ABSTRACT

Automated Segmentation in Medical Imaging

Hanh Hong Nguyen, M.S.

Advisor: Keith Evan Schubert, Ph.D.

Computed Tomography (CT) is one of the most common medical diagnostic imaging techniques. Since the first clinical CT scanner was installed in the 1970s, there are about 30,000 CT scanners installed worldwide. Despite the vast number of scanners and the improvement in image quality, the demand of accuracy in differentiating distinct regions and types of tissues in the treatment area of the patient body still remains. Some materials are easy to identify while some are not due to their thin shape and/or the limited resolution of the scans. This thesis addresses the problem by investigating several image segmentation techniques to achieve fast performance and better quality tissue assignment.

Automated Segmentation in Medical Imaging

by

Hanh Hong Nguyen, B.S.

A Thesis

Approved by the Department of Electrical and Computer Engineering

_____

Kwang Lee, Ph.D., Chairperson

Submitted to the Graduate Faculty of
Baylor University in Partial Fulfillment of the
Requirements for the Degree
of
Master of Science in Biomedical Engineering

Approved by the Thesis Committee

_____

Keith Evan Schubert, Ph.D., Chairperson

_____

Linda Olafsen, Ph.D.

_____

Byron Newberry, Ph.D.

Accepted by the Graduate School
December 2017

_____

J. Larry Lyon, Ph.D., Dean

*Page bearing signatures is kept on file in the Graduate School.*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

# ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisor, Dr. Keith Schubert. He has been patiently giving me guidance, encouragement and support since I sat in his Image Processing class the first semester of my junior year. I would like to thank my committee members, Dr. Linda Olafsen and Dr. Byron Newberry for their guidance and valuable feedback. Lastly, I would like to thank my parents for their unconditional love and constant support.

*to my parents*

*for always being there for me*

# CHAPTER ONE

# Introduction

## *1.1 Motivation*

Proton therapy is the most precise and advanced form of radiation therapy that uses a beam of protons to target a tumor in a human body either to kill or to stop the reproduction of those diseased cells. A proton ionizes human tissue as the result of its energy loss. It deposits a dose along the travel path and stops when all energy has been lost. The ionization of human tissue forms ions and free radicals, thus it causes damage or even kill cells at the destination where proton stops. The advantage of proton therapy over other types of radiation therapy is that a proton carries a long low dose of radiation, hence diseased cells will be destructed while normal cells are exposed to a smaller dose of radiation (Schulte, Levy, Lee, Neupane, Shihadeh, Slusarczyk, Schubert, and Slater 2005). A typical beam of protons travels in non-linear paths throughout the body. Each individual proton within one beam needs to be processed separately to maintain sharpness in the reconstruction. The reconstruction becomes a time and resources consuming task. Beside from using proton, other particle such as Helium ions are also utilized and studied.

The Baylor Proton Computed Tomography (pCT) Preprossessing/Image Reconstruction program reconstructs pCT images from scan measurements acquired from experiments or simulations. (Schultze, Karbasi, Giacomelli, Plautz, Schubert, and Schulte 2015) It is shown that the software can run in under 2 minutes with the set of the pediatric phantom head data. The reconstructed image (Figure 1.1) gives decent amount of information, but improvements in edge detection and segmentation can happen in order to have better quality image for medical simulations and early stage diagnosis.

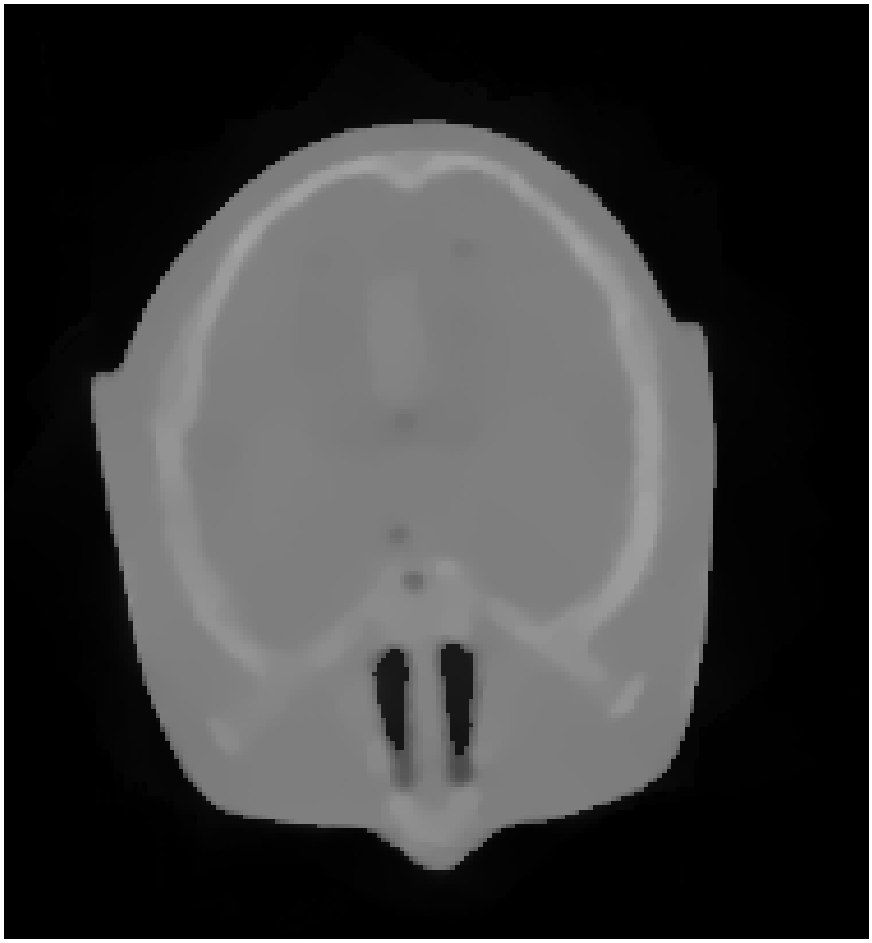

Figure 1.1. The output of Baylor pCT software, one scan slice of the pediatric head phantom, ran in August 2016 beam at Chicago Proton Center.

Intensity Modulated Proton Radiation Therapy (IMpRT) is a therapy that delivers a highly conformal dose to a tumor or other target even in close proximity to sensitive and critical organs. IMpRT is delivered in several dose fractions to shape the dose to the target using multiple directions with different narrow proton beams or pencil beams and to modulate the intensity of each beam in case it passes through the critical organ or area at risk.

To deliver doses correctly, it is essential to assign regions in the CT scans. One simple ways to do this is to use the Hounsfield unit (HU) intervals as threshold values and assign them to specific tissue as in Table 1.1. However, this assignment

is not the perfect solution for this problem dues to the standard deviation of noise, the overlap of some material, and artifacts in CT scans.(Padhye 2013)

Table 1.1. Tissue Categorization according to Hounsfield Unit Value

| HU Interval | Type of Tissue |
|---|---|
| $[-1000, -800)$ | air |
| $[-800, -700)$ | sinus |
| $[-700, 40)$ | soft tissue |
| $[40, 90)$ | brain |
| $[90, 150)$ | spinal disc |
| $[150, 200)$ | trabecular bone |
| $[200, 1000)$ | cortical bone |
| $[1000, 2000)$ | tooth dentin |
| $\geq 2000$ | tooth enamel |

Identifying accurately materials and regions in CT scans has been the demand in various aspects in the medical field. With the current medical and plan treatment system, a physician needs to manually identify and mark the boundaries for the region of interests in every scan to create the proton treatment plan. If materials and regions can be detected accurately and automatically, it can be helpful for both pCT image reconstruction and IMpRT treatment planing. (Sharma and Aggarwal 2010)

## *1.2 Thesis Overview*

The primary purpose of this thesis is to propose hybrid methods to identify materials and regions in CT scans. To accomplish this, the following issues were addressed:

(1) Review of the most common image segmentation methods used in medical imaging.

(2) Propose hybrid methods that can improve the quality of differentiating type of tissues.

Chapter Two gives the background information about Computed Tomography, the history and its principles of operations as well as the overview of image

segmentation. Nowadays there are thousands of clinical CT scanners installed and operated worldwide. Major improvements have been made in generating medical images since the 1970s. However, tissue identification still faces obstacles such as noise, low contrast, mixed pixels and uncertain regions. Image segmentation is the process of splitting up a digital image into multiple segments to simplify or provide more meaningful information about that image. By assigning a tag to every pixel in one image, those pixels with the same tag will have the same visual attribute. If the segmentation process becomes automatic and accurate, it will be a helpful and time-saving tool in treatment planing, verification imaging, and hopefully early diagnostic medicine. Chapter Three contains the overview of image segmentation techniques. The main principle of each technique will be discussed as well as the advantages and disadvantages each technique. Chapter Four presents the testing method and the results of the implementations discussed in Chapter Three. Chapter Five summarizes the improvements to tissue segmentation and the future direction of this work.

# CHAPTER TWO

## Background

This chapter will discuss about the two essential principles that needed for this thesis as well as the introduction of the phantom that used in this research work. Section 2.1 goes over the history and principle of Computed Tomography. Section 2.2 gives a brief introduction about image segmentation. Section 2.3 introduces the phantom.

### *2.1 Computed Tomography*

Since the discovery of x-rays in 1895, there have been tremendous studies and new inventions utilizing x-rays in medicine. Beside plain x-rays images, Computed Tomography (CT) is one of the main procedure that uses special x-rays machine to create specific pictures or scans of an interested region inside the body. Tomography is derived from Greek words "tomos", which means a cut or section, and "graphia", which means to write or record.

In 1955, Allen M. Cormack probably built the first CT scanner. While observing the planing of radiation treatment, Cormack realized the important of knowing the x-ray attenuation coefficient distribution inside the body. To improve the accuracy of radiation treatment, he did the experiments to reconstruct the attenuation coefficient of tissues in 1956 and again in 1963. Even though his work did not gain much attention at the time, it was still remarkable in CT history. In 1967, Godfrey N. Hounsfield started to develop the first clinical CT scanner at the Central Research Laboratories of EMI, Ltd, in England. It took nine days to acquire the data and then a hundred fifty minutes to reconstruct the image on a computer at that time period. Hounsfield independently of Cormack discovered that measurements of x-rays taken from different directions inside the body could reconstruct the internal structure. The

first clinical CT scanner was installed in 1971 at Atkinson-Morley Hospital in England. The first clinical data acquisition took about four and a half minutes, which was significantly faster than the original prototype. For the foundation work in CT, both Cormack and Hounsfield received a Nobel Prize in Medicine in 1979. Since the first clinical CT scanner installed until now, the quality of scans and time needed to taken CT scans have evolved and advanced.(Hsieh 2003)

CT uses x-rays to create a series of two-dimensional images of of the internal stuctures of the scanned object or patient around a single axis of rotation. The x-ray source in a CT scanner rotates around the patient as it moves through a gantry. A gantry is the largest system in a CT scanner. It holds components such as the scanned object support system, the x-ray tube, and detectors. Detectors are placed at the opposite side from the position of x-ray source. CT scanners in the 1970s only have one or a few detectors while the current CT scanners contain an arc or a whole ring of detectors. This setup allows reduction of the scan time and the exposure of radiation. At the present time, there are seven generations of CT scanner design, which are distinguished based on the relation of the x-ray source and detectors.

CT is a projection radiography method because it represents the projection of a 3-D object onto a 2-D image (often called slice). CT collects multiple projections of the same positions from different positions by moving the x-rays around the object. The systems have digital detectors on the opposite side and they are used to reconstruct the cross sectional slices of the object of interest. CT generates tomographic images after reconstruction, not projections or "shadows" of the object. (Lee 2006)

Some of the artifacts that can appear in CT images are aliasing, beam hardening, electronic or system drift, x-ray scatter, and motion. Artifacts occur mainly because of patient motion, improper preparation of scan protocol or calibration of the machine. Those artifacts often reduce the image quality. (Prince and Links 2006)

The Hounsfield Unit (HU) is the x-ray attenuation unit commonly used in CT scan interpretation (Gücük and Üyetürk 2014). The HU scale is the standard scale to calibrate CT scanner. The type of tissue has the same HU value, range from $-1024$ to 3071 HU. The formula to calculate HU of tissue $A$ is computed by using its linear attenuation coefficient $\mu_A$, and linear attenuation coefficients of water $\mu_{water}$ and air $\mu_{air}$.

$$HU_A = 1000 \frac{\mu_A - \mu_{water}}{\mu_{water} - \mu_{air}} \tag{2.1}$$

### *2.2 Image Segmentation*

Image segmentation is one of the first and important steps in image analysis. It is the process of splitting up a digital image into multiple segments based on a discontinuity or similarities in characteristics. Image segmentation is crucial in image analysis because it can provide more meaningful information about an image. It connects the bridge between low and high level image processing. In the identified segments, or regions, every pixel should share the same characteristics. In a math-oriented definition, image segmentation is the process of partitioning the image into non-overlapping regions such that each region is homogeneous and connected, and the union of adjacent regions is not homogeneous.

Segmentation is a process of partitioning $I$ into $n$ sub-regions. Let $I$ be the image and $R_1$,$R_2$,...,$R_n$ be sub-regions such that

- $\bigcup_{i=1}^{n} R = I$
- $R_i$ is a connected region $\forall i$
- $R_i \bigcap R_j = \emptyset \quad \forall i, j, i \neq j$
- $P(R_i) = \text{TRUE} \quad \forall i$
- $P(R_i \cup R_j) = \text{FALSE for } i \neq j$

In which $P$ is the logical predicate defined over the points in set $R_i$ and the null set $\emptyset$. (Gonzalez and Woods 2007)

Image segmentation has various practical applications such as:

- Object detection: face detection, in-use vehicle or traffic (pedestrian or brake light) detection, and map (roads, trees, lakes,...) detection from satellite images.
- Recognition tasks: face and finger prints recognition (Fabijanska, Kuzanski, Sankowski, and Jackowska-Strumillo 2008)
- Traffic control systems (Walad and Shetty 2014)
- Medical imaging
    * Detecting of tumors detection and other pathology
    * Surgery planning and virtual surgery simulation
    * Diagnosis
    * Study of anatomy

Over the years, many algorithms, techniques, and methods of image segmentation have been developed and studied to fit the specific need in the particular field. Depending on the specific need or requirement, an image can be subdivided into regions in different levels of details. In many cases, segmentation process can provide a certain amount of information about the region or object of interest.

For this thesis, the interest of image segmentation is in medical image. In theory, one medical image could be divided into regions by defining different HU values to different tissues then differentiate the image to get those different regions. However, this method is not sufficient because of noise and artifacts of CT images.

### *2.3 Pediatric Head Phantom*

A digital head phantom Figure 2.1 had been created in the Department of Radiation Medicine at Loma Linda University by Dr. Reinhard Schulte's research group to study and optimize the performance of various pCT reconstruction and IMpRT solution algorithms. The head phantom is based on an anatomical phantom of the head and neck of a 5-year-old child. The phantom is a realistic model in term

Figure 2.1. Commercial Pediatric Head Phantom in the CT Scanner at the Department of Radiation Medicine, Loma Linda University Medical Center

of anatomical details and electron density of the materials. The anatomy of this phantom contains a homogeneous brain with ear canals, sinus cavities, teeth with enamel and dentin, cervical spine and spinal disc.

The phantom was scanned in the 64-detector-row CT scanner (Light speed, GE Health care, Waukesha, WI). The scan consists of 200 slices and each slice is organized into a matrix of 512-by-512 image pixels. There are two thickness levels of 1.25 mm and 0.625 mm. In three-dimensional, this scan creates a digital space comprised of the order of 50 million voxels.

### *2.4 Bone Anatomy*

In many previous studies, the main goal was to automate segmenting of the human head structure based on hard and soft tissue. In this thesis work, the main target is to identify the bone structure (cortical and trabecular bones) from CT scans as these create the greatest effect on the proton path. The four general categories of bones are long bones, short bones, flat bones, and irregular bones. (Clarke 2008)

Human skull is a flat bone that contains 22 different bones such as the 8 cranial bones and 14 facial skeleton bones. Most of facial bones are flat or small bones.

Within one bone, there are cortical and trabecular regions, see Figure 2.2. Cortical bone is called hard bone because it is dense and solid, while trabecular bone is known as spongy because of its honeycomb-like soft structure. Both cortical and trabecular bone are composed of osteons.(Clarke 2008)

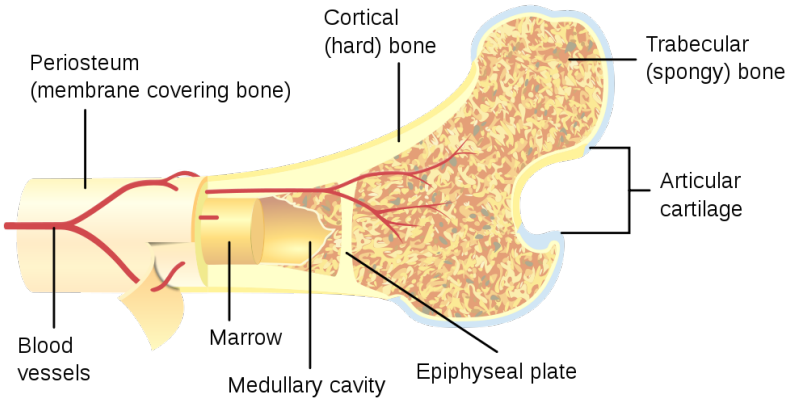


Figure 2.2. Cross section of bone anatomy (Pbroks13 2007)

# CHAPTER THREE

## Image Segmentation Techniques

This chapter will examine some of the major techniques of image segmentation. Pixel-based segmentation, which uses gray-level values to classify each pixel into its own category, is examined in Section 3.1. Edge-based segmentation, which detects an abrupt change in intensity of the image, is discussed in Section 3.2. Region-based segmentation techniques, which considers gray-level of the neighboring pixels to identify the region, are investigated in Section 3.3.

### *3.1 Pixel-based Segmentation*

Pixel-based segmentation uses gray-level values to classify each pixel into its own category. It is the simplest method in segmentation. Thresholding is one technique of pixel-based segmentation. As the name pixel-based suggested, it classifies each and every pixel in an image into categories, then assigns a pixel to be or not to be in one region.

In the simple case, an appropriate threshold will separate the object from the background and other regions. For example, there is a pixel $(x, y)$ in image $I$ with threshold $T$. If the intensity of $(x, y)$ is equal or higher than $T$, $(x, y)$ will be considered as part of the object.

There are two types of thresholds, local and global thresholds. If threshold $T$ is used consistently throughout the image, $T$ is a global threshold. Otherwise it is a local threshold, and only applies for a particular area of the image.

Thresholding is fast and easy to understand as well as implement. However this technique only works well with images whose background is clear and distinct from the object. In particular, global thresholds will not work with uneven background-illuminated images. While that will not affect local threshold methods because of the

multi-level thresholds, thresholds are sensitive to noise. Thus the separation between the object and its background will be corrupted in more complicated noisy images.

Figure 3.1 shows the example of thresholding segmentation technique. The black and white picture of coins on the left is the input image. On the right is the result image. The threshold establishes to classify between the background and the individual coin. The coins are clearly distinct from the background but the details inside each coin are hard to distinguishable.

(a) Original Image (b) Segmented Image after Thresholding

Figure 3.1. An example of Thresholding

### *3.2 Edge-based Segmentation*

Edge-based segmentation detects edges and links them together to form regions. Edge detection operates based on detecting discontinuities or similarities in gray-scale level, color, texture, etc.

Edge detection is the process of identifying and locating sharp discontinuous boundaries in an image. Those boundaries are abrupt changes in pixel intensity. Edge detection is naturally understandable as edges are usually the important feature in an image to separate regions. However the results from edge detection cannot be use directly. Edge linking is necessary to present the closed region of interest.

The general principle of edge detection is to convolve an image with an operator to identify gradients in that image while restoring values of zeros in uniform regions. (Padhye 2013) Because of the importance, edge detection is an active research. There are numerous edge detection operators to choose from, depending on the specific need. Some of the determining factors to choose the suitable operator are:

(1) Edge profiles: This factor is important because edge detection depends on the change in intensity and not every edge has the same type of changes in its intensity. Step edge has a abruptly change from one value on this side to another value on the opposite side while ramp edge has a steady change from one to another value of intensity. Line edge has a value of intensity change in a period then returns to the previous value. Roof edge is a line edge with the steady change in intensity over finite distance. (Ramesh, Rangachar, and Brian 1995)

(2) Edge orientations: An operator at first can search for diagonal, horizontal and vertical edges because those are easier to determine.

(3) Noise: Image noise is random but significant in medical imaging. Noise can be defined as unwanted pixels and often it makes variation in the image to reduce meaningful information. The change due to noise is not an edge but it might be detected by an operator. Reducing noise also blurs and distorts edges.

The edge set produced by an edge detector can be divided into three types: correct edges which are the real edges, false edges which do not exist, and missing edges which do exist but are not detected. False edges are false positives while missing edges are false negatives. (Ramesh, Rangachar, and Brian 1995)

Edge detection techniques can be classified into categories based on derivatives of the image. There are first-order derivative operators, second-order operators, and optimal edge detection operator such as Canny edge detection.

### 3.2.1 First-Order Derivative Operators

First-order derivative operators, as known as gradient operators, find edges by searching for the maximum and minimum in first-order derivative of an image. Some of the well-known first-order derivative operators are the Roberts cross operator, the Sobel operator, and the Prewitt operator. These operators utilize two kernels to compute an approximation of the gradient of the image intensity at each pixel in vertical and horizontal directions.

Let $I(x, y)$ be the image intensity function of the input image $I$. In theory, the gradient magnitude $G(x, y)$ of the image intensity function can be computed as

$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \tag{3.1}$$

in which $G_x(x, y)$ and $G_y(x, y)$ are partial derivatives at pixel $(x, y)$. In common practice, an approximation of the magnitude is computed as

$$|G(x, y)| = |G_x(x, y)| + |G_y(x, y)| \tag{3.2}$$

which is computationally faster and still acceptably works as derivatives. The approximation is zero when intensity is constant and it is proportional to the direction of the change in intensity of areas which pixel values are variables.

The direction of the edge at pixel $(x, y)$ is perpendicular to the gradient vector $G(x, y)$. The direction of gradient vector $\alpha(x, y)$ is the angle of $G(x, y)$ at pixel $(x, y)$

$$\alpha(x, y) = tan^{-1}\frac{G_y(x, y)}{G_x(x, y)} \tag{3.3}$$

In the first-order derivative operators, two masks $G_x(x, y)$ and $G_y(x, y)$ are convolved with the input image $I$ to calculate the approximations of the derivatives in vertical and horizontal directions.

The Roberts cross operator is one of the first edge detector that was introduced by Lawrence Roberts in 1963. The Roberts cross operator's convolution masks are a two-by-two matrices that respond maximally to edges running at 45 degrees to the pixel grid with one mask for each of two perpendicular orientations,

$$G_x(x,y) = \begin{bmatrix} -1 & -0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad G_y(x,y) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \tag{3.4}$$

The Sobel operator is presented by Irwin Sobel and Gary Feldman in 1968 as an "Isotropic 3x3 Image Gradient Operator". The operator computes an approximation of the gradient of the image intensity function at each pixel. The result is the corresponding gradient vector or the norm of that vector. The Sobel operator's convolution masks are two three-by-three matrices,

$$G_x(x,y) = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad G_x(x,y) = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}. \tag{3.5}$$

The Prewitt operator works similarly to Sobel operators with two three-by-three masks,

$$G_x(x,y) = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad G_y(x,y) = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}. \tag{3.6}$$

Gradient operators are good to detect sharp edges in which the pixel intensity levels change rapidly. When the intensity level changes slowly from lower to higher value, the other operators should be considered

### *3.2.2 Second-Order Derivative Operators*

Second-order operators or Laplacian operators determine edges by zero crossing at the second-order derivative of an image. Zero crossing of the second derivative of a function is identified as the presence of a maximum of that function. The maximum location corresponds to the location of edge in the image.

The two-dimensional Laplacian operator for function $f$ is defined as,

$$\nabla^2 f = \frac{\partial f}{\partial x^2} + \frac{\partial f}{\partial y^2} \tag{3.7}$$

The approximations of the second derivatives along x and y directions are computed by the following mask,

$$\nabla^2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{3.8}$$

When more weight is needed for the center pixel, the mask to use is,

$$\nabla^2 = \begin{bmatrix} 1 & 4 & 1 \\ 4 & -20 & 4 \\ 1 & 4 & 1 \end{bmatrix} \tag{3.9}$$

After convoluting the input image with the Laplacian mask, minimum and maximum are found within the neighborhood of a pixel. If the minimum is smaller than zero and the maximum is greater than zero, the pixel is a zero crossing. Sometimes false zero crossing can be detected due to the existence of random noise.

Laplacian of Gaussian (LoG) detects edges by doing the four following steps,

- Smoothing the input image to remove unwanted noise by convoluting the input image with a Gaussian mask $G(x, y)$ with the width $\sigma$

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{3.10}$$

- Enhancing edges by applying the Laplacian operator
- Finding zero crossings to determine edge locations
- Thresholding the zero crossings to keep only strong edges

LoG is defined as,

$$LoG(x, y) = -\frac{1}{\pi\sigma^4}[1 - \frac{x^2 + y^2}{2\sigma^2}]e^{\frac{x^2+y^2}{2\sigma^2}} \tag{3.11}$$

in which $\sigma$ is a standard deviation of Gaussian filter.

One simple LoG mask is

$$LoG = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ -1 & -2 & -16 & -2 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \tag{3.12}$$

The larger size of the LoG mask requires longer computation time because of the complexity of the computaton. The difference of two LoG or the Difference of Gaussian is the approximation of LoG.

$$DoG(x, y) = \frac{e^{\frac{x^2+y^2}{2\sigma_1^2}}}{2\pi\sigma_1^2} - \frac{e^{\frac{x^2+y^2}{2\sigma_2^2}}}{2\pi\sigma_2^2} \tag{3.13}$$

in which $\sigma_1$ and $\sigma_2$ are standard deviations of Gaussian filters.

### *3.2.3 Optimal Edge Detector*

The last category is the optimal edge detection operator. The Canny edge detector, which was developed by John Canny in 1986, is an edge detection method that uses multiple operators to mark a wide range of edges in an image. It responded to real edges and gave no response to non-edges. The distance between each edge pixel and the real edge is minimized. Each edge also has a single response to eliminate the multiple detection to one edge. Canny edge detector is recognized as the optimal edge detector because it is significantly better than other existing edge detectors at the time. The following criteria are considered: (Canny 1986)

(1) Good detection: There should be a high chance to mark read edge points, and also a high chance to mark non-edge points.

(2) Good localization: The marked edge points should be as close to the center of true edges as possible.

(3) Single edge response: There should be one and only one response to a single edge.

There are four steps of Canny edge detection. The first step is to filter any noise in the input image before doing any edge detection. A Gaussian filter is used to smooth any existing noise. After choosing or calculating a convolution mask, Gaussian smoothing can be performed using the regular convolution method. The dimension of a convolution mask is smaller than the dimension of the input image, thus it can be understood as the mask slides throughout the image. The size of the Gaussian mask affects its sensitivity to noise as well as the smoothness of the output. This first step is similar to the initial step of LoG.

The second step is to determine the intensity of gradients by computing gradient magnitude and its orientation by using approximation of partial derivatives to determine the regions with high spatial derivatives. In practice, simple mask filters, such as Roberts, Prewitt, or Sobel, can be used to compute the gradient of the image

since there is no specific function for either an edge or image. The gradient orientation is also calculated by using the gradient in $x$ and $y$ directions.

The third step is to apply non-maxima suppression to the gradient magnitude and eliminating non-maxima pixels. Edges occur at pixels that the gradients are at maximum, thus every other point not at maximum gradient should be discarded. To discard those non-maxima pixels, the magnitude of gradient at every pixel is compared with the magnitude of gradient at the neighboring pixels. If it is not greater than its neighbors, then it will be suppressed. The result of this step is thinned edges at the points of the greatest local change.

The final step is to detect edges by hysteresis with two thresholds, the lower value sets non-edge to zero value while everything above the upper threshold is set as an edge. The edge magnitudes that lie below upper threshold but above lower threshold are retained if and only if they connect to other edge that is above the upper threshold.

Some parameters that can affect the performance of Canny edge detection are the size of Gaussian filter and the values of thresholds. The first step of choosing a smoothing filter is important because it sets the initial input for the whole detection process. For example, a large sized filter can smudge out more than desired area in the image, resulting loosing small edges. On the other hand, small sized filters can be too small and too detailed so that they give minor blurring and many small, sharp edges will be detected.

Figure 3.2 shows the example of Canny edge detection in step by step from the input image to the edge traced image. The input Figure 3.2a of this example is the black and white picture of coins. Each of the following figures from Figure 3.2b to Figure 3.2f illustrates one step of the Canny edge detection technique. At first, the input image is smoothed by apply a three-by-three Gaussian filter to remove any noises. Then the gradient magnitude and its orientation are computed by using

approximation of partial derivatives in $x$ and $y$ directions. A large amount of edges appear in this step. In Figure 3.2c, non-maxima suppression is applied to the gradient magnitude to eliminate non-maxima pixels. By the end of this step, the thinner edges at the points of the greatest local change are attained by the end of this step.

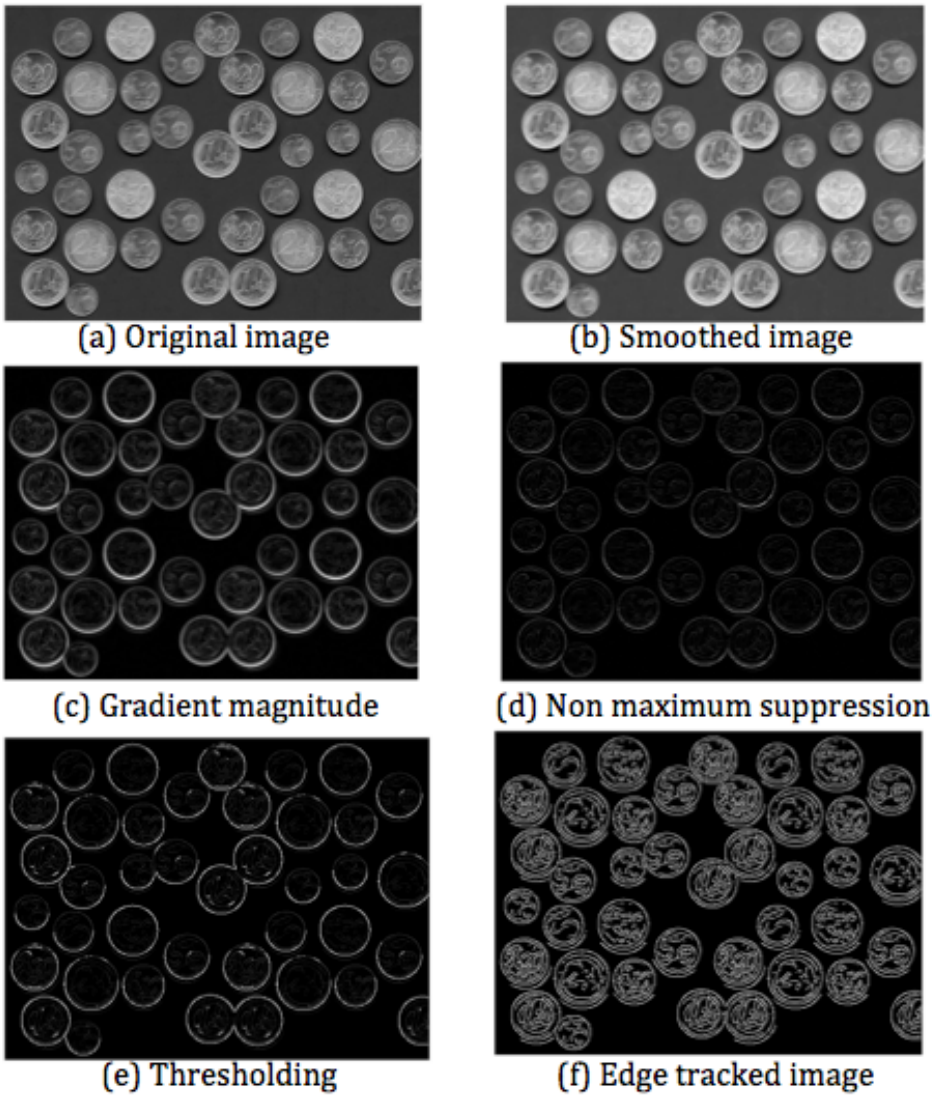


Figure 3.2. An example of Canny edge detection

Three-dimensional Canny edge detection is the extension of the original Canny edge detection. It also follows four steps of the original Canny edge detection, but instead of one scan, a set of three scans are entered as an input.

### *3.3 Region-based Segmentation*

Region-based segmentation considers gray-level of the neighboring pixels to identify the specific region. It is the process of dividing an image into regions while considering the neighboring pixels instead of looking for boundaries or any thresholds. The basic formulation of region-based segmentation is about completeness, connection, disjoint, satisfaction, and segmentation.

Region growing is one of the region-based segmentation methods that combines pixels or small regions with similar properties into a larger region. The general idea of region growing is to check the neighboring pixels of the initial seed points, then decide if those neighboring pixels have the same or different property with the seed points.

There are two steps in region growing. At first, a set of seed points is chosen. If one or more of the neighboring pixels can satisfy the checking criteria such as intensity, gray-level, color, boundary or etc. then it establishes a region along with the original seed point. The process repeats until all of the pixels are checked. Thus, this is an iterative process.

The example in Figure 3.3 is to explain the seed region growing method. The criteria for this test is the value of pixel. If the value of a neighboring pixel is as same as a seed point, that neighboring pixel is added to a seed point and one region is established. The process stops when there is no change in two successive iterations. The input of this example is Figure 3.3a - a 9-by-9 image that contains random pixels value from 1 to 9. The seeds are chosen as randomly selected seeds with distinct values. Four seeds show in Figure 3.3b. The next step is to grow a region around one seed. If the intensity value of the neighboring pixel is equal the intensity value of

that seed, then one region is formed as in Figure 3.3c. The checking process continues until every pixel is assigned to one region.

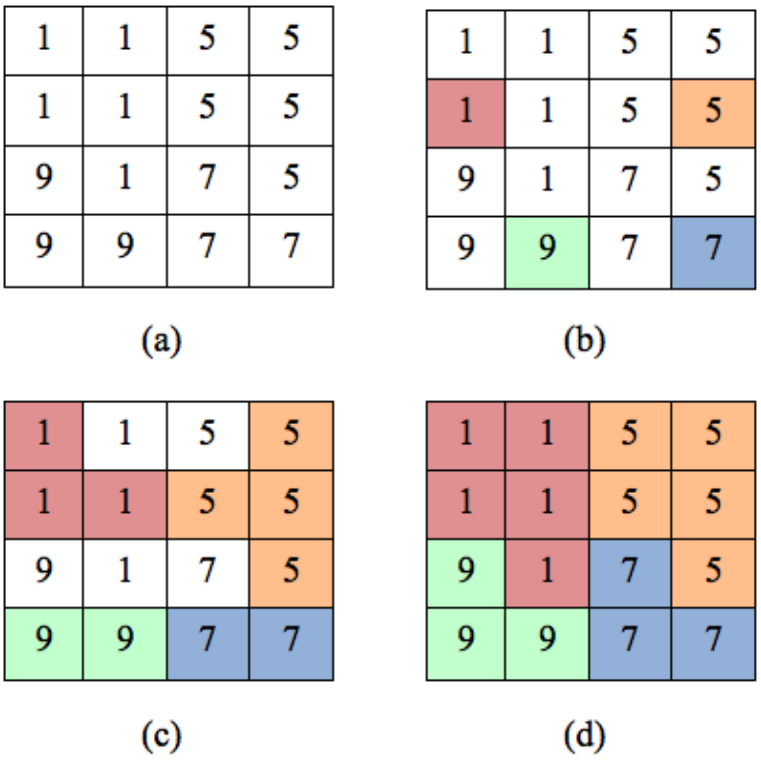


Figure 3.3. An example of region growing

Region growing is better than edge-based and thresholding techniques in noisy images where edges and thresholds are difficult to determine. Figure 3.4 is the result of random seeded region growing segmentation from the same input image as the example of Canny edge detection in Figure 3.2 above. A pointer walks through all pixels of the input to find the peak values. Those peak values are selected as seeds. Then the difference of a seed point and its neighboring pixel values is computed. If the difference is smaller than the defined threshold, which is calculated as ten percent of the difference of the maximum and minimum of the pixel values, then that neighboring point could be classified into the same region as the seed point.

The process of checking repeats until the difference of pixel value is greater than the defined threshold. When it happens, the first region is marked with a set of boundaries, those boundaries then are used as a new seed points. The region growing process ends when every pixel has been assigned to the distinct region.

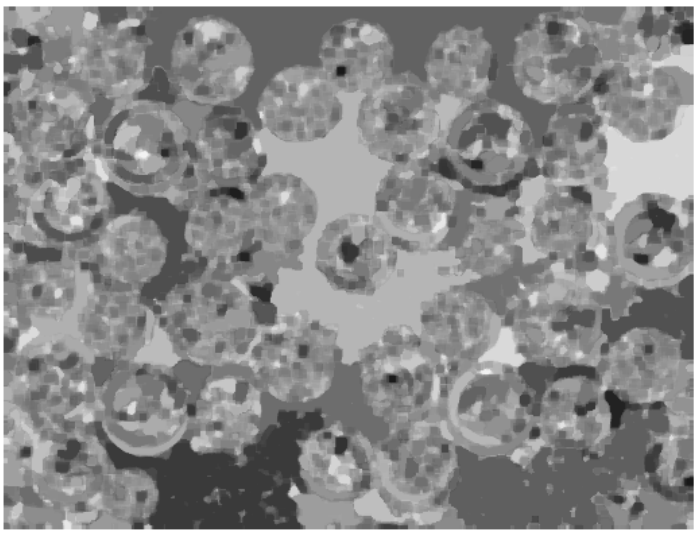

Figure 3.4. An example of random-seeded region growing implementation

# CHAPTER FOUR

## Results

This chapter presents the implementation results of the original Canny edge detection, the three dimensional Canny edge detection, seed region growing and the new hybrid techniques.

### *4.1 Canny Edge Detection*

The main purpose of edge detection is to reduce the amount of details in an image, while providing the structural details so that it can be used for image processing. Canny edge detection is one of the efficient technique to achieve this purpose. By using the built-in Canny edge detection function in MatLab, Figure 4.1, shows the result. Canny edge detection is computationally more expensive than other techniques but it gives better results in detecting smaller and weaker edges. Canny algorithms can be adjusted and optimized by parameters $\sigma$, which is the standard deviation for the Gaussian filter, and the threshold values, $T_{low}$ and $T_{high}$. $\sigma$ is directly proportional to the size of the Gaussian filter, as the value for $\sigma$ increases so does the Gaussian filter's size.

Figure 4.1 shows the result of Canny edge detection with optimized parameters from MatLab. Even though the result looks good to the human eyes, there are ambiguous regions in which the edges are thinner or broken up (circled in red). When performing solely the Canny edge detection, edges are identified but the edges of middle regions with more complicated details are not clear as the outer regions. When the detail is missing, the result of this implementation cannot be used in any purpose such as clarification for medical diagnosis or planning treatment. Thus, the optimal Canny edge detection is not the optimal solution.

Figure 4.1. Original Canny Edge Detection, slice 100

Figure 4.2 shows the result of Canny edge detection in different ranges of threshold. The lower range of threshold gives the result of the left image while the higher range of threshold gives the result of the right image. The original thresholds are set using MatLab built-in Canny function. The lower range of threshold is calculated as a quarter of the original threshold range. The higher range of threshold is twice as great as the original threshold range. As expected from the definition of threshold, the lower threshold range does not trace as many edges as the higher range. However the higher threshold range also collects more edges and scattering dots in the middle of the scan than needed. The first impression from this result is that the lower range of threshold would not work because it allowed Canny to keep so many noise as edges. The higher range of threshold would not work either since it created many gaps between edges.

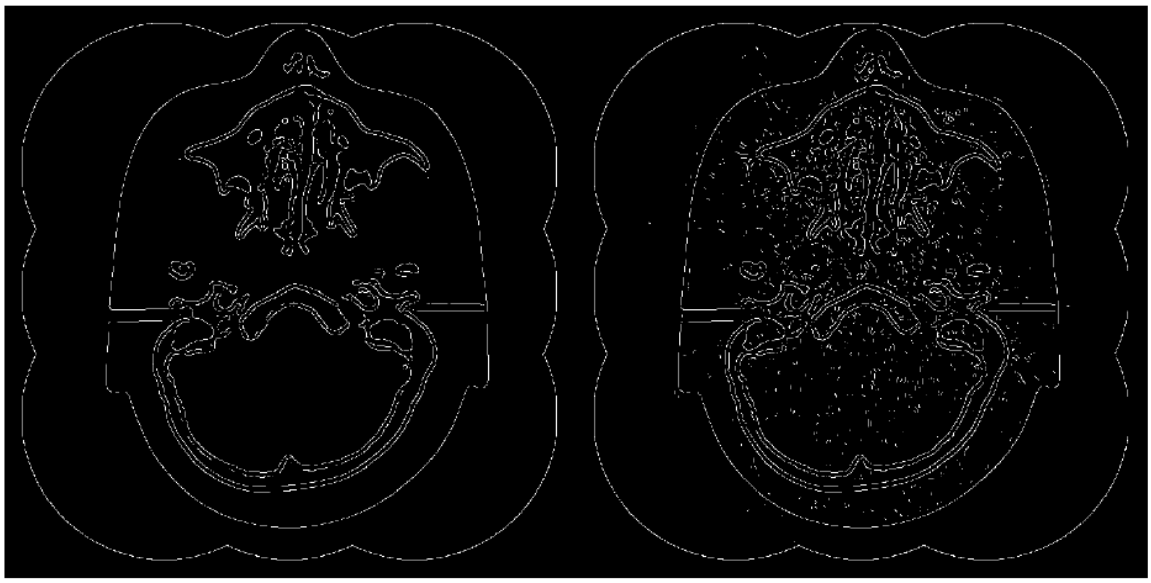

Figure 4.2. Original Canny Edge Detection, low vs. high threshold ranges, slice 100 of the pediatric phantom

This implementation shows that Canny edge detection works for detecting edges in general. From its results, the general idea of regions and edges can be determined. Thus, to optimize and automate the technique for more advanced image processing and medical verification, three-dimensional Canny edge detection is investigated.

### *4.2 Three-Dimensional Canny Edge Detection*

As mentioned in previously, three dimensional Canny edge detection is the extension of the original Canny edge detection. The implementation of the three-dimensional Canny edge detection is more time consuming since instead of a two-dimensional scan, a set of three scans are entered as an input. Three-dimensional Canny edge detection gives more connected and defined edges within the same set of parameters.

Figure 4.3 and Figure 4.4 are the results of two-dimensional vs three-dimensional Canny edge detection side by side. The input slices are slice 20 and slice 80. These

slices are one of the earlier slices out of 200 slices of the pediatric head phantom. These slices are not as complicated comparing to later slices but they still contain reasonable details such as the brain soft tissue, the cortical and trabecular bones, and the nasal cavity inside the skull to be investigate.

Slice 20 shows the image of the homogeneous brain with the skull. The most interested part of this slice is to differentiate the cortical layer around the inside of the skull. Figure 4.3 is the zoom in of the result of the implementation. In the two-dimensional result on the left, at he bottom of the scan, the inside edges of the back bones of human head are not continuous. From common knowledge, those edges should be continuous. On the other hand, the result of three-dimensional Canny edge detection on the right shows clear and more defined edges of those bones.

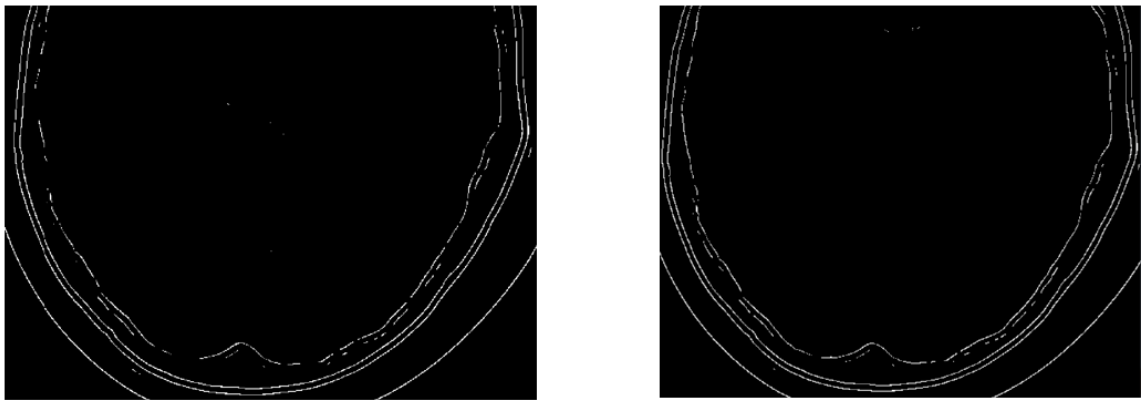

Figure 4.3. 2-D vs. 3-D Canny Edge Detection, zoom in slice 20 of the pediatric phantom

The same thing happened in Figure 4.4 which shows the results in slice 80. When three-dimensional Canny edge detection can give true weak edges (near the nasal cavity), two-dimensional Canny edge detection misses most of the lower part of the nasal cavity. With the missing part, the result of the two-dimensional Canny edge detection cannot be used in any of the applications in practical medical area.

The three-dimensional Canny edge detection shows promising results because it is better at detecting edges of small and thin bones.

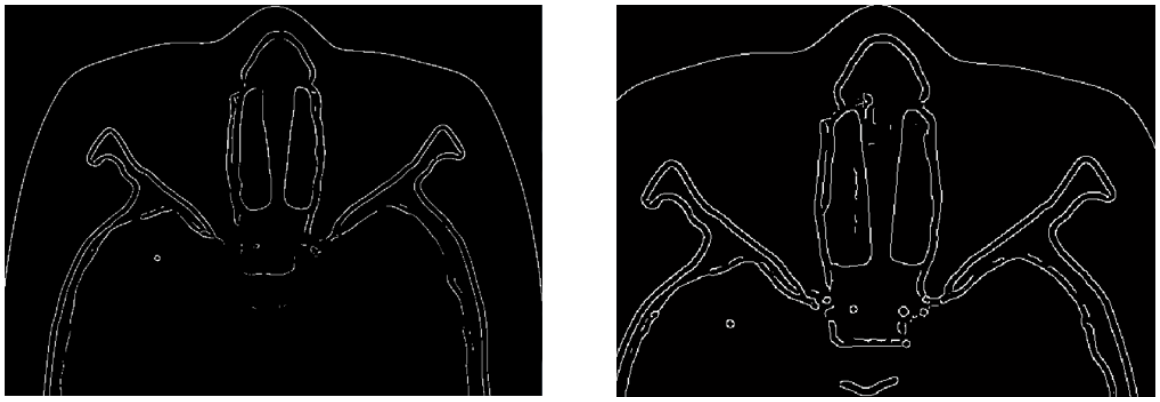

Figure 4.4. 2-D vs. 3-D Canny Edge Detection, slice 80 of the pediatric phantom

This implementation of three-dimensional Canny edge detection shows that the three-dimensional approach works for detecting edges in general. From the results, the general idea of regions and edges can be determined. Thus, to optimize and automate the technique for more advanced image processing and medical verification, three-dimensional Canny edge detection is investigated.

### *4.3 Seed Region Growing*

Seed region growing is the last segmentation method to look at. It is the implementation of region growing with the simple seed picking process - to choose a seed according to histogram peaks and unlabeled regions. A pointer will walk through all pixels of the input to find the peak values. Then the difference of pixel value of the initial seed point and its neighboring points is computed. If the difference is smaller than the defined threshold, then the neighboring point could be classified into the same region as the seed point. The process repeats with the next pixel until the difference of pixel value is greater than the defined threshold. When it happens, the first region is marked with a set of boundaries, those boundaries then are used

as a new seed point. The growing process stops when all pixels in image have been assigned to the distinct region.

Figure 4.5 is the seeded region growing result in slice 60 of the pediatric phantom. The first impression from the result is that the image is too noisy. A lot of smaller regions are defined. In some regions, the material is supposed to be the same, but the result shows differently. It is hard to say if seed region growing is worse or better than Canny edge detection because this technique still provides the information needed to detect different regions within the image of the brain. Perhaps with adjusting the method of selecting the original seeds, seed region growing will give better performance.

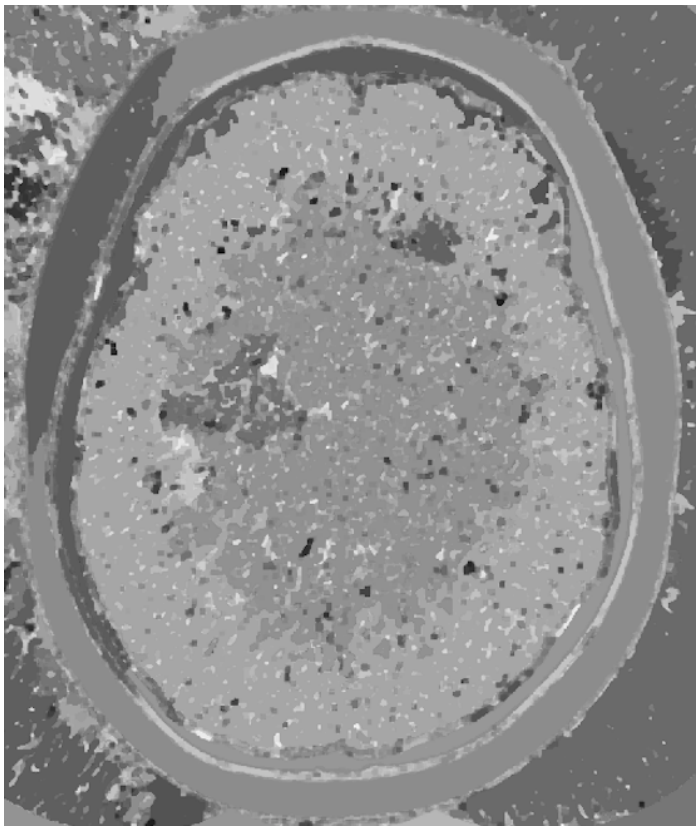

Figure 4.5. Seeded Region Growing, slice 60 of the pediatric phantom

## *4.4 Hybrid Techniques*

### *4.4.1 Three-Dimensional Canny Edge Detection and Thresholding*

The first hybrid technique is a combination of three-dimensional Canny edge detection and the predefined thresholds using Hounsfield unit, i.e. cortical bone falls in the interval of 200 to 1000.

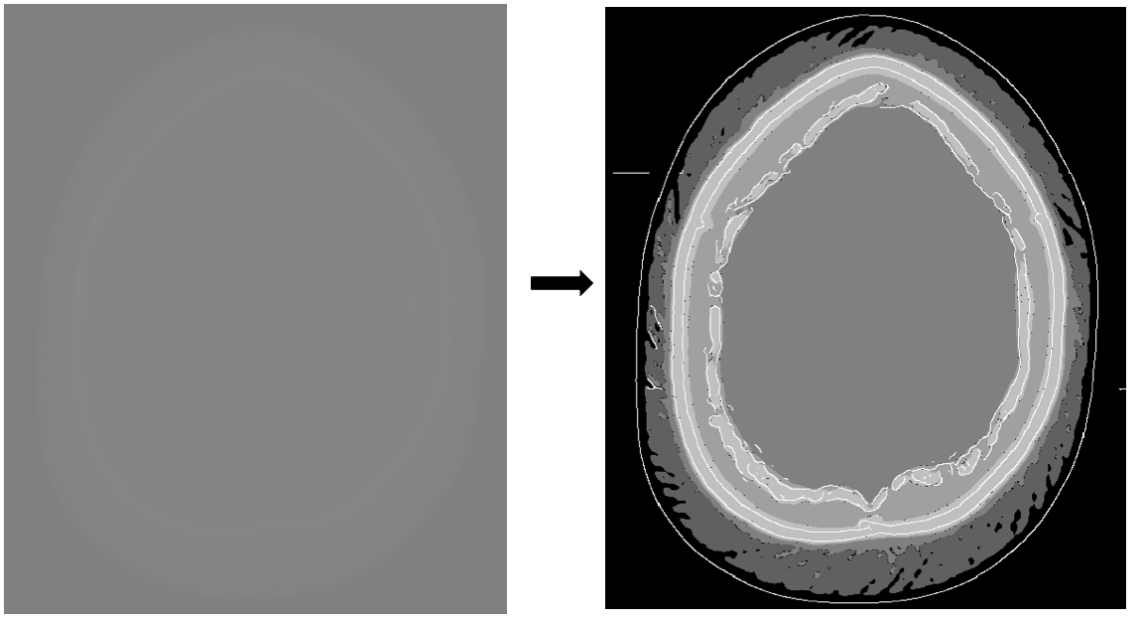

Figure 4.6. The original vs hybrid of 3-D Canny edge detection and threshold, slice 20 of the pediatric phantom

Comparing between the raw scan and this hybrid result in Figure 4.6, the result in Figure 4.6 contain information about the edges and major regions. Most of the regions toward the center of the image are identified as different materials. On the outside, there are some disconnections between the edges and the regions. The edges, especially weaker edges, are more defined in this first hybrid technique. The implementation of the three-dimensional Canny edge detection and thresholding technique shows the improvement in differentiating between hard (bone layers) and soft (brain) tissues. Most of the soft tissues are detected but not all. There are

patches around the skull that are supposed to be soft tissues but they are marked as the background in the result. Hence, the second hybrid method is brought into the considerations.

### *4.4.2 Seeded Region Growing Based on Three-Dimensional Edge Detection*

The second hybrid technique is a seeded region growing technique with the seed is chosen from the three-dimensional Canny edge detection result. This technique is unique because the growing seed is chosen from the three-dimensional edge detection image. The three-dimensional edge detection was used instead of the original Canny edge detection because it gave more defined and connected edge as discussed above in Section 4.2.

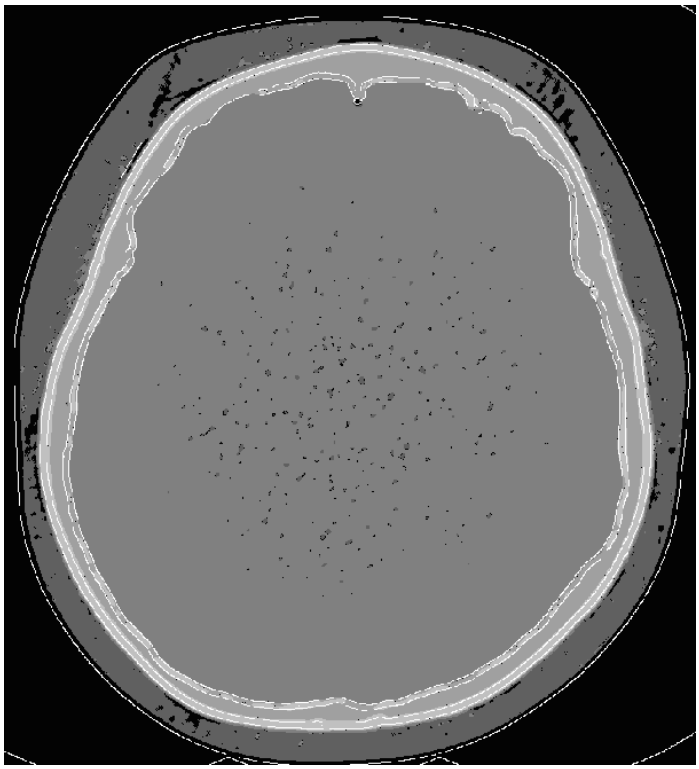

Figure 4.7. Seeded Region Growing using 3-D edge detection: slice 60

Figure 4.7 is the result of the second hybrid technique applied slide 60. Slice 60 was chosen to show because it is one of the earlier slice. It shows the image of the homogeneous brain with the skull. From the intentions to distinguishing between the cortical layers of the skull. Comparing to the seeded region growing in Figure 4.5, the regions are cleaned up as the brain and brain fluid are grouped together. The edges of the skull shows up much clearer. There are the distinct layers of the cortical bones.

# CHAPTER FIVE

# Conclusion

## *5.1 Summary*

For this thesis, different types of image segmentation techniques were study to automatically segment and differentiate tissues of the pediatric head phantom CT scans. Canny Edge detection and three-dimensional Canny edge detection was implemented to enhance the segmented results of threshold-based and region growing-based segmentation as the hybrid techniques. The archived result is that the hybrid techniques gives better detailed and more continuous of the inner cortical bone edges in human head phantom scan.

In the future, more efficient and accurate technique can be developed by integrating iterative algorithms onto the existing hybrid technique.

## *5.2 Future Works*

Even though the current implemented techniques have given more accurate edges and materials identification in testing results, there are still some artifacts that showed in some regions which do not belong to that respective region. These artifacts make the output image contain non uniform regions. Some future directions can be done to overcome challenges already seen and discussