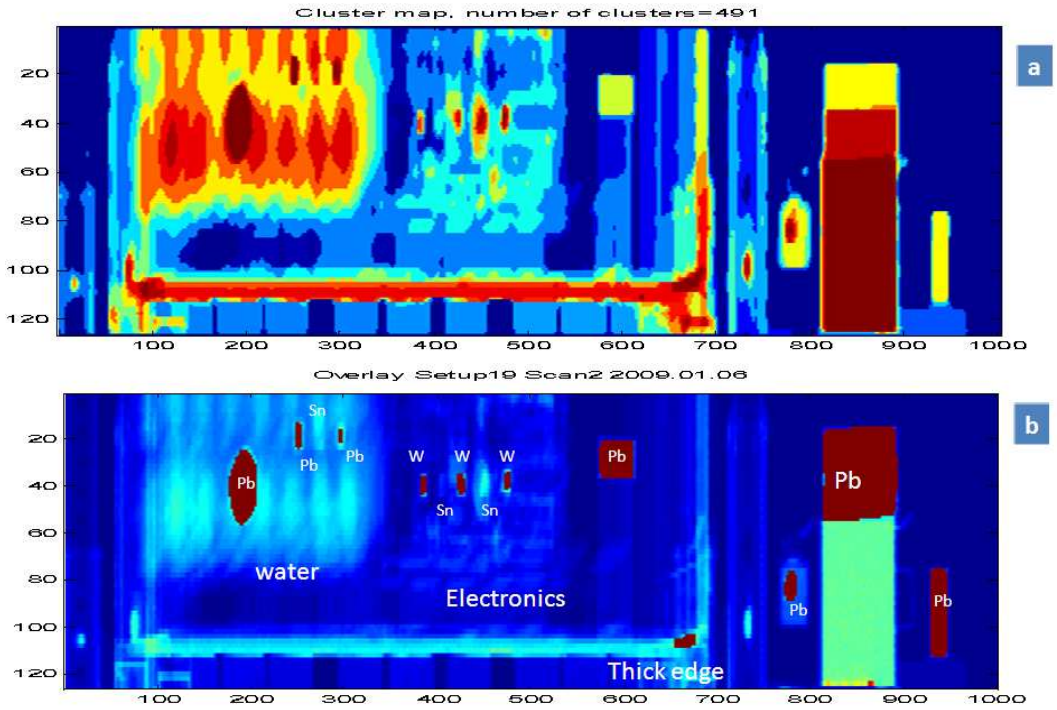


Figure 5.14 (a) The cluster map of the cargo container scan. (b)The overlay of material discrimination result and a cargo X-ray scan image.

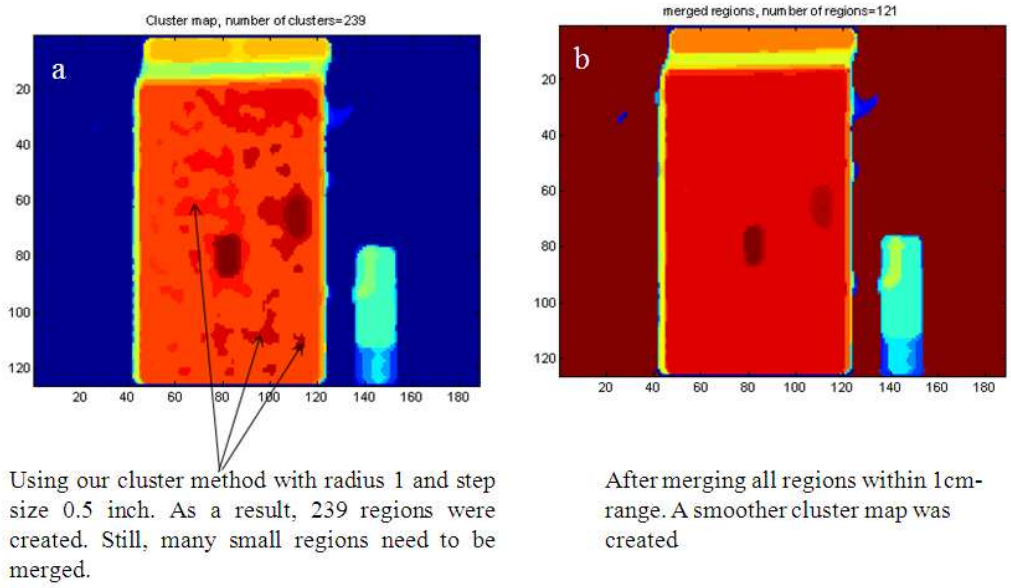


Figure 5.15 The labeled cluster map before and after region merging.

5.6 Conclusion

In this chapter, we discussed a new clustering approach. This new approach is simpler than the traditional k-means and leader clustering and yet it provides greater accuracy in segmentation. Unlike k-means, we do not need to adjust the centroids in each iteration because we use a set of fixed centroids which are the levels of quantization; nor do we need to worry about what number *k* should be used. Unlike the leader cluster algorithm that requires a predefined radius which might limit the size of an object in an image, we let a cluster (region) grow as large as necessary.

The result of clustering and region filling segmentation is very helpful in material discrimination algorithms. The clustering or region filling segmentation result can serve as masks to eliminate false identifications of high Z material. The result of masking a noisy cargo Z image shows that the clustering and masking approach is a very powerful tool for a good material discrimination imaging system.

Chapter 5, in part, is adapted from the material as it appears in the Proceedings of SPIE (2009), Fu, Kenneth; Guest, Clark; Das, Pankaj “Segmentation of suspicious objects in an X-ray image using automated region filling approach.” The author of this dissertation is the primary investigator and author of this paper.

6

Layer Separation for Dual Energy Imaging

6.1 Introduction

In Chapter 4, we discussed a method that used a global minimum ratio for computing high Z probability; and the concept of masking to eliminate false alarm regions was also introduced. The problem with that method is that it still requires human visual inspection to select a region of interest in an X-ray image. In other words, it is not fully automated. The assumption that the ratio of one specified material is a constant was employed in that model. This assumption is challenged when the spectrum of X-rays is not well-controlled to a mono-energetic beam. Also, in Chapter 4, we applied Bojorkhom's method of high Z discrimination by finding all pixels with value above a threshold. That threshold is the ratio of steel, and was in fact, a function of attenuation (equivalent thickness). We simplified his model by applying the same ratio to all pixels regardless of the thickness of the object. In this case, mid-Z objects can have higher probability than some shielded high Z objects. The remedy for that is adaptive thresholding and masking that screen out objects under a similar background. An upside is this method does not require the system to be stable. An unstable system can have different readings of an object's ratio at different time. Therefore, the ratio vs. attenuation lookup table may work on one day but fail to work when there is a tiny change in initial conditions. Since adaptive thresholding and

masking only look for relatively high probability regions, the performance will not be very sensitive to the minor changes of a system.

In this chapter, we propose another approach to boost the accuracy of the performance of an automated dual-energy imaging system. This method for high Z material discrimination can more accurately identify a region of shielded high Z objects. The method in this Chapter uses a ratio lookup table to find possible high Z regions. In order to find more accurately the shielded object, the lookup table is not based on just the ratio of steel but the ratio of steel-shielded high Z objects. Once the relationship between the attenuation and ratio has been established, we use that to find the minimum ratio threshold for each pixel and compute the probability. Masking is still applied to the probability image, but the method is slightly different. One assumption for building the ratio vs. attenuation lookup table is that metals have higher ratio than other materials, and high Z objects are shielded by metals. Therefore, for regions where high Z materials are hidden, the ratio will be higher than the ratio of common metals. In actuality, this is not totally true. In some situations, a measured attenuation will have a ratio lower than that of the same attenuation caused by steel shielded objects. This may cause misses of detection in the low-Z-shielded regions. The layer separation technique was developed to resolve this problem to boost the detection rate.

6.2 Background

Recall in Chapter 4 we discussed the high Z detection method in Bjorkholm's paper [14]. In this paper, the threat detection is based on a lookup table of the threshold ratios vs. attenuation. For various thickness of steel, the average ratio $\bar{R}$ and the standard deviation σ are measured and calculated through a series of experiments. Then, for a measured attenuation, the threshold T is assigned to each pixel based on its high energy attenuation. Detection is called when n, the total number of pixels that pass the threshold, exceeds a preset value N_{max} in a 9x9 examination window.

For the *n* pixels that pass the threshold, the probability of passing the threshold for each of them, $p_{i,}$ is obtained by integrating the normal distribution from $(T - R)/\sigma$ to ∞. And the probability of an 81-pixel square region having *n* pixels passing their thresholds is given by

$$P(n|81) = \frac{81!}{n!\,(81-n)!} p_i (1-p_i)^{81-n}\,. \tag{6.1}$$

The false alarm probability or each examination window is the sum of $P(n|81)$ for all $n > N_{max}$:

$$P_{false} = \sum_{n=N_{max}}^{81} P(n|81) \tag{6.2}$$

It is clear that the probability of false alarm is a function of threshold *T*. The probability of false alarm will decrease as the threshold increases. While Bjorkholm's method correctly points out the need to consider statistical error by computing the probability of each pixel, his model of defining false alarm as shown in Equations (6.1) and (6.2) is, nevertheless, of questionable value. One obvious reason for that is p_i is not identical for all pixels. A binomial distribution requires that all trials have the same probability of success. In this case, all p_i's must be the same. This is simply not the case. Also, the curve of N_{max} vs. P_{false} will be a bell-shape curve centered at $N_{max} = 40$ when $p_i = 0.5$. This creates a problem that the probability of false alarm decreases even when N_{max} increases. It contradicts the assumption that the higher the threshold (of the number of detected pixels) is, the lower the probability of false alarm would be.

Both the measured attenuation and ratio are a combined effect of multiple objects in the path of the X-ray beam. The combined ratio is obtained by:

$$R = \frac{\mu(E_H, Z_{eff})t}{\mu(E_L, Z_{eff})t} = \frac{\sum_{material\ i} \mu_i(E_H, Z_i)\ t_i}{\sum_{material\ i} \mu_i(E_H, Z_i)\ t_i}. \tag{6.3}$$

The ratio is dominated by the term with greatest attenuation. Therefore, if one object is much thicker than the other, or has much greater attenuation coefficient, that object

will dominate the ratio value. Therefore, when a high Z object is heavily shielded by steel, the ratio of the shielded region will be less distinguishable. Figure 6.1 shows curves of ratios of different 100cc materials shielded by steel using Equation (7.3) and values from Table 6.1. Lead (Pb) and tungsten (W) are classified as high-Z materials, while tin (Sn) is classified as a mid-Z metal. It can be seen that as the steel becomes thicker, the three curves will converge to the value of the ratio of steel. The differentiability in ratio of the three objects will be lost when the steel becomes thicker than 25 cm, because of the statistical noise. What makes the problem more complicated is that the system does not know the thickness of any of the material. All the information available is the attenuation, which in our case, is expressed in copper thickness. There is no way to tell how thick an object really is unless we know its attenuation coefficients, and from there, we convert the copper thickness to the material thickness. To do that requires knowledge of the material, which is something we don't have.

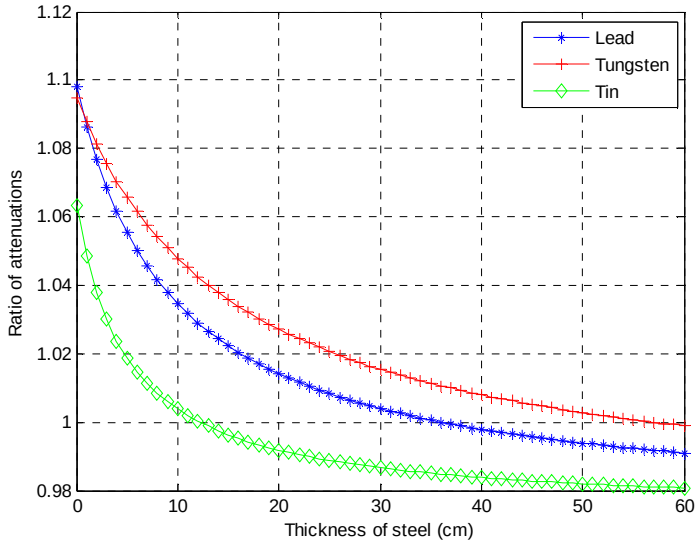


Figure 6.1 Theoretical curves of ratios (for mono-energetic 9 Mev / 6 MeV spectra) of different 100cc cubic materials shielded by various thicknesses of steel.

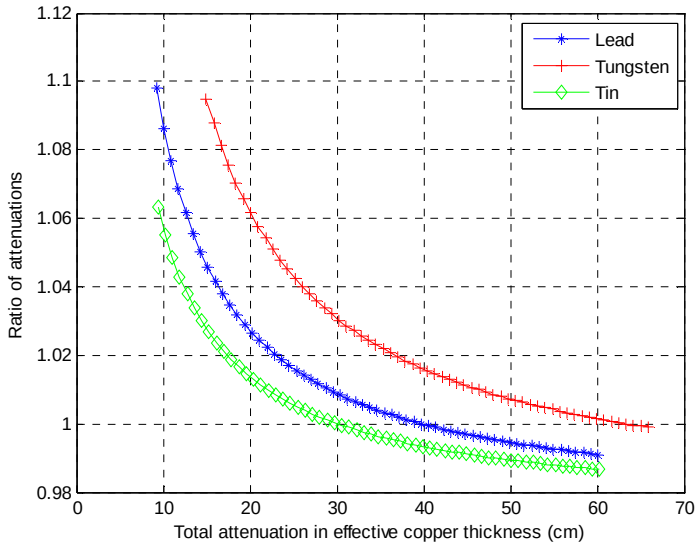


Figure 6.2 Theoretical curves of ratios vs. measured attenuation of different 100cc cubic materials shielded by various thicknesses of steel. Attenuation is expressed in term of copper thickness.

The idea of a ratio lookup table is based on the fact that for the same attenuation observed, the higher atomic number material will have higher ratio of high-and low-energy attenuation coefficients. As shown in Figure 6.2, the theoretical ratio vs. attenuation curves, a curve between the tin curve (green) and tungsten (red) can serve as the curve of minimum threshold. However, computing the probability based on the threshold obtained from a lookup table of steel-shielded object's ratio can be problematic for identifying other layered objects. For example, based on values in Table 6.1, if a 2-cm thick lead object is shielded by 5cm of steel, the total attenuation of a 9MeV energy scan will be

$$\mu(E,Z)t = \sum_{material\ i} \mu_i(E,Z_i)\ t_i$$

$$= \mu_{steel}(9MeV, 26) * 5cm + \quad \mu_{lead}(9MeV, 82) * 2cm$$

$$= \left[2.97 * \left(\frac{cm^2}{g}\right) * 7.86\left(\frac{g}{cm}^3\right) * 5cm + 4.82\left(\frac{cm^2}{g}\right) * 11.35\left(\frac{g}{cm}^3\right) * 2cm\right] * 10^{-2}$$

=2.2614. (6.4)

The same attenuation can be achieved by other shielding with *t* cm of water.

$$\mu_{steel}(9MeV, 82) * 2cm = \mu_{water}(9MeV, 10) * t$$

$$\rightarrow\ t = \frac{\mu_{steel}(9MeV, 82) * 2cm}{\mu_{water}(9MeV, 10)} = 20.2\ cm \tag{6.5}$$

Now, compute the (9MeV/6MeV) ratio of the steel shielded object, the ratio will be

$$R_1 = \frac{\left[2.97 * \left(\frac{cm^2}{g}\right) * 7.86\left(\frac{g}{cm}^3\right) * 5cm + 4.82\left(\frac{cm^2}{g}\right) * 11.35\left(\frac{g}{cm}^3\right) * 2cm\right]}{\left[3.05 * \left(\frac{cm^2}{g}\right) * 7.86\left(\frac{g}{cm}^3\right) * 5cm + 4.39\left(\frac{cm^2}{g}\right) * 11.35\left(\frac{g}{cm}^3\right) * 2cm\right]}$$

$$= 1.03. \tag{6.6}$$

The (9MeV/6MeV) ratio of the water shielded object, the ratio will be

$$R_2 = \frac{\left[2.31 * \left(\frac{cm^2}{g}\right) * 1\left(\frac{g}{cm}^3\right) * 20.2cm + 4.82\left(\frac{cm^2}{g}\right) * 11.35\left(\frac{g}{cm}^3\right) * 2cm\right]}{\left[2.77 * \left(\frac{cm^2}{g}\right) * 1\left(\frac{g}{cm}^3\right) * 20.2cm + 4.39\left(\frac{cm^2}{g}\right) * 11.35\left(\frac{g}{cm}^3\right) * 2cm\right]}$$

$$= 1.003. \tag{6.7}$$

If an unshielded tin object causes the same attenuation, its thickness is

$$t_{Sn} = \frac{\mu_{steel}(9MeV, 26) * 5cm + \ \mu_{lead}(9MeV, 82) * 2cm}{\mu_{Sn}(9MeV, 10)}$$

$$= \frac{2.2614}{3.82 * 10^{-2}\left(\frac{cm^2}{g}\right) * 7.31\left(\frac{g}{cm}^3\right)} = 8.1\ cm. \tag{6.8}$$

The ratio of the object will be

$$R_3 = \frac{\left[3.82\left(\frac{cm^2}{g}\right) * 7.31\left(\frac{g}{cm}^3\right) * 2.12cm\right]}{\left[3.583\left(\frac{cm^2}{g}\right) * 7.31\left(\frac{g}{cm}^3\right) * 2.12cm\right]} = 1.0661. \tag{6.9}$$

For the purpose of high Z detection, the ratio threshold T is a value greater than the ratio of steel. It can be obtained by measuring the ratio of 2-cm thick lead behind various thicknesses of steel. As demonstrated in equations (6.4) to (6.7), for a measured attenuation of 2.61 caused by 5cm of steel and 2 cm of lead, the ratio threshold from the lookup table will be $1.03 \pm \delta$ for δ being a tiny variation. However, for the same attenuation caused by 2cm of lead shielded by 20cm of water, using that threshold is too high and will eliminate the possibility of detecting the lead object. Moreover, (6.8) and (6.9) show that an eight-centimeter thick non-high Z object, tin, can cause the same attenuation with ratio being 1.066. The fact that $R_1 > R_2$ means in order to correctly detect the lead, we need to lower the threshold; however, the fact that $R_3 > R_1$implies the need to increase the threshold in order not to have false alarms caused by a thick tin.

In reality, curves ideally separated by up to 40cm copper thickness as shown in Figure 6.2 do not really exist, because different materials have different hardening effects. The curve of tin will cross over the other two cures when the total attenuation is about 25cm. Also in reality, organic materials (low-Z), which are very common in cargo containers, act the same as water, having the same shielding effect as shown in (6.7). Using a single threshold (obtained from the curve of the same material) will not effectively detect the shielded high Z objects.

To solve the problem of choosing the threshold, information of whether an object is shielded and the type of the shielding material is important. The techniques discussed in Chapter 5 are very helpful in developing the approach to achieve that goal. The so called “layer separation” algorithm defines objects and backgrounds using the hybrid clustering approach. The idea is to define the property of each region (based on the ratio) and separate objects from their background if they are shielded. Therefore, low-Z shielded objects can be identified and their ratios and probabilities can then be adjusted to obtain a more desirable result. This method is proven to be able to significantly improve the accuracy by suppressing false alarms and increasing the detection rate. It will be discussed in the next section.

Table 6.1 Attenuation coefficients and ratios for some materials. (Values obtained from NIST website)

	H_2O (Z=10)	Steel (Fe) (Z=26)	Sn (Z=50)	W (Z=74)	Pb (Z=82)	U (Z=92)
density(g/cm^3)	1	7.86	7.31	19.3	11.35	18.68
μ_{9MeV}/ρ	2.31E-02	2.97E-02	3.82E-02	4.61E-02	4.82E-02	4.583E-02
μ_{6MeV}/ρ	2.77E-02	3.05E-02	3.583E-02	4.21E-02	4.39E-02	5.195E-02
$r = \frac{\mu_{9MeV}}{\mu_{6MeV}}$	0.8339	0.9738	1.0661	1.0950	1.0979	1.1335

6.3 Methodology

The objective is to develop an algorithm for a dual energy system that can discriminate all hidden high Z objects in a cargo container. The ability to find objects behind metal was already demonstrated in Chapter 4. The more important goal is to find the objects with various kinds of shielding.

6.3.1 Image creation

The X-ray images are obtained through a 9/6 MeV dual energy system moving along the container at a speed of 33 inches per second (ips). The strength of the source is enhanced by 30% over that originally used in Chapter 4. The vertical resolution of the system is about 0.5 inch. The X-ray source emits X-ray pulses at a repetition rate of 400Hz with high energy and low energy generated alternately. Each detector of the system is calibrated individually. The number of photons received by each detector is then converted to the equivalent thickness of copper. For each scan, two equivalent thickness images are created, one from high energy pulses and one from low energy pulses, denoted $H(x, y)$ and $L(x, y)$ respectively. The calibration process also creates two lookup tables for standard deviations of each measured thickness for high and low energy scans. The ratios of the two images are taken to create a ratio image, which can be expressed as

$$R(x,y) = \frac{H(x,y)}{L(x,y)}. \tag{6.10}$$

Before the ratio is taken, the two X-ray images are filtered by a Wiener filter to enhance the quality of ratio information.

A lookup table of minimum ratio for high-Z versus total attenuation (in terms of equivalent copper thickness) is created, in this case by using the ratios slightly less than the ratios for 100cc of lead behind varying thicknesses of steel. For different systems, the values of the lookup table will be different. Even a slight change in the same system will yield different results. Figure 6.3 (a) shows the lookup table we use for creating results demonstrated in this chapter. The minimum size of interest is 100cc cube of lead, which is approximately 4.6 cm thick on each side. Its equivalent copper size is 6.9cm. For pixels with a measured thickness less than 6.9 cm in the high energy image, the minimum ratio requirement will be the ratio of an unshielded lead cube (flat part of curve). This is based on the assumption that if an object is thinner than 6.9cm of copper, it is either a larger object which is not as dense as lead (i.e. lower Z) or a smaller, denser object. Each equivalent thickness on the curve shown is the sum of 6.9cm of lead plus the thickness of shielding steel.

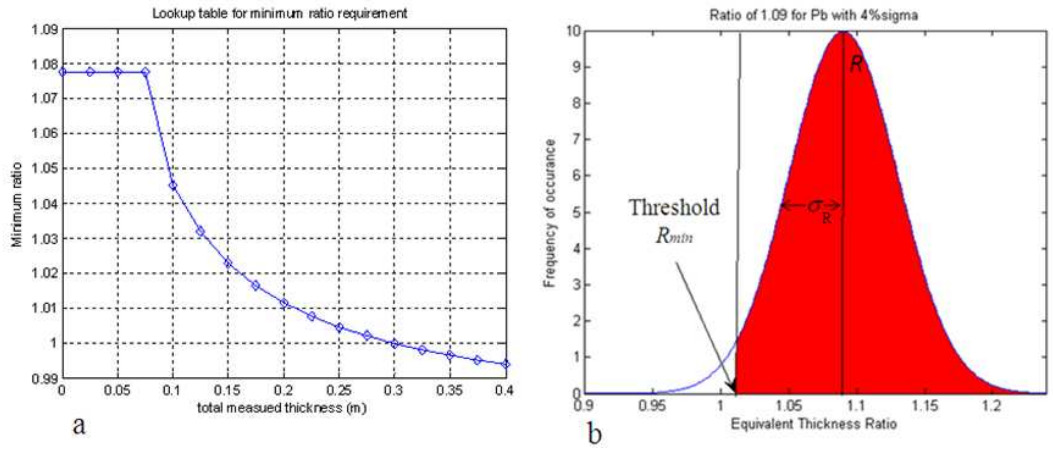


Figure 6.3 (a) Curve of minimum ratio versus total thickness requirement of being high-Z. (b) For a measured R=1.09 and equivalent copper thickness 19cm, the probability of being high Z is the area of the red region.

The certainty of a pixel being high Z is obtained by calculating each pixel's probability of having ratio greater than the minimum ratio threshold. As discussed, assuming the noise is Gaussian-like, the probability is obtained by integrating over the normal distribution from $(R_{min} - R))/\sigma_R$ to ∞. The probability can be expressed as

$$P(x,y) = 1 - 0.5\left[1 + \text{erf}\left(\frac{R_{min}(x,y) - R(x,y)}{\sqrt{2}\,\sigma_R(x,y)}\right)\right] \tag{6.11}$$

where $\sigma_R$is the standard deviation of the measured ratio, and R_{min} is the minimum ratio requirement, which is a function of thickness, obtained from Figure 6.3(a). If a pixel has calculated thickness equivalent to 19cm of copper in the high energy image, and measured ratio R=1.09, its minimum ratio for being high Z will be 1.01 according to Figure 6.3(a), and the certainty of that pixel being high Z will be the area of the red region shown in Figure 6.3(b). If an observed ratio is equal to the minimum ratio threshold, the probability will be 0.5. If it is higher than that threshold, the smaller the standard deviation σ_R, the higher the probability will be.

6.3.2 Image segmentation

The entire image is segmented into regions, or clusters, based on the high energy attenuation. Pixels with similar equivalent thickness and ratio will be grouped together as a cluster. The segmentation method used here is the hybrid clustering method, which is a mixture of *k*-mean and seed growing segmentation approach discussed in Chapter 5.

For a quick review, flow charts of the segmentation algorithm are shown in Figure 6.4 and Figure 6.5. For hybrid clustering, each pixel of an input image *I(x, y)* is compared to the value to its neighboring regions. Four *d* by *d* filters Q_1 to Q_4 compute the average values of four quadrants centered at (*x*, *y*). A (2*d*+1) by (2*d*+1) filter *Q* computes the average value of the four quadrants. One of the outputs of the five filters, which is closest to the value of *I*(*x*, *y*), will be quantized and assigned to the new image I_{out}. The quantization is done by choosing a "step size", the non-quantized values will be rounded to the closest level of quantization, which is an integer multiple

of step size. In image I_{out}, the connected pixels with the same value form a region, or cluster; each cluster is then labeled by an integer.

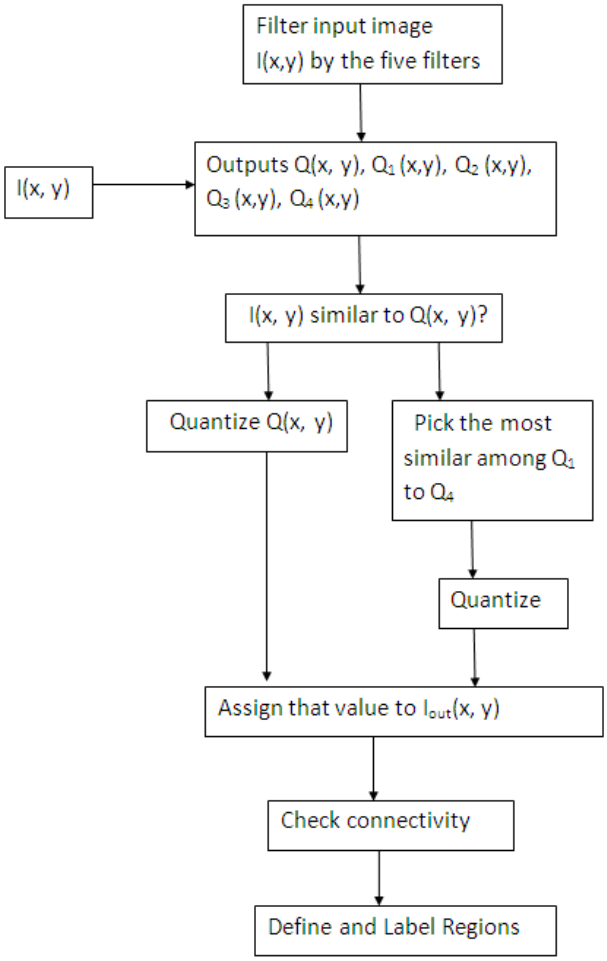


Figure 6.4 Flow Chart of hybrid Clustering Algorithm.

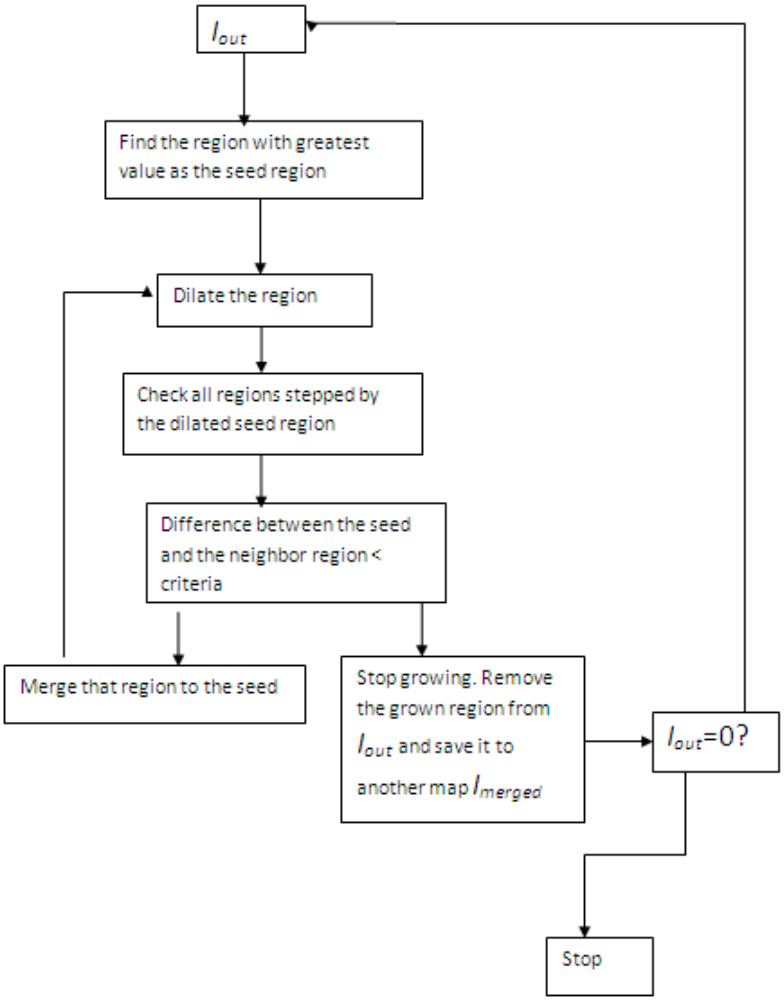


Figure 6.5 Flowchart of region merging algorithm.

Region merging, as shown in Figure 6.5 is performed on I_{out} to refine the clusters. Neighboring clusters will be merged together when their difference is within a preset value. This is done by using the thickest region as a seed. Starting from the seed, we check all adjacent regions and merge only the regions that meet the criteria. After a new merged region is formed, we check the neighboring regions and go through the same procedure again. The region stops growing when no more regions that meet the criteria can be found. The next seed is chosen from the second thickest regions and the procedure is repeated. The algorithm stops when no more seeds can be found.

6.3.3 Layer Separation

Layer separation is performed on regions that are shielded by low Z materials and those regions only. The reason for doing so is that the lookup table is developed using mid-Z shielding (shielding material with ratios ~1.0). Low-Z shielding (ratios less than 1.0) results in lower total ratios for shielded objects. Based on experimental results, the ratio range for these regions are 0.92(paper) ~0.97 (bottle water). The method of layer separation is illustrated by Figure 6.6. An object of thickness t_1 is placed behind another kind of material of thickness t_2. The total attenuation M' is:

$$M' = \mu_1 t_1 + \mu_2 t_2 \tag{6.12}$$

while the attenuation caused by the background is only

$$M = \mu_1 t_1, \tag{6.13}$$

Also, we have $M' > M$ since X-ray goes through more layers. Therefore, for each region that the attenuation is greater than its background, we can assume that the object is shielded by the background object as demonstrated in Fig. 6.6 (a). Therefore, the attenuation caused by the object alone will be

$$M'' = M' - M = \mu_2 t_2 \tag{6.14}$$

The ratio of Object 2 in a (E_H, E_L) dual energy system is

$$R = \frac{M''(E_H)}{M''(E_L)} = \frac{M'(E_H) - M(E_H)}{M'(E_L) - M(E_L)} = \frac{\mu_2(E_H)t_2}{\mu_2(E_L)t_2} \tag{6.15}$$

The attenuation or the equivalent copper thickness of the background is calculated by (1) identifying the location of the shielded object, (2) dilating the object by the morphological operation, (3) subtracting the original region from the dilated region. The remaining ring-shaped region is the background, or shielding for the object, and the thickness of that region is the thickness of the material shielding the object of interest. Figure 6.6(b) to (d) illustrates the procedure. The red region in (b) is the object of interest buried in a blue background. In (c), the red region is dilated by

applying a 3x3 square structuring element. The red ring region in (d) is the region 1-pixel outside from the object region, the average attenuation of that region is used as the attenuation of the first layer (t_1), which is subtracted to obtain the attenuation of the second layer (t_2).

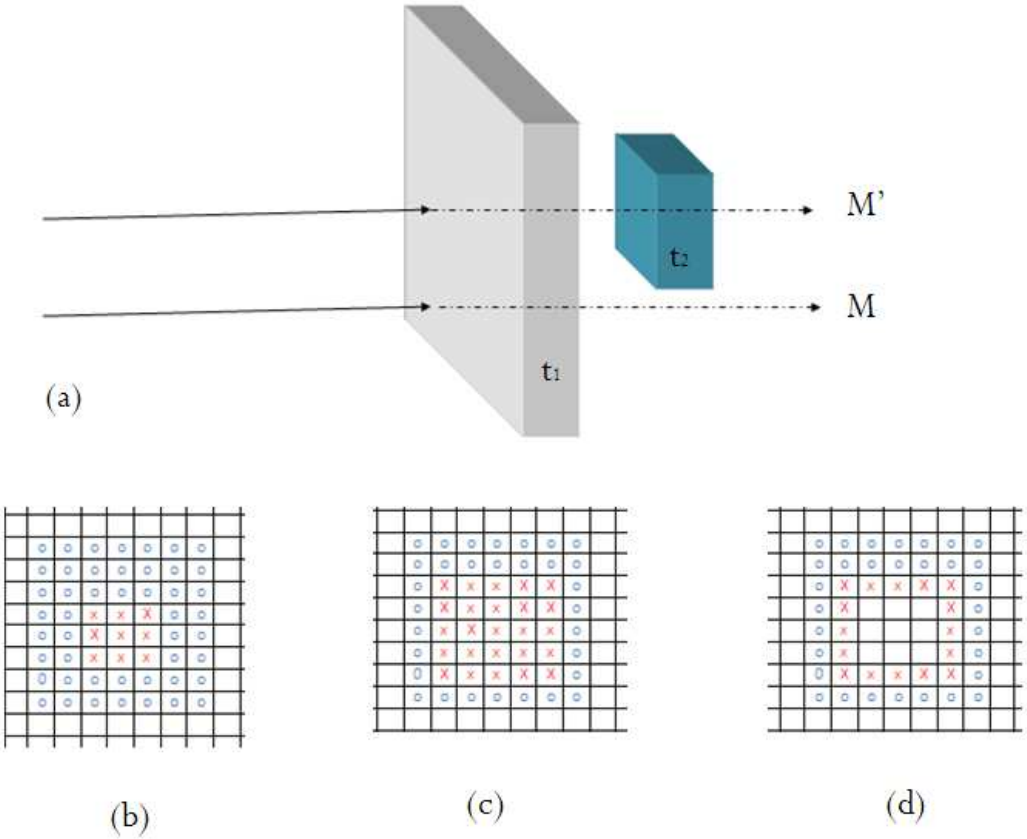


Figure 6.6 (a) X-rays pass through two layers of object. (b) to (d) indicate the procedure of finding the thickness of background.

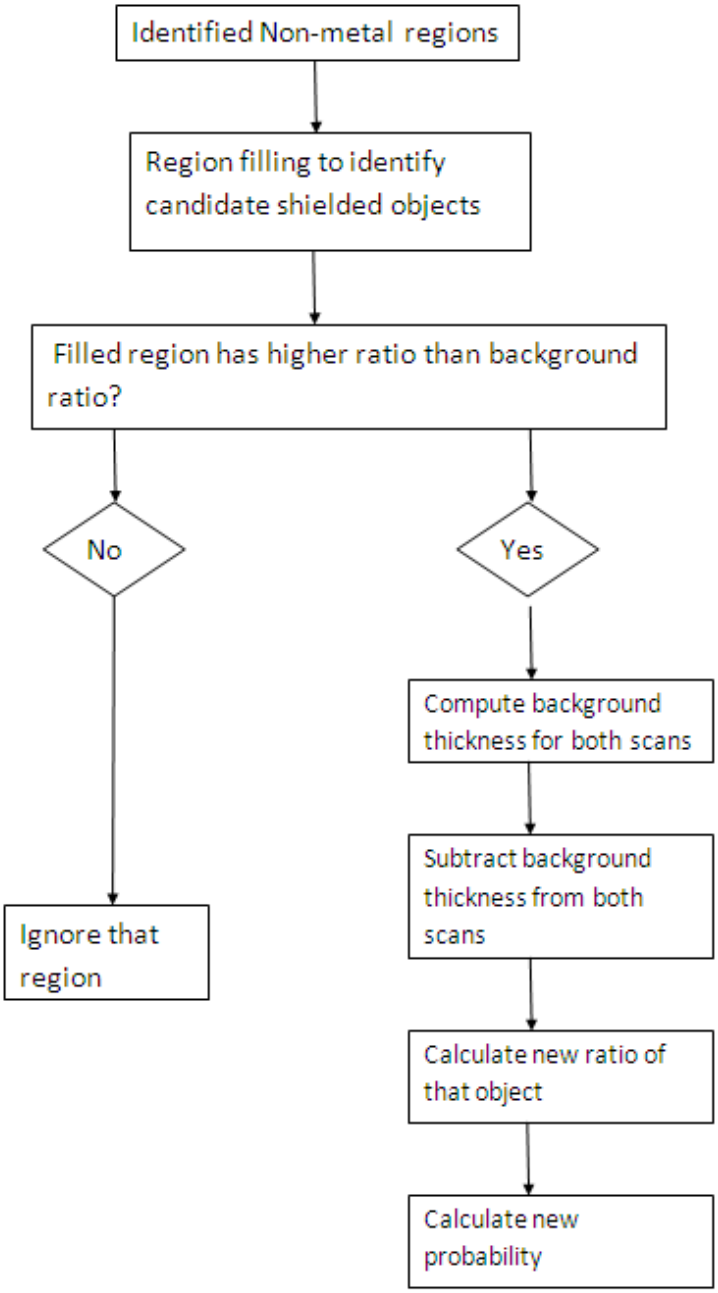


Figure 6.7 Flowchart of layer separation algorithm.

Using the layer separation method described above, the attenuation ratio for the object of interest can be more accurately calculated. Although noise and scattering will still affect the precision of the newly calculated attenuation, this method can help detect high-Z objects where they might otherwise be hidden. As discussed, we only need to apply this layer separation method to regions where the lookup table created for metal shielded regions fails. Figure 6.7 illustrates the procedure of applying layer separation on such regions.

For each cluster obtained from the results of segmentation, we calculate the average ratio of that region. All regions with ratio less than 0.96 are identified. Regions with ratio less than 0.7 can be disregarded since hydrogen (Z=1) has R=0.82, any pixels with ratio lower than that are unreasonable. Those pixels are assumed to be air or materials with very low attenuation. The ratios of high-Z objects shielded by those materials will not be significantly affected according to Equation (6.3). After all those low atomic number material regions are identified, we perform the region filling algorithm. All regions being filled are the candidate regions for objects of interest. If a filled region has greater ratio than its background, it can possibly be metal. We also ignore regions whose sizes are too small to be considered a threat. Layer separation is then performed on the possible metal regions. Figure 6.8(b) shows a ratio map of a cargo scan shown in Figure 6.8 (a). The red regions in Figure 6.8 (c) indicate the regions of low ratios. Two holes are left in the bigger region meaning that there are possibly two objects being shielded. They are the object of interest (OOI.) Figure 6.8 (d) shows the regions being filled by filling algorithm. Both of those regions have average ratios greater than their background.

The ratios of the layer separated objects are calculated. The lookup table for minimum ratio threshold is then mapped to the new calculated equivalent thickness. Because greater error will be associated with the new thickness, the standard deviation of the new thickness is assumed to be greater. To minimize noise, we take the average over all pixels in that region to obtain the thickness and its corresponding standard

deviation. That is, all pixels in the low Z shielded region will be assigned a single value of thickness.

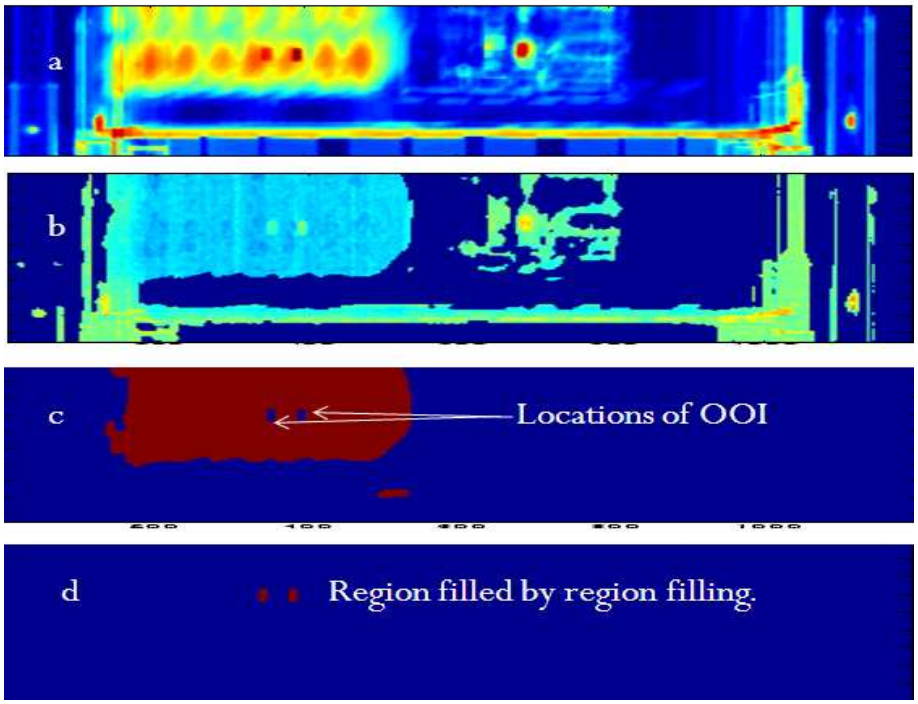


Figure 6.8 (a) A cargo container image. On the left side of the container are boxes of bottled water. (b) The ratio image. The light blue regions are water. Two green regions inside the water region are object of interest. (c) Regions of ratio<0.97. (d) Result of region filling. The filled regions are water shielded.

6.3.4 Decision Making

If a pixel has probability greater than the minimum probability threshold, we call it a detection. Noise from scattering and statistical errors will cause false detection. We eliminate false detection pixels by a democratic procedure: if in a certain region, the number of pixels detected exceeds a certain threshold, all pixels of that region will be called detected. Otherwise, none of the pixels in that region will be counted as detected. By doing this we eliminate many false alarms. We name this kind of false alarm reduction method “masking” in Chapter 4 since the regions are used as masks to map all detected pixels to make decisions.

6.4 Results

Figure 6.9 (a) shows a cargo X-ray image in colormap, with the greatest value red and the smallest value blue. There are a total of eight objects being tested inside the cargo container. Object 1, 2, 3 and 4 are placed behind boxes of bottled water. Object 1 is a 400cc lead ball, and objects 2 and 4 are just 100cc cubes of lead. Object 3 is a 200cc steel cube. Objects 5 to 9 are placed behind electronics. Objects 5 and 7 are 100cc of lead; objects 6 and 8 are 100 cc cubes of tin, which is a material of high ratio that can easily cause misclassification. The purpose of putting those objects there is to test how well the algorithm can distinguish the non-heavy metal tin from heavy metals. Object 9 is a 100cc cube of tungsten, which in our classification belongs to heavy metal group. Outside the cargo container, there are two batteries sitting on stands. Batteries contain a high percentage of lead. They are therefore very likely to be classified as a threat object. On the right of the image, there are steel plates shielding a pile of bricks. On top of the bricks are lead bricks. Finally, a lead brick is standing next to the right of the steel plates. Those unshielded lead bricks are used to test how well the system can identify the free air unshielded objects.

Figure 6.9(b) shows the result of hybrid clustering. The pixels whose equivalent thicknesses are within a half inch difference are grouped together. Each color in this map represents an object (region). The assumption here is that most objects inside the container are relatively uniform in thickness. For objects with shapes having various thicknesses, they can be clustered into different regions. Figure 6.9 (c) is the ratio map of the high and low energy scans. The ratio in the water region is about 0.932. The average ratio of object 1 is 1.006, approximately the same as the ratio of steel. The average minimum ratio for that region is 1.012. Therefore, it won't be classified as a high Z object using the method in Bjorkholm's paper. Figure 6.9 (d) shows the new ratio image after we applied the layer separation on that region. The new ratios for objects 1 to 3 are recalculated. The new ratios for objects 1, 2 and 4 are 1.19, 1.092, and 1.0677 respectively. As a result, most of the pixels in those regions of interests pass the minimum ratio requirement. Ideally, we will have the new calculated

ratio values approximately equal to that of the unshielded lead bricks. Due to statistical errors, the new ratio of object 1 ends up to be 0.11 greater than that.

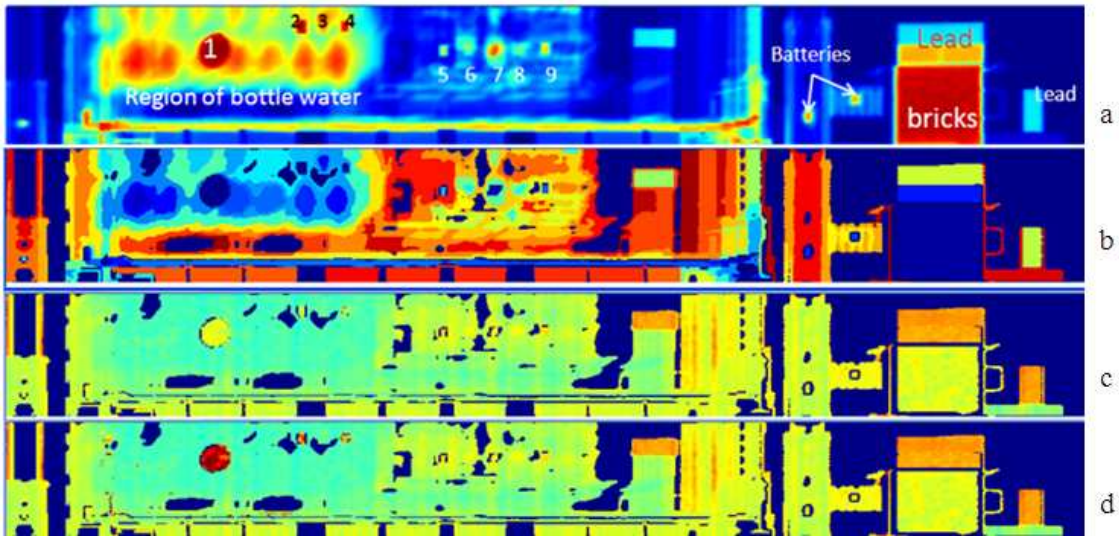


Figure 6.9 (a) The X-ray image of a container scan, which is shown in a colormap with minimum being blue and maximum being red. Nine objects being tested are placed in different regions. (b) The cluster map of image(a). (c) Ratio map of (a). (d) Ratio map corrected by layer separation. The ratios of three objects in the water zone are recalculated.

Figure 6.10 illustrates the improvement that layer separation can make for high Z material classification. The image in Figure 6.10 (a) is the same image as Figure 6.8 (a) shown in inverted grayscale. Object 3 is more discernable in this kind of display. The second row indicates all the pixels whose probability of ratio exceeding the minimum ratio requirement are greater than or equal to 0.5. All those pixels are shown in red. We can see that noise has created many false classification pixels. Especially in the steel plate and brick region, the standard deviation is greater due to the low penetration caused by greater object thickness. The image in Figure 6.10 (c) is the result of applying masking to eliminate the false alarms. For each cluster, that region is called a high Z region only if the total number of identified heavy metal pixels exceeds 40% of the total number of pixels of that region. A significant improvement in the brick region can be seen. This masking method is very useful in eliminating false

alarms. While having great effectiveness, it is conceptually less complex and intuitively makes more sense than using a 9 by 9 examination window. The two batteries outside the container are alarms since they contain a lot of lead. An inspector can eliminate the possibility of that kind of object being threat when additional information is available.

After layer separation is applied, the ratios of the four objects in the water zone are recalculated. The probabilities of those objects having ratio exceeding their thresholds are also recalculated and segmented. As a result, shown in Figure 6.10 (d), the three high Z objects produce alarms. The 200cc steel cube still has a relatively smaller ratio. This image shows a very successful result of material classification for the practical customs use purpose.

The superiority of our approach can be demonstrated by comparing the results shown in Figure 6.10 (d) and Figure 6.10 (e). Figure 6.10 (e) shows the effective Z-image of the container created by the conventional method, where red represents high effective Z value and blue represents low effective Z. The water region in Figure 6.10 (e) indicates a very low atomic number. The regions of the three water-shielded objects also appear to be very low Z objects according to this Z image. This illustrates how the conventional method fails to make a practical high Z detection.

A series of experiments were conducted to test our method. Most of the low Z shielded objects can only be identified through the use of layer separation. Occasionally, false classification happens in those regions, but it occurs less than 8% of the time, which is a much better performance than current systems that can only provide information about the effective atomic number but cannot identify shielded objects.

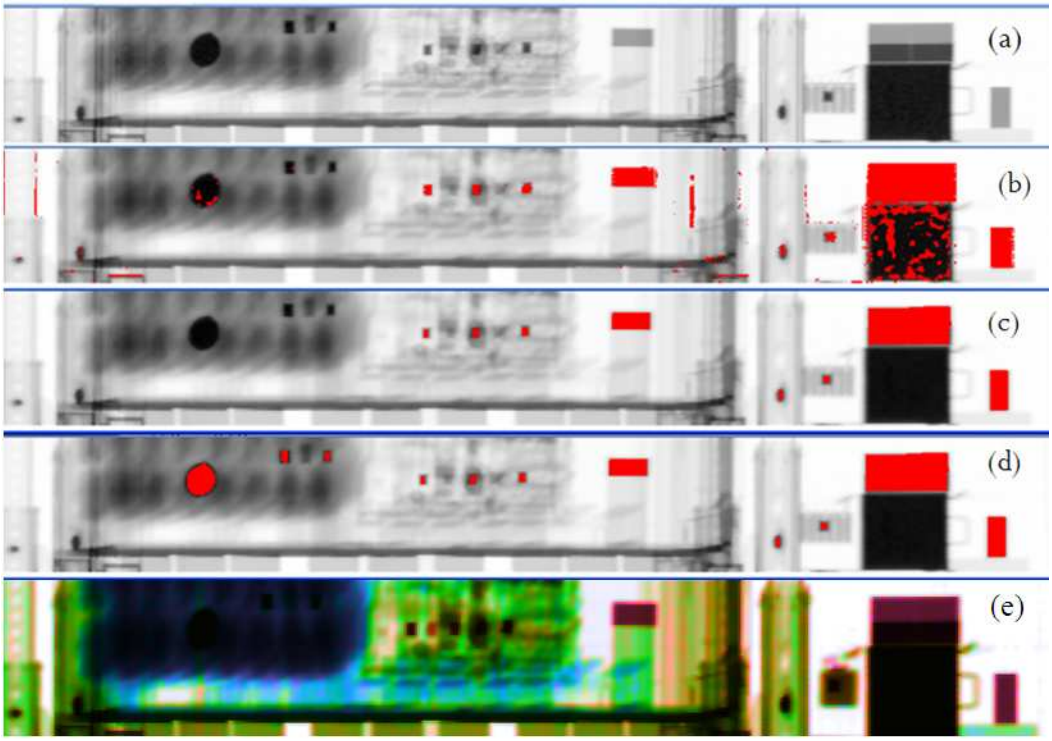


Figure 6.10 (a)The same scan shown in Figure 6.9 is displayed in inverted grayscale;(b)all pixels with 50% probability of being greater than the minimum ratio requirement are indicated in red. At this stage, no layer separation or masking is done. Random red pixels can be seen on the edges of objects. Also, not all pixels in a heavy metal region are red. (c) Result of applying mask to eliminate false alarm pixels. Random pixels are gone and heavy metal objects solidly in red (d) Result of applying layer separation in the water region to recalculate the probability. All three heavy metal objects are detected this time. (e) z-image obtained from conventional approach.

6.5 Conclusion

In this chapter, we proposed a method for shielded high Z material detection. Our approach overcomes the shortcomings of previously proposed methods. By applying hybrid clustering and masking, an object is used as an examination window. In this way, we can more effectively remove the false alarms caused by large objects. By applying layer separation, we are not limited to only measuring the effective Z value of each pixel like all the current systems do. We can actually find a hidden threat. Because statistical noise is not the only source for false alarm, factors such as

random scattering can also affect the measurement of thickness or ratio in a cluster and several occasions of false classification are still possible.

The concept of layer separation is very important for a dual energy system. Besides the method we introduced in this chapter that adjusts the measured ratios and maps them to the threshold map, one can also compute the probability by creating a different lookup table for each different shielding material. By identifying the ratio of an object's background, we can then switch to the corresponding lookup table for high Z discrimination. This will be one aspect of our future work.

Although decent results can be obtained, currently, the outputs of the system still need human assessment as a safeguard. However, our approach shows a promising step toward a fully automated classification system.

Chapter 6, in part, is a reprint of the material as it appears in the Proceedings of SPIE/IS&T (2020), Fu, Kenneth; Ranta, Dale; Guest, Clark; Das, Pankaj "Layer separation for material discrimination cargo imaging system." The author of this dissertation is the primary investigator and author of this paper.

7

Processing X-ray Images in the Wavelet Domain

7.1 Introduction

In previous chapters, we focused our discussion on processing X-ray images in the spatial domain. The convolution of spatial filters and images is equivalent to multiplication of the Fourier Transforms of the two. Processing images in the spatial domain has the same effect as processing them in the frequency domain. There are three major parts in developing an algorithm for material discrimination: noise reduction, visual quality enhancement, and machine learning. So far, the noise reduction parts of our algorithms are all done by applying a Wiener filter. The Wiener filter, by estimating local variance, (or local spectra amplitude), achieves MSE optimization. This kind of filter is reported to be very effective in processing X-ray images [47]. In this chapter, we will exploit the application of processing X-ray images in a totally different domain—the wavelet domain.

Recently, processing images in the wavelet domain has gained popularity. Wavelets have a broad range of image processing applications such as coding, compression, denoising and image fusion. Wavelet denoising methods are proposed to deal with noise associated with radiography imaging such as X-ray imaging (Poisson), or magnetic resonant (MR) imaging (Rician). Here, we propose a wavelet

based algorithm for a dual energy imaging system for high Z detection. The concepts of wavelet denoising and fusion will be briefly discussed in Section 7.1.1 and Section 7.1.2. In Section 7.2, we discuss our wavelet algorithm. Our method for high Z detection is still the statistical model aided by adaptive masking. A final review of the statistical model and a new interpretation will be given in Section 7.2. The main point of this chapter is to show how wavelet domain processing can be applied to our system, and to investigate whether processing our images in the wavelet domain will provide any advantage that processing images in pixel domain does not.

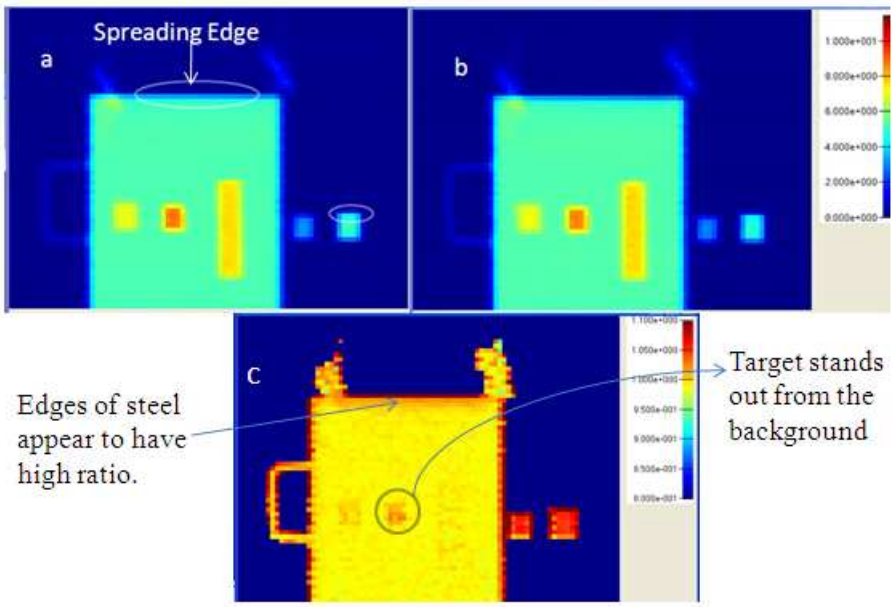


Figure 7.1 (a) low energy X-ray image (b) high energy X-ray image (c) ratio image.

7.1.1 Wavelet domain denoising for X-ray images

Figure 7.1 shows three images from a dual energy system: the high- and low-energy thickness map, and the ratio map. In thickness maps, regions are mosaic-looking. The reason for this is that greater thickness causes less penetration. As discussed in Chapter 2, a major source of noise in X-ray imaging is quantum noise, which is modeled by the Poisson distribution. The Poisson nature of the system is due to the variation of number of photons emitted at different time. While scanning with a

moving source, different locations will receive various counts of photons even when the scanned object is uniform.

For Poisson noise, the standard deviation is the square root of the mean and, the signal to noise ratio is proportional to the standard deviation. When penetration is good, the variation between counts is insignificant. The result of poor penetration is low counts of photons, which means smaller *SNR* and greater uncertainty. The problem of noise in low-count photons can be dealt with using wavelet domain processing [71-74]. The ratio map in Figure 7.3(c) shows the problem of the edge effect. We previously proposed using morphological filtering and adaptive masking to solve this problem. In this chapter, we also attempt to resolve this issue in wavelet domain.

A wavelet ψ(t) is a waveform function that, when scaled and shifted, forms bases that in linear combination can represent another function in a so-called wavelet domain. Those bases are called child wavelet or baby wavelet. The positive real number *a* represents the scale, or dilation, of the wavelet; the integer *b* represents the shift in *t*. For discrete signals, discrete wavelet transforms (DWT) are applied, and each baby wavelet is in the form of

$$\psi_{m,n}(t) = a^{-\frac{m}{2}}\psi(a^{-m}t - nb). \tag{7.1}$$

where the integer *m* represents the level of the transform, and integer *n* is the shift in time. Commonly used values for (*a, b*) are (2, 1). Going through high- and low-frequency wavelet filters, a signal *S* can be decomposed to an approximation A_1 and detail D_1. Following the same procedure, the approximation A_1 can be split into A_2 and D_2. This process is iterated for each level of A_i. For a total of *J* levels of decomposition, the signal can be represented as a linear combination of the approximation and all levels of detail:

$$S = A_J + \sum_{i=1}^{J} D_i \ . \tag{7.2}$$

In Equation (7.2), the approximation A_J is a low frequency, high-scale (greater *m*) component of the signal, while the $D_i s$ are the high frequency, low-scale components. In the time domain, wavelets are defined by the wavelet function ψ(t) and its complement, the scaling function $\varphi(t)$.The wavelet function is, in effect, a bandpass filter, which represents the details of each scale. The scaling function is a lowpass filter, which filters the lowest level of the transform and ensures that the whole spectrum is covered. From the concept of Equation (7.2), using the wavelet functions and scaling function as bases, a time domain signal *f*(*t*) has the expression:

$$f(t) = \sum_n c_n^J 2^{-\frac{J}{2}} \varphi(2^{-J}t - n) + \sum_{j=-\infty}^{J} d_n^J 2^{-\frac{J}{2}} \psi\left(2^{-j}t - n\right), \tag{7.3}$$

where the scaling coefficients are

$$c_n^j = 2^{-\frac{j}{2}} \int f(t)\, \varphi\left(2^{-j}t - n\right) dt, \tag{7.4}$$

and wavelet coefficients

$$d_n^j = 2^{-\frac{j}{2}} \int f(t)\, \psi\left(2^{-j}t - n\right) dt\,. \tag{7.5}$$

The coefficient of each vector base and is the inner product of the function $f(t)$ and the base wavelet function. The first term in Equation (7.3) is the approximation component of the signal at scale *J*; the second term is the detail component, which is the sum of all details at all levels $j \leq J$. A great benefit that wavelet domain presentation has over Fourier domain presentation is it provides information of changes in the signal in different time frames. In the frequency domain, we can only tell the strength of different frequency components, but the information of the relationship between frequency and time is not given.

For a two-dimensional image, there are three detail functions for each level ***i***: D_i^H, D_i^V, and D_i^D, representing the details in horizontal, vertical and diagonal directions, respectively. The process of splitting the image or signal into multiple levels of wavelet representations is called wavelet decomposition. A decomposition method that gives fuller scope of the image is wavelet packet analysis, which decomposes not only the approximations, but the details as well.

One advantage of doing DWT is that it concentrates the energy of the signal into a small number of coefficients, and the low SNR part of the signal will be distributed into a large number of coefficients. The idea is that by discarding a larger number of coefficients belonging to low SNR, noise is removed. By performing the inverse transform, a denoised image can be obtained in the spatial domain. This noise reduction method is called wavelet shrinkage, which was first proposed by Donoho *et al.* [64] The process that discards coefficients smaller than a certain threshold is called thresholding.

Thresholding can be hard or soft. Hard thresholding sets any coefficient less than or equal to the threshold to zero. That is, only large observations are retained, and therefore, hard thresholding will give a cruder result. Soft thresholding subtracts the threshold from coefficients that are greater than the threshold, and sets coefficients less than the threshold to zero. By shrinking the coefficients, soft thresholding gives a more continuous function of the data. Thresholding is expressed as:

$$\eta_H(w,t) = \begin{cases} w & |w| \geq t \\ 0 & |w| < t \end{cases} \tag{7.6}$$

$$\eta_S(w,t) = \begin{cases} w-t & w \geq t \\ 0 & |w| < t \\ w+t & w \leq -t \end{cases}, \tag{7.7}$$

where η_H and η_S are thresholded wavelet coefficients after hard and soft thresholding respectively; *w* is the original wavelet coefficient and *t* is the threshold.

There are several ways to determine the threshold, such as VisuShrink, RiskShrink, and SureShrink[64-66]. In this Chapter, we use SureShrink. Stein's

unbiased risk estimate (SURE) is an unbiased estimator of mean square error (MSE). It guarantees that the MSE will always be less than $\sigma\sqrt{2\log n/n}$ for a data set of length n with variance σ^2 . That limit of mean square error is used as the threshold of the wavelet coefficients. One advantage of using SureShrink is that it is an automatic procedure [65-66] to determine the threshold for a wavelet at each level j, instead of using a single threshold for all levels of wavelets.

7.1.2 Wavelet Fusion

Image fusion refers to the process of combining two or more images into a single image. The images are usually the same scene acquired from different instrument modalities or captured by different techniques. By combining different features into a single image, image quality can be enhanced or analysis can be made more easily. Image fusion has important applications in biomedical imaging, remote sensing, microscopic imaging, security vision systems, and computer vision.

The principle of image fusion using wavelets is to merge the wavelet decompositions of the two original images using fusion methods applied to approximation coefficients and detail coefficients. The two images must be of the same size and are supposed to be associated with indexed images on a common colormap. For images of two different sizes, the images must be resized and properly registered before fusion.

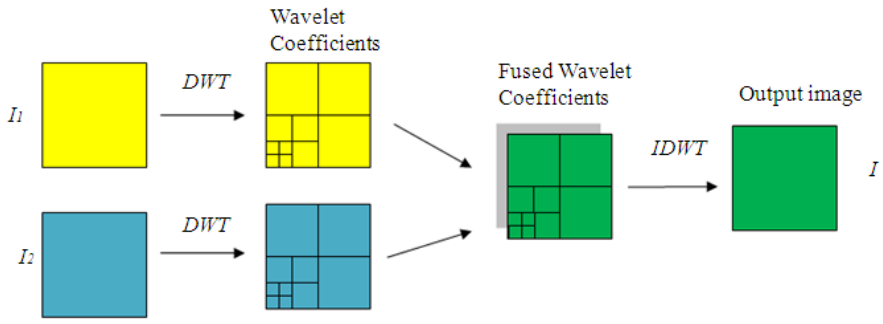


Figure 7.2 Wavelet fusion system overview

The process of wavelet fusion is depicted in Figure 7.2. Two registered images I_1 and I_2 undergo the discrete wavelet transform. The wavelet coefficients are fused by a predefined fusion rule. The inverse discrete wavelet transform is taken on the fused wavelet coefficients.

The fusion rule is chosen based on the purpose of the fusion. Several simple rules include maximum, minimum, mean and random. Maximum rule takes the maximum of the two coefficients; conversely, the minimum rule chooses the minimum. Mean takes the average of the two coefficients. Random rule randomly picks one of the coefficients. The rules applied to approximate and detail coefficients do not need to be the same. Intuitively, applying the maximum rule to both details and approximations can obtain an enhanced image. Beside the simple rules, more complex rules have been proposed. For example, Burt and Kolczynski [61] proposed a "weighted average" method in which combination of coefficients is done adaptively. At each position, weights are assigned to the source images, and the combined result is the weighted sum of the coefficients. In addition to fusing the absolute value of the wavelet coefficients, Hill *et al* proposed a method for fusing the complex coefficients. It is reported in [62] that the image quality is enhanced by such approach.

Chen *et al* in [25] proposed applying wavelet fusion to airport security dual energy imaging. It is claimed that the detection rate can be significantly improved. However, the threat detection method is not provided in the paper and how the wavelet fusion helps the decision making is also not demonstrated. It is also not clear what kind of fusion rule is applied in their work. It is, nevertheless, demonstrated in their work that wavelet fusion can enhance the quality of X-ray images.

7.2 Data and Methodology

Image creation

We still use a 9 to 6 MeV dual energy system. Two equivalent thickness images are created corresponding to high- and low-energy X-rays, respectively. The intensity

of the images is a function of X-ray attenuation. The total attenuation of an object is expressed by its copper equivalent thickness.

Wavelet shrinkage denoising on X-ray images

Two kinds of information are needed for a DE system: a regular X-ray image of good quality and contrast for visual inspection, and a ratio image that gives information for automated threat detection. As shown in Figure 7.1, the edge artifacts on the ratio image can cause misclassification along the edge. The problem is harder to solve when the ring is wider than one pixel. A lowpass filter, such as a Gaussian filter, may be able to smooth the edge, but it would smear and erase the target, especially when the target signal is weak. A median filter is also good at eliminating the edge effect when the edge is only one-pixel wide. If the edge is thicker than 2 pixels, a median filter that is able to eliminate the edge effect will also reduce the strength of the signal. Wavelet shrinkage denoising is performed on both high- and low-energy X-ray images. The denoised images are used to create the ratio image. In this case, wavelet “D6” or “db6” from Daubechies family is used. The highest decomposition level is 3. Using a high decomposition level does not yield different visual results. Soft thresholding is chosen since it gives a smoother, finer result when SureShrink is used.

Ratio image

The ratio image is crucial to the performance of a DE system. In our model, a desirable ratio image must be limited to an appropriate level of detail. The objective here is to create a smooth ratio image that provides the average ratio information for each region in the image. In other words, similar ratios will cluster in a region belonging to the same object. A blurry but smooth ratio image is preferable to the images that keep the extreme values from the edge artifacts. The denoised X-ray images used to create the ratio image will not be used for visual inspection. The ratio in free air will have extreme values since the copper equivalent thickness of air is extremely small. The ratio of free air could end up being from thousands to infinity. To avoid this problem, pixels with thicknesses thinner than one-tenth of an inch are set to be zero on the ratio image. Since the thickness of the container is more than the 0.1-

inch cutoff thickness, ignoring the thin air regions will not have any negative effect on the result.

Probability image

Using the property that the same attenuation caused by a different material will have a different ratio, a curve of the attenuation of high-Z object versus ratio is created. A minimum size lead cube (100 cc) is placed behind different thicknesses of steel, the total attenuation caused by both steel and lead and the effective ratio are recorded. This data is used as a training set for identifying pixels with similar thickness and ratio. Once the relationship between the observed attenuations and ratios is analyzed, a curve of ratio versus attenuation can be obtained by interpolation. An example of such curve is shown in previous chapter, Figure 6.2. If the same attenuation is caused by an object with an atomic number higher than that of lead, its ratio value will fall above the curve generated by the training set. In other words, the curve from the training set is the minimum ratio requirement for a pixel with observed attenuation to be classified as high-Z material. Two hypotheses are used here: $\boldsymbol{H}_0$ and $\boldsymbol{H}_1$. The former represents the class of non-high Z, the latter represents the high-Z class. For a pixel with 9 MeV X-ray attenuation ${T_H}^{eq} = T$, and ratio R, the probability of that pixel belonging to class $\boldsymbol{H}_1$ is

$$P_1 = P(H_1|T) = 1 - 0.5\left[1 + \text{erf}\left(\frac{R_{min} - R}{\sqrt{2}\,\sigma_R}\right)\right] \tag{7.8}$$

where R_{min} is the minimum ratio requirement for the observed ${T_H}^{eq}$ to be high Z. This probability is based on the assumption that the noise can be modeled as Gaussian. The standard deviation of the ratio σ_R is obtained by

$$\sigma_{\mathrm{R}} = R\sqrt{\left(\frac{\sigma_H}{{T_H}^{eq}}\right)^2 + \left(\frac{\sigma_{\mathrm{L}}}{{T_L}^{eq}}\right)^2}\,. \tag{7.9}$$

where ${T_H}^{eq}$ and ${T_L}^{eq}$ are the attenuations represented in equivalent thickness for high- and low-energy X-ray pulses, respectively. The standard deviations of high- and low-energy thickness $(\sigma_H, \sigma_{\mathrm{L}})$ are derived from the calibration data, taking into account the

number of pulses used to form each pixel. The probability of hypothesis $\boldsymbol{H}_{\circ}$ when observing attenuation $\boldsymbol{T}$ is therefore $P(H_0|T) = 1 - P(H_1|T)$.

A high-Z probability image shows the probability of each pixel having a ratio greater than its minimum ratio threshold. The concept of computing probability in Equation (7.8) is different than the probability $P_{high\ Z}$ calculated in Chapter 4 where a global minimum threshold ratio is used. In that case, the computed probability is independent of the observed attenuation. The decision is solely based on the observed ratio. From the discussions in Chapter 6, it is shown that the lookup table method can produce a more accurate prediction. Moreover, the prediction can be more accurate when information of the object's background is provided. For two scenarios that ω_1 and ω_2 represent the background being mid-Z and low-Z respectively, $P(H_1|T, \omega_1) > P(H_1|T, \omega_2)$because the curve is based on mid-Z shielded objects. $P(H_1|T, \omega_2)$ is obtained through layer separation, and layer separation will introduce errors to the calculation of probability.

Fusion, clustering and masking

We apply hybrid clustering to the wavelet-fused image. Fusing the high- and low-energy images can provide a smoother image while preserving sharp edges. The wavelet used here is still "D6." The wavelet level chosen here is 3. The fusion rule for both approximate and detail wavelets is "mean", meaning the output is the average of the two coefficients. The fusion rule "mean" is chosen over "maximum" because "maximum" does not give an improved result over the original two X-ray images. For pixels belonging to mid-Z and low-Z regions, the low-energy image will have greater equivalent thickness, and the opposite is true for pixels of high Z regions. The maximum rule will just replace pixels of low-z and mid-Z regions in the high-energy image with the values of their low-energy image counterparts. In this case, the contrast is actually depressed and nothing gets smoothed.

Hybrid clustering is an image segmentation method used to group pixels within the same intensity level. Contiguous pixels with similar intensities will form a region or cluster. Ideally, each cluster represents either an object or the background. A region is used as a mask on the probability image. Regions with only insignificant

percentage of high probability pixels will not be considered high-Z regions. Masking is on a region-by-region basis; if a significant amount of pixels are classified as high-Z pixels, the whole region causes an alarm.

Decision making

The algorithm for high-Z detection is summarized in the flow chart in Figure 7.3. High energy and low energy images are both processed by wavelet shrinkage denoising. For the high-energy image, the minimum ratio for each pixel is calculated based on the previous experimental data. The two denoised images are used to obtain the ratio image and, subsequently, the probability image. In parallel, the unprocessed X-ray images are fused by wavelet fusion to obtain a smoother image for clustering. Clustering is used to map the probability image for decision making. Two kinds of decision are being made. One is for pixels, and the other is for the whole cluster region. A pixel is a high-Z pixel if $P_1 \geq 0.5$. For each region, if more than 30 percent of its pixels are classified as high Z, the whole region will be a high-Z region. The 30 percent criterion is chosen based on experiments. For strong signals, it is very likely the case that 100 percent of the pixels in a region are high-Z pixels. For a heavily shielded object, only about 30 percent can pass the minimum ratio requirement.

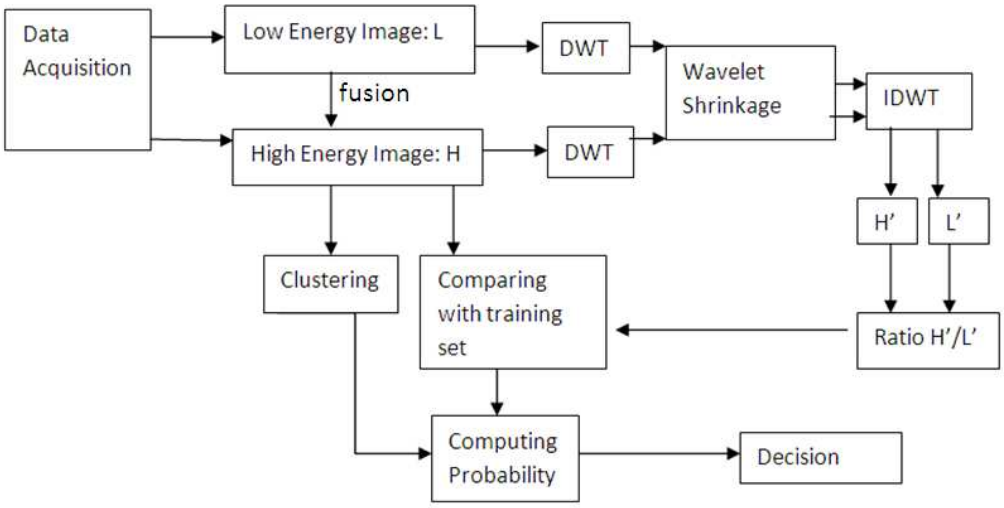


Figure 7.3 Flowchart of the wavelet-based DE imaging system. (IDWT=inverse discrete wavelet transform)

7.3 Experimental Results

Test for unshielded objects discrimination

The photos in Figure 7.4 show the setup for testing the performance in differentiating non-shielded objects. From left to right, the objects are a 400 cc lead cube, a 100 cc tungsten cube, a 100 cc lead cube, 150 cc and 100 cc of tin, and 150 cc of steel. Only the first three objects are high-Z materials. The X-ray images are unprocessed (a.2) and wavelet-denoised (a.3) high-energy X-ray images, respectively. Edge artifacts can be seen in those X-ray images. The wavelet-denoised X-ray image is blurred after processing. The three images in column (b) are the unprocessed ratio image (b.1), the thin-air, background-removed ratio image (b.2), and thin-air, background-removed ratio image obtained from wavelet-denoised X-ray images (b.3). The top image shows that the ratios of air are extremely high. Those unreasonably high values are mostly caused by random scattering. We force the ratios in regions thinner than a half-inch of copper to be zero to obtain the bottom two images. A salient difference between the processed (b.3) and unprocessed (b.2) ratio maps is that the image (b.3) is much smoother, especially on the edges of the objects. The benefits of applying wavelet denoising can be better illustrated in Figure 7.4(c). Figure 7.4 (c.1) shows the high-Z probability image obtained from the ratios in Figure 7.4(b.2). The image shown in Figure 7.4(c.2) is the high-Z probability image obtained from the wavelet-denoised X-ray images. All objects are unshielded, and the penetration is good, the signal is strong and the uncertainty of ratio is low. As a result, most of the pixels have either very high probability or zero probability. In Figure 7.4 (c.1), the edge artifacts caused several false positive pixels on the edge of the three non-high-Z objects. The wavelet-denoised image in (c.2) does not show the false positives associated with the edge artifacts.

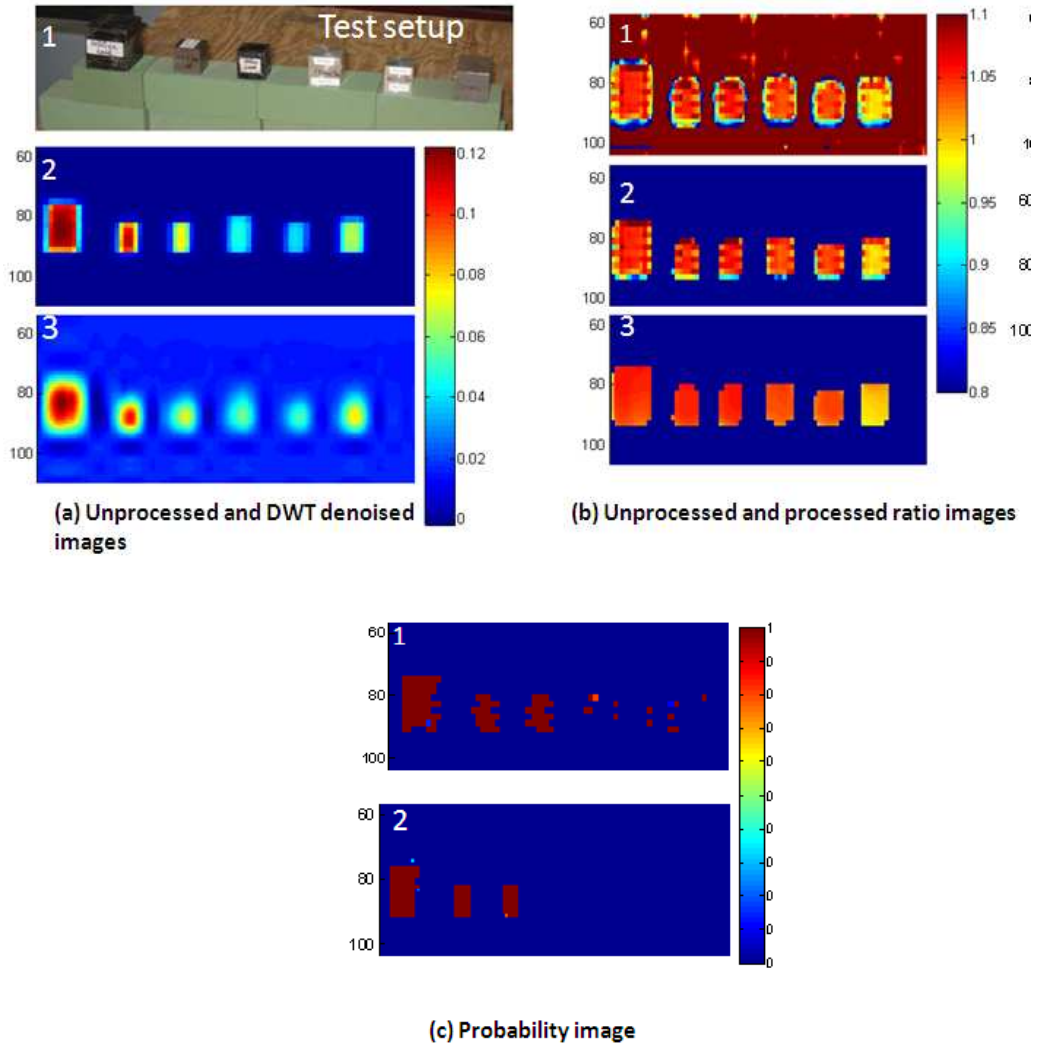


Figure 7.4 (a) Top: photo image of setup. Middle: high-energy X-ray image. Bottom: DWT denoised X-ray image. (b) Top: ratio image. Middle: setting all the pixels in top image to zero if that pixel's location has thickness less than half-inch of steel in top image. Bottom: ratio image of wavelet-denoised image. (c) Top: probability image from unprocessed X-ray images. Bottom: probability image from wavelet-processed X-ray images.

Test for shielded objects discrimination

Figure 7.5 shows the test of shielded object detection (photo of the target setup is shown in Figure 7.4(a). Three objects are placed behind 10 inches of steel. The two on the left are high-Z objects: a 100 cc lead cube and a 100 cc tungsten cube. A 100 cc tin cube is placed next to them for possible false positive tests. Bricks of lead are

positioned on top of the steel plates. Figure 7.5 (a) shows the unprocessed high-energy X-ray image, the experimental configuration. Figure 7.5 (b) is the noisy ratio image. Based on that ratio image, we do not have any confidence about whether any high-Z object is behind the steel. Also seen in this image are the sawtooth-like artifacts on the edges of the steel. Figure 7.5(c) shows the probability image obtained from the ratio image 7.5(b). The probability image is noisy. The edge artifacts caused the edges of steel plates to appear to have high probability.

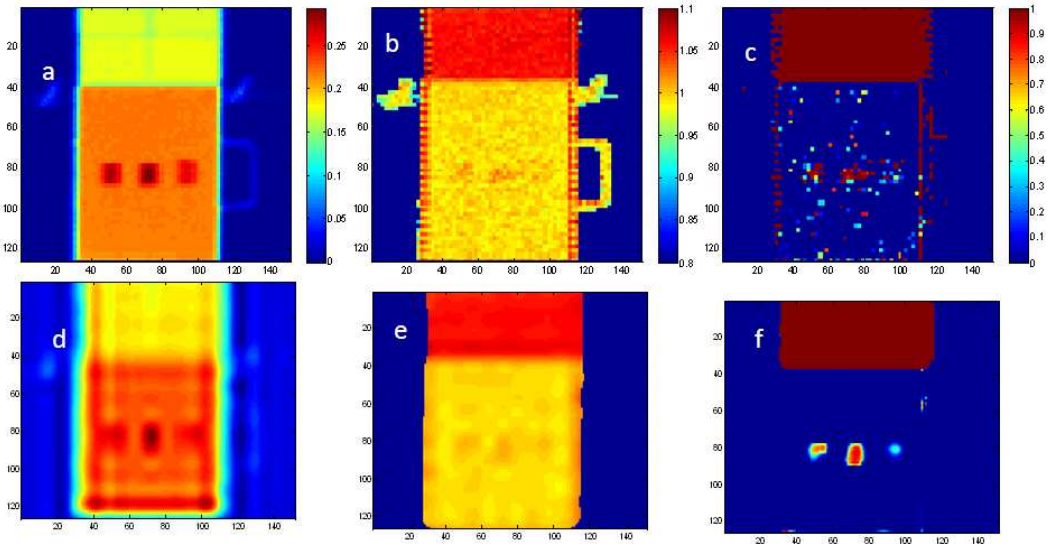


Figure 7.5 (a) high-energy X-ray image of test configuration shown in Figure 7.7 (a). (b) Ratio image (c) probability image. (d) Wavelet-filtered high-energy X-ray image. (e) Ratio image of wavelet-denoised images. (f) Probability image.

Figure 7.5 (d) is the wavelet-denoised high-energy X-ray image. It is too blurry for X-ray inspection. However, this image is not produced for visual inspection but for material discrimination computation. Fig. 7.5 (e) shows the ratio image of the two wavelet-denoised X-ray scans. The sawtooth effect is smoothed by the wavelet denoising. Figure 7.5(f) is the probability image for the filtered ratio image. A significant difference between the probability map (f) and the probability map (c) is that not only does the image (f) appear to be cleaner, but also that all the regions of non-high-Z objects do not produce high-probability pixels. Because they are not

shielded by any non-high-Z object, the regions of the lead bricks appear to have very high probability.

Fusion, clustering and masking

Figure7.6(c) shows the result of wavelet fusion. Comparing this result with the original, unprocessed high-energy X-ray image shown in Fig. 7.6(b), the steel plate regions are now smoother. This is the enhanced image that can be used for visual inspection, together with the result of high-Z detection. A clustering algorithm is applied on this fused image. Neighboring pixels with differences within 1.25 inches of copper equivalent thickness will be grouped together. Figure 7.6(d) is the clustered fusion image. All subimages (regions) will be labeled, and each subimage will be used as a mask for decision making.

Figure 7.7 illustrates the steps of masking. The first stage shows that the high-Z image will be masked by the cluster map on the bottom for decision making. The cluster map is a label image that assigns an integer to each distinct region. Each color in the cluster represents a number. The result of masking is shown in the second step. The shape of the high-Z map is now in accordance with the shapes of the clusters. Pixels outside a high-Z cluster will be discarded, while the original non-high-Z pixels in a high-Z cluster will be filled in. Overlaying the output of masking with the wavelet fusion-enhanced image gives the final output. This overlay image not only tells if there is an alarm, but also tells the user what object to look for.

Comparing Wavelet De-noising with the Wiener filter

Figure 7.8 shows two probability images obtained from a pair of the Wiener filter-processed X-ray images (left) and from a pair of wavelet-denoised images (right). The image in the middle is the result of applying a 5 by 5 median filter on the first probability image. One high-Z object is shielded by 10 inches of steel. Both cases do not have a very strong signal of high Z because of low X-ray penetration. The result of the Wiener filter shows that edge effects still exist after the filtering. When the median filter is applied to reduce the edge effects, the signal from the high-Z object is also weakened. Its wavelet-denoised counterpart, though it has a weaker

signal for the object, has taken care of the edge artifacts by itself. By selecting a suitable wavelet, we can obtain a smoother image, although the signal is weaker.

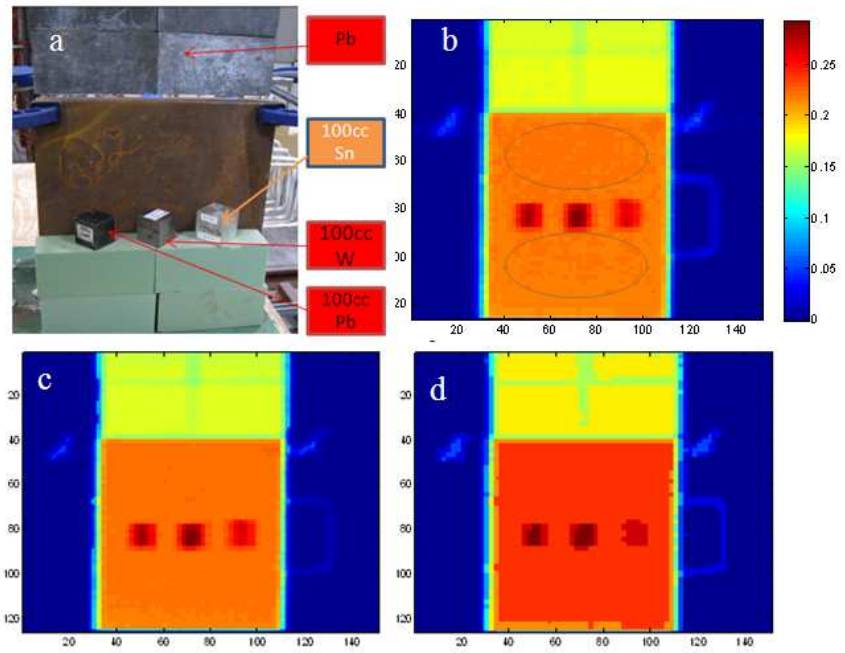


Figure 7.6(a) Experimental setup: Two high-Z objects and two non-high-Z objects are shielded behind 10 inches of steel. (b) The high energy image. (c) The result of wavelet fusion. Comparing this image with (b), the circled steel plate region is much smoother in this image. (d) The clustering result of applying hybrid clustering to image (c).

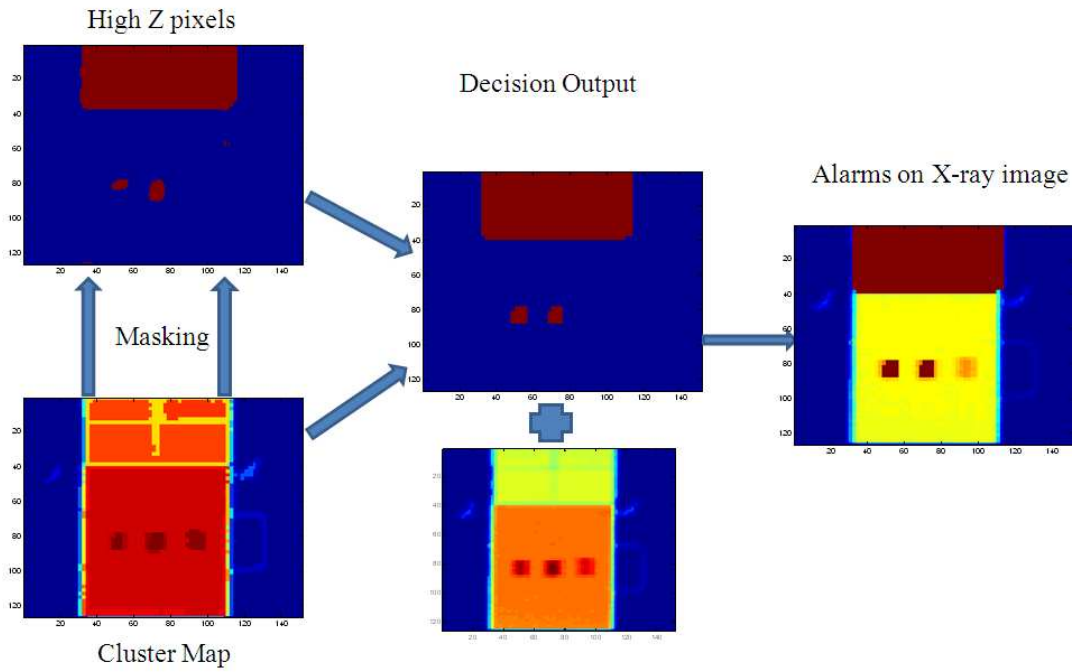


Figure 7.7 First column: cluster map and high-Z map. Second column: the output of masking on the top will be overlaid with the fused image to create a high-Z alarm for inspectors, shown in the last column.

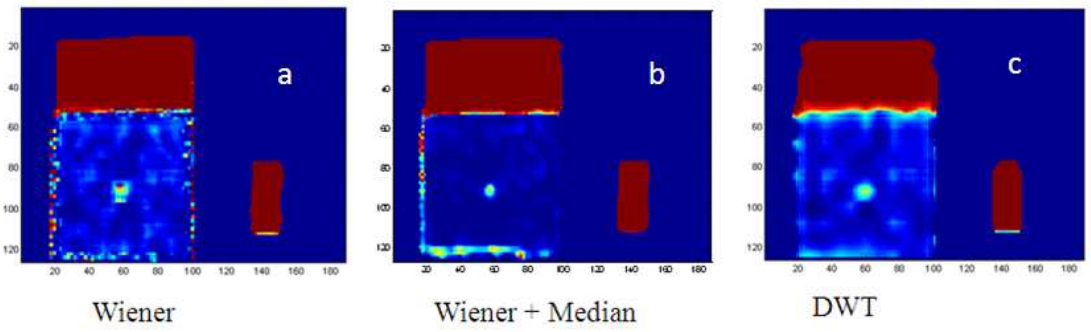


Figure 7.8 (a) Probability image obtained from a pair of Wiener filter-processed thickness images, (b) the median-filtered image of (a), and (c) the probability image from a pair of DWT denoised thickness images.

7.4 Conclusion

In this chapter, we introduced a wavelet–based approach to enhance the performance of a dual energy imaging system. In reducing the amount of Poisson noise, the false negative rate has significantly dropped. Also, the false positives caused by edge effects can be effectively removed. By selecting a proper wavelet, we found that wavelet domain noise reduction can be a very effective alternative for the conventional Gaussian filter or Wiener filter.

Our algorithm in this chapter is a dual-action approach. We separate the tasks of enhancing image quality and the detection rate. We found that we do not need to enhance the detection rate by enhancing the quality of X-ray images to keep the edge details. We choose a wavelet that creates blurrier but smoother images after wavelet denoising. This method can eliminate the false alarms caused by artifacts.

A similar result can also be obtained from a Wiener filter. While the Wiener filtering on X-ray images can reduce the X-ray image noise and improve the detection rate, it often needs to be combined with morphological filtering or median filtering on the probability image to eliminate the edge effects that can lead to the elimination of already weak object signal due to heavy shielding or low X-ray penetration quality.

By taking advantage of having two images in a DE system, we use wavelet fusion to obtain a denoised image, which can later be used for image segmentation. There is no meaningful difference between a Wiener filtered X-ray image and wavelet fused image. Our results show that there is a potential for applying a wavelet-based approach in the application of dual energy imaging.

Chapter 7, in part, is adapted from the material as it appears in the Proceedings of SPIE/IS&T (2010), Fu, Kenneth; Ranta, Dale; Guest, Clark; Das, Pankaj “The application of wavelet denoising in material discrimination system.” The author of this dissertation is the primary investigator and author of this paper.

8

Conclusion

The material discrimination function of cargo imaging systems is a vital technology for national security. The ability to detect contraband material hidden inside a cluttered cargo container can prevent tragedies from happening. Although the theory of using dual energy systems for material discrimination is well developed, the technology to build a practical system has still not fully matured.

The key to realize such an imaging system is digital signal and image processing technology. With help from those fields, the X-ray image can be enhanced, and the detection rate of target materials can be significantly improved. In this thesis, we have demonstrated several novel methods for improving the performance of such systems. All those methods have given promising results. However, more precise algorithms are still needed to achieve an optimal detection rate.

8.1 Achievements

For a single pixel in an X-ray image, the only information we can obtain is the intensity at the point. However, we have no way of telling how many layers of materials are in the path that photons pass through. Therefore, the ratio of absorptions from high and low energy scans cannot be the single indicator for screening out high atomic number materials. The current proposed DE systems can only provide the effective atomic number to the user but not really an automatic threat detection system. We have provided solutions for this inadequacy.

The Poisson noise from the X-ray system will add uncertainties to the measured ratio. A model to calculate the probability of a pixel being high Z is proposed. Based on that model, we have developed several algorithms to enhance the signal and reduce the false detection rate.

An adaptive masking model that uses a modified automatic thresholding technique is a good approach to detect high Z hidden materials in a small region. Using such an approach, a threshold to determine the probability of triggering a threat alarm is not needed, and at the same time, the false alarm rate is greatly reduced.

For more complicated scenarios, we need an algorithm that can do a more precise job of region finding. We have developed a new image segmentation approach that has the advantages associated with both *k*-means clustering and seed growing, but without encountering the disadvantages associated with them. An approach that uses region filling to detect shielded objects also has been developed. The detected objects can be used as masks for the adaptive masking algorithm. We can then use those objects for further region analysis.

Different from the adaptive masking approach that finds high atomic number materials on a region by region basis, a pixel by pixel analysis algorithm is also proposed. For a single pixel for which an equivalent thickness is calculated, the theoretical minimum ratio of high- and low-energy scans can be calculated. This minimum value is the theoretical threshold that a pixel has to pass to be called high Z. False alarms are eliminated if the number of high Z pixels within a region is less than a preset threshold.

One significant improvement to the system is the layer separation approach. It provides additional information for making judgments. Although the method has not been perfected yet due to the limited data we have, the experimental results show a promising future of this approach. Future work will focus on perfecting this roach.

Lastly, the applicability of wavelet-based methods has been tested. Ours is one of the few works that apply wavelets to a security inspection system, especially in

dual-energy imaging. We have shown that wavelet denoising and fusion can be a good substitute of other traditional approaches.

8.2 Future work

We have not exploited all possible ways to improve the performance of the system. The first, and probably the most important task is noise reduction of the X-ray image. In addition to the Wiener filter and wavelet denoising, there are still different approaches such as neural network or independent component analysis (ICA). The approach of adaptive wavelet coefficient thresholding based on regional statistical properties is also worth trying.

For shielded objects, layer subtraction is a good way to reconstruct the ratio of absorptions of objects in each layer. However, the 3-D information of the container will probably provide a more accurate result. For two-dimensional images, the best way to do layer subtraction is to subtract the thickness of the background from the thickness of the foreground. That relies on a good method of image clustering, which still has its limits. In addition to the region growing-based approach for image segmentation, we believe that a combination of several different approaches will give us an optimal performance.

The data we have is from a system that still has room for improvements. The improvement of hardware includes control of scattering, stabilizing the system, and enhancing the penetration quality. Once the hardware side is improved, more sophisticated algorithms can be developed.

8.3 Closing Remarks

The technology of dual-energy imaging for cargo material discrimination is still in its fledgling state. We are doing the pioneering work of finding the solutions for some fundamental problems. Still, there are challenges waiting to be met.

On the Christmas Eve of 2009, Umar Farouk Abdulmutallab made the headlines. He successfully passed the security check point at the Schiphol Airport, Amsterdam by sewing the explosive material known as PETN (pentaerythritol tetranitrate, or pentrite) in his underwear, attempting to detonate the bomb device on Northwest Airlines' Flight 253. The stories such as Mr. Abdulmutallab's have demonstrated the importance of a material discrimination system. Although our work in this thesis was based on detecting radioactive materials, similar ideas can be applied to explosive detection for airport security as well.

We hope our work can provide some ideas to readers who are also working in this field and want to help further fortify our national security. We are confident that the current challenges will be solved in the near future and this technology will keep evolving in advance of any terrorist strategies.

Appendix

Useful Matlab Functions

A. For Noise Reduction

A1. Wiener filter

The "wiener2" function in Matlab was implemented according to Eq.(3.24) to Eq.(3.27). We use that for high and low energy images noise reduction. Image I is filtered by an M- by N Wiener filter by typing

```
J = wiener2(I,[M N]) ;
```

The default values for M and N are 3 and 3. The high and low energy images denoted H and L respectively and their filtered version H1 and L1 is obtained by

```
H1 = wiener2(H) ;

L1 = wiener2(L) ;
```

Ratio image R is then H1./L1. The operation "./" means point by point division.

A2. Median filter

The function for median filter in Matlab is " medfilt2." The output image J is obtained by

```
J = medfilt2(I,[M N]) ;
```

where I is the input image and M and N defined the size of the filter window.

A3. Wavelet shrinkage

The wavelet denoising in this thesis was done by the following way:

First, compute the default values for denoising by using the command "ddencmp"

```
[thr,sorh,keepapp,crit] = ddencmp('den','wp',x);
```

The first parameter ‘den’ indicates the operation we want to do is de-noising. The second parameter is the wavelet approach. By selecting ‘wp’in stead of ‘wv’, we are using the wavelet packet, in which both approximate and detailed are used for decomposition. The third argument x is the image. The output of the function will return the global threshold ‘thr’, the method of thresholding, “sorh” =”s” for soft threshodling and “h” for hard thresholding respectively. The output keepapp will be either 1 or 0. When it equals one, the approximation will be kept. This is the case when wavelet package was chosen. The last output “crit”=”sure” if “wv” is used and equals “threshold” when “wp” is selected.

Then, the denoising is done by using the wpdencmp command, which is a function for wavelet packet denoising and compression. Using the parameters obtained above and other inputs, we have

```
xd = wpdencmp(x,sorh,level,wave,crit,thr,keepapp);
```

where x is the image, level is the level of decomposition, wave is the wavelet use. In our case, level=3, and wave=’db5’.

B. Morphological operations for masking

The matlab command for dilation, erosion, opening, and closing are “imdilate”, “imerode”, ‘imopen” and “imclose”. They all have the same structure. First, a structuring element needs to be defined. In the case that a 3 by 3 square structuring element is used for those operations, we have a structuring element str to be

```
str=strel(“square”, 3) ;
```

Then, a dilation operation on an image I will be achieved by

```
J=imdilate(I, str);
```

It is the same thing for all other morphological operations. To find the background of an object, we can simply use the following method:

```
[L M]=bwlabel(I);
mask=zeros(size(L));
mask(L==i)=1;
mask1=imdilate(mask,str);
background=mask1-mask;
```

here, on a binary image I, each distinct region is assigned a label by using command "bwlabel", the value of the labels are stored in the array L, and output M is a scalar, giving the total number of regions. A new image "mask" is initiated by setting all of its pixels to zero. To dilate the object i, we assign a value 1 to mask on the pixels that L equals i. After that, its dilated version mask1 is created by using the imdilate function. The extended area of the object is region of the object's background. By having the binary image of the background region, we are able to analyze the properties of the background, i.e. ratio, thickness etc. For example, the background thickness of object number 5 will be

```
[L M]=bwlabel(I);
mask=zeros(size(L));
mask(L==5)=1;
mask1=imdilate(mask,str);
background=mask1-mask;
BG_thickness= background.*High_energy;
BG_thickness2= background.*Low_energy;
Ave_ BG_thickness= mean(nonzero(BG_thickness));
Ave_ BG_thickness2= mean(nonzero(BG_thickness2));
```

Therefore, layer separation can be achieved by

```
New_object_thickness=mask.*High_energy-Ave_ BG_thickness;
```

In the line above, mask.*High_energy gives a thickness image that provides only the information on the object region, which is defined by mask. Then each pixel in the image is subtracted by an value, which is the average background thickness we

obtained from the previous line above. Similarly, the adjusted low-energy thickness will be

```
New_object_thickness2=mask.*Low_energy-Ave_ BG_thickness;
```

And the new ratio in that region is obtained by

```
R_new= New_object_thickness./New_object_thickness2.
```

C. Clustering algorithm for segmentation

The following code is used to perform the clustering method in chapter 5. Three inputs are needed: radius of the filter window “rd,” high energy image being filtered “hi”, and step size “step.” Five filters q1 to q4 and Q are defined. Output is I.

```
for i=rd+1:r-rd
    for j=rd+1:c-rd
        q1=hi(i:i+rd,j-rd:j);
        q1v=mean(nonzeros(q1))*100/2.54/step;
        q2=hi(i-rd:i,j-rd:j);
        q2v=mean(nonzeros(q2))*100/2.54/step;
        q3=hi(i-rd:i,j:j+rd);
        q3v=mean(nonzeros(q3))*100/2.54/step;
        q4=hi(i:i+rd,j:j+rd);
        q4v=mean(nonzeros(q4))*100/2.54/step;
        qv=[q1v q2v q3v q4v];
        Q=hi(i-rd:i+rd,j-rd:j+rd);
        Qv=round(mean(nonzeros(Q))*100/2.54/step);
           if .9<hi(i,j)*100/2.54/step/Qv &
hi(i,j)*100/2.54/step/Qv<1.1
                     temp(i,j)=Qv;
           else
                  dif=abs(hi(i,j)*100/2.54/step-qv);
                  [y,I]=min(dif);
                  temp(i,j)=round(qv(I));
           end
    end
end
I=temp; I=I+qb;
```

Bibliography

[1] Chen, G., "Understanding X-ray cargo imaging," Nuclear Instruments and Methods in Physics Research, vol. 241, pp.810-815 (2005)

[2] Rebuffel, "Dual energy X-ray Benefit and limits," Insight - Non-Destructive Testing and Condition Monitoring, vol. 49, Issue 10, 589-594, (2007)

[3] Novikov, V.L., Ogorodnikov, S.A., Petruni, V.I., "Dual Energy Method of Material Recognition in High Energy Introscopy System," International Workshop on Charged Particle Linear Accelerators: Problems of Atomic Science and Technology, 93-95, 1999

[4] Ogorodnikov, S., Petrunin, V., "Application of High-penetrating Introscopy System for Recognition of Materials," Proc. EPAC, 2583-2585, (2000)

[5] Ogorodnikov, S., Petrunin, V., "Processing of interlaced images in 4–10 MeV Dual-energy customs system for material recognition". Phys. Rev. ST Accel. Beams 5, pp. 104701-1-104701-11 (2002)

[6] Ogorodnikov, S., Petrunin, V., "Processing of interlaced images in 4–10 MeV Dual-energy customs system for material recognition". Technical Report, Not published.

[7] Ogorodnikov, S., Petrunin, V., Vorogushin, M., "Radioscopic Discrimination of Materials in 1÷10 MeV Range For Customs Applications." Proc. EPAC, pp. 2807-2809, (2002)

[8] Tomasi, C., and Manduchi, R., "Bilateral Filtering for Gray and Color Images," Proceedings of IEEE International Conference on Computer Vision, Bombay, India, (1998). http://users.soe.ucsc.edu/~manduchi/Papers/ICCV98.pdf

[9] Devarajan, H., and Nyikal, H., "Bilateral Filters," Image Scaling and Bilateral Filtering 2006 course.
http://scien.stanford.edu/class/psych221/projects/06/imagescaling/

[10] Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Introduction to Data Mining, http://www-users.cs.umn.edu/~kumar/dmbook/ch8.pdf

[11] P.A. Vijaya, M. Narasimha Murty 1, D.K. Subramanian, "Leaders–Subleaders: An efficient hierarchical clustering algorithm for large data sets," Pattern Recognition Letters 25, pp. 505–513, (2004)

[12] Bjorkholm, P.J., "Detection of weapons of mass destruction," Proceedings of SPIE, Vol. 5408, pp. 46-51, (2003)

[13] Bjorkholm, P.J., “Dual energy detection of weapons of mass destruction,” Port Technology International, June/July, 22-26, (2004)

[14] Bjorkholm, P.J., “WMD detection “, Cargo Security International, June/July, 23-25, (2005)

[15] Bjorkholm, P.J., “Detection of weapons of mass destruction,” Port Technology International, PT 18 -5/1.
http://www.varianmedicalsystems.com/media/security_and_inspection/resources/articles/pdf/Detection_of_WMD.pdf

[16] Bjorkholm, P.J.,U.S. Patent No. 6,069,936 (2000)

[17] Bjorkholm, P.J., “The X files part 1,” Cargo Screening International, pp 49-50, June (2004)

[18] Bjorkholm, P.J., “The X files part 2,” Cargo Screening International, pp 48-49, August (2004)

[19] Bjorkholm, P.J., “The X files part 3,” Cargo Screening International, pp. 58-59, October (2004)

[20] Bjorkholm, P.J., “The economics of cargo screening,” Port Technology International, 31-39_2.
https://electronics.wesrch.com/User_images/Pdf/GP9_1197150276.pdf

[21] Bjorkholm, P. J., and Wang, T. R. “Explosives detection using three dimensional computer assisted image analysis”, SPIE Vol. 1824, pp. 122-126, 1992.

[22] Reed, W., “Implementing X-ray cargo screening,” Cargo Security International, pp.106-108, June/July (2005)

[23] Macdonald, R., “Design and implementation of a dual-energy X-ray imaging system for organic material detection in an airport security application” Proc. SPIE, Machine Vision Applications in Industrial Inspection IX, Proceeding Vol. 4301, pp.31-41 (2001)

[24] Auditore, R., “Going towards the Dual Energy X-ray Radiographic System for Material Recognition Purposes,” Proc. of PAC07, pp. 2754-2756, (2007)

[25] Chen, Z., Zheng, Y., Abidi, B.R., Page, D.L., and Abidi., M., “A Combinational Approach to the Fusion, De-noising and Enhancement of Dual-Energy X-ray Luggage Images,” Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05), Vol. 19, 1063-69 (2005)

[26] Zhang, G, Chen, Z., and Zhang, Li., "An H-L curve method for Material Discrimination of dual energy X-ray inspection systems," 2005 IEEE Nuclear Science Symposium Conference Record, Vol. 1, pp. 326-328 (2005).

[27] Zhang, G., Zhang, L., and Chen, Z., "Exact Reconstruction for Dual Energy Computed Tomography using an H-L Curve Method" 2006 IEEE Nuclear Science Symposium Conference Record, M14-462, pp. 3485-88 (2006)

[28]Wang, X.W., Li, J., Kang, K., Tang, C., Li, Z., Chen, Z., Li, Y., and Zhong, H., "Material Discrimination by high energy X-ray dual imaging," High energy Physics and Nuclear Physics, Vol. 31, No. 11, Nov. (2007)

[29] Wang, T.W., and Evans, J.P.O., "Stereoscopic dual energy imaging for target materials identification," IEE Proc. of Vision and Image Signal Processing, Vol. 150. No. 2, pp.122-130 (2003)

[30] Liang, J., Abidi, B., and Abidi, M., "Automatic X-ray image segmentation for threat detection," Proc. of Fifth International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2003), pp. 396 - 401 (2003)

[31] Paranjape, R., "Segmentation of handguns in dual energy X-ray imagery of passenger carry-on baggage," IEEE Canadian Conference on Electrical and Computer Engineering, vol.1, pp 377 – 380 (1998)

[32] Singh., S., "Explosives detection systems (EDS) for aviation security," EURASIP Signal Processing, vol. 83, 31 – 55 (2003)

[33] Sukovic, P. and Clinthorne, N.H., "Design of an experimental system for dual energy X-ray CT," Conference Record IEEE Nuclear Science Symposium, vol.2, pp. 1021 - 1022 (1999)

[34] Sukovic, P. and Clinthorne, N.H., "Penalized Weighted Least-squares Image Reconstruction for Dual Energy X-ray Transmission Tomography," IEEE Transactions on Medical imaging, Vol. 19, No. 11, 1075-1081(2000)

[35] Ying, Z., Nadiu, R., Guilbert, K., Schafer, D., and Crawford, C.R., "Dual Energy Volumetric X-ray Tomographic Sensor for Luggage Screening," IEEE Sensors Applications Symposium, pp. 1-6 (2007).

[36] Chartrand, R., Asaki, T., "Background Radiography for Border Inspectors," Presentation for the Los Alamos National Laboratory Background Radiography Team.

[37] Gonzalez, R., and Woods. R., "Digital Image Processing 2nd Edition," 595-626, Prentice Hall, (2002)

[38] Lim, J.S., "Two-Dimensional Signal and Image Processing," 454-460, Prentice-Hall (1990)

[39] Adams, V., Bischolf, L., “Seed Region Growing”, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 16, No. 6. June (1994)

[40] Criminisi, A., “Region Filling and Object Removal by Exemplar-based Image Inpainting,” IEEE Transactions on Image Processing, Vol. 13, No. 9, pp. 1200-1212, SEP (2004)

[41] Malm, P, “Development of a hierarchical k-selecting clustering algorithm-application to allergy”. Master’s thesis, 30ECTS, Engineering Biology programme, University of Linkoping, Nov. (2007)

[42] Veenman, C., “A Maximum Variance Clustering Algorithm,” IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 24, no. 9, pp. 1273-1280, (2002)

[43] Rui, Y., “Automated region Segmentation Using Attraction-Based Grouping in Spatial-Color-Texture Space,” Proc. International Conference on Image Processing, Vol. 1, pp. 53-56,(1996)

[44] Mat-Isa, N., “Seed Region Growing Features Extraction Algorithm—Its Potential Use in Improving Screening for Cervical Cancer,” International Journal of the Computer, the Internet and Management, vol. 13, pp. 61-70 (2005)

[45] Chen, Y., “A hybrid method of seeded region growing and region hue-area information fusion for object segmentation under patterned background,” 10th international Conference on Information Fusion, pp.1-6 (2007)

[46] Amer,A. “New binary morphological operations for effective low-cost boundary detection,” Internal Journal of Pattern Recognition and Artificial Intelligence, vol. 17, No. 2, pp. 1-13 (2002)

[47] Aach, T., Schiebel, U., and Spekowius, G., “Digital image acquisition and processing in medical X-ray imaging,” SPIE Journal of Electronic Imaging vol.8 (1), pp. 7–22 (1999).

[48] Aach, T., Schiebel, U., and Spekowius, G.,, “A Lapped Directional Transform for Spectral Image Analysis and Its Application to Restoration and Enhancement,” EURASIP Signal Processing, vol. 81, pp. 2347-2364 (2000).

[49] Hakan Öktem , Karen Egiazarian, Jarkko Niittylahti, and Juha Lemmetti, “An approach to adaptive enhancement of diagnostic X-ray images,” EURASIP Journal on Applied Signal Processing, Vol. 2003, pp. 430-436 (2003)

[50] Chartrand, R. “Background Radiography for Border Inspections,” presentation for the Los Alamos National Laboratory Background Radiography Team.

[51] Schultz,LJ, “Image reconstruction and material *Z* discrimination via cosmic ray muon radiography,” *Nuclear Instruments and Methods in Physics Research A* **519**, pp. 687–694 (2004).

[52] Schultz, LJ, “Cosmic Ray Muon Radiography for Contraband Detection” Proceedings of AccApp’03, San Diego, CA, June 2003. http://www.lanl.gov/physics/pdfs/muon2.pdf

[53]Chen, Y., Maitre, M., Fang, T., “Transparent Layer Separation for Dual Energy Imaging,” Proceeding of ICIP, pp. 821-824, 2008.

[54]Aach, T., Kunz, D., “Spectral Estimation Filters for Noise Reduction in X-ray Fluoroscopy Imaging,” Proc. of EUSIPCO-96, Tri-este, pp 10–13, Sept. 1996

[55] Til Aach, Erhardt Barth and Claudia Mayntz, “Motion-Compensated Defect Interpolation for Flat-PanelDetectors,” Proc. Image Processing SPIE Vol. 5370, pp.1352-1363, (2004)

[56] Han, X., Chen, Y., Kitamura, K., Ishikawa, A, “An ICA Based Noise Reduction for PET Reconstructed Images,” Proc. Conference on International Information Hiding and Multimedia Signal Processing, pp.1655-1660, (IIH-MSP 2007)

[57]Li, J., Han, X., Algorithm of ICA-based Poisson-noise reduction and its application to CT image.” Proc. of SPIE Vol. 6789, pp. 908-1 to 908-8, (2007)

[58] Jenneson, P.M., “Large vessel imaging using cosmic-ray munons,” Nuclear Instruments and Methods in Physics Research, A 525, pp. 346-351, (2004)

[59] Bogdan J. Matuszewski, Lik-Kwan Shark, Martin R. Varley, “Region-based wavelet fusion of ultrasonic, radiographic and shearographyc non-destructive testing images,” http://www.ndt.net/article/wcndt00/papers/idn263/idn263.htm

[60] Gang Hong, Yun Zhang, “The Effects of Different Types of Wavelets on Image Fusion,”http://www.isprs.org/congresses/istanbul2004/comm4/papers/474.pdf

[61] P.J. Burt, R.J. Kolczynski, “Enhanced image capture through fusion.” *Proceedings of the 4th International Conference on Computer Vision*, pages 173–182, 1993.

[62] Paul Hill Nishan , Paul Hill , Nishan Canagarajah , Dave Bull, “ Image Fusion using Complex Wavelets “Proc. 13th British Machine Vision Conference (2002)

[63] F. Bektas Balcik, E. Sertel, “Wavelet-Based Image Fusion of Landsat ETM Images: a Case Study for Different Landscape Categories of Istanbul,” www.isprs2007ist.itu.edu.tr/29.pdf

[64] Donoho, D. and Johnstone, I., “Adapting to Unknown Smoothness via Wavelet Shrinkage,” Journal of the American Statistical Association, Vol. 90, pp.1200-28, (1995).

[65] Donoho, D, “Threshold Selection for wavelet Shrinkage of Noisy Data,” Proc. of the 16th Annual International Conference of the IEEE, Engineering in Medicine and Biology Society, pp. 24-25, (1994)

[66] Donoho, D., “De-Noise By Soft-Thresholding,” IEEE Tractions on Information Theory, vol. 41, No 3, pp. 613-627, (1995).

[67] Stephane G. Mallat, “A Theory for Multiresolution Signal Decomposition: The Wavelet Representation,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. II, No. 7. pp 674-693, (1989)

[68] Carl Taswell, “The What, How, and Why of Wavelet Shrinkage Denoising,” Technical Report, Computational Toolsmiths. < www.toolsmiths.com >

[69]L. Birg´e and P. Massart, “From model selection to adaptive estimation,” Research Papers in Probability and Statistics: Festschrift for Lucien Le Cam, D. Pollard, E. Torgersen, and G. Yang, Eds. New York: Springer-Verlag, pp. 55-87, (1996).

[70] G. Strang and T. Nguyen, “Wavelets and filter banks” Wellesley-Cambridge Press, (1996)

[71] Wink, A.M. and Roerdink, B.T.M., “Denoising Functional MR Images: A Comparison of Wavelet Denoising and Gaussian Smoothing,” IEEE Transaction on Medical Imaging, Vol. 23, No. 3, pp 374-387, (2004)

[72] Nowak, R. “Wavelet-Domain Filtering for Photon Imaging System,” IEEE Transactions on Image Processing, Vol. 8, No 5, pp 666-678, (1999).

[73] Nowak, R., “Wavelet-Based Rician Noise Removal for Magnetic Resonance Imaging,” IEEE Transactions on Image Processing, Vol. 8, No 10, pp 1408-1419, (1999).

[74] Wang, L., “Noise Removal for Medical X-ray Images in Wavelet Domain,” Electrical Engineering in Japan, vol. 163, pp 37-46, (2008).

[75] S. Grace Chang, Bin Yu, and Martin Vetterli,“Spatially Adaptive Wavelet Thresholding with Context Modeling for Image Denoising,” IEEE Transactions On Image Processing, Vol. 9, No. 9, pp 535-539, (2000).

[76] S. Grace Chang, Bin Yu, and Martin Vetterli, “Image Denoising via Lossy Compression and Wavelet Thresholding**,”** Proceedings of the 1997 International Conference on Image Processing (ICIP), Vol. 1, pp. 604-607, (1997)

[77] S. Grace Chang, Bin Yu, and Martin Vetterli, "Resolution Enhancement of Images Using Wavelet Transform Extrema Extrapolation," Proc. Int. Conf. Acoustics, Speech, Sig. Process, vol. 4, pp. 2379-2382, (1995).

[78] Ghael, S, Sayeed, A.M., and Baraniuk, R.G., "Improved wavelet denoising via empirical Wiener filtering," *Proc. SPIE, Wavelet Applications in Signal and Image Processing V*, vol. 3169, pp. 389–399, Oct. 1997.

[79] P. Bao and L. Zhang, "Noise Reduction for Magnetic Resonance Images via Adaptive Multiscale Thresholding," IEEE transaction on Medical Imaging vol. 22, No. 9, pp 1089-1099, 2003

[80] Bismuth, V., and Vaillant, R., "A Devise Enhanceing and denoising Algorithm for X-ray Cardiac Fluoroscopy," In ICPR, 19th International Conference on Pattern Recognition, 2008. http://figment.cse.usf.edu/~sfefilat/data/papers/TuBCT9.39.pdf

[81] Mario Mastriani, "Denoising and Compression in Wavelet Domain via Projection onto Approximation Coefficients," International Journal of Signal Processing. http://www.waset.org/journals/ijsp/v5/v5-1-3.pdf

[82] Oh, J., Hwang, H., Lee, C., and Kim, Y., "Morphological Clustering Filter for wavelet Shrinkage Improvement," International Journal of control, Automation, and Systems, Vol.1 No.3, pp. 390-394, Sep. 2003

[83] Chendeb, M., Khalil, M., Duchene, J., "The use of Wavelet Packets for Event Detection," Proc. the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Francisco, CA, USA, pp. 161-164, September 2004.

[84] Elke Van de Casteele, Dirk Van Dyck, Jan Sijbers, and Erik Raman, "The effects of beam hardening on resolution in X-ray microtomography," In *SPIE Medical Imaging*, 2004. http://webhost.ua.ac.be/visielab/papers/elke/spie04.pdf

[85] Kurt Rossmann, "Point Spread-Function, Line Spread-funtion, and Modulation Transfer Funtion," Radiology 93, pp. 257-272, 1999

[86] J. Sylwester and B. Sylwester, "Reconstruction of Images with Poisson Noise," ACTA ASTRONOMICA Vol. 49, pp. 189-199 (1999)

[87] http://www.as-e.com

[88] http://www.saic.com

[89] http://www.Heimannsystems.com

[90] http://www.invision-tech.com

[91] http://www.vividusa.com

[92] http://physics.nist.gov/PhysRefData/XrayMassCoef/tab3.html

[93] Voosen, P. “NUCLEAR SECURITY: Rushed, costly DHS port venture failed to live up to hype,” E&E news http://www.eenews.net/public/Greenwire/2009/07/24/2

[94] http://www.aapa-ports.org/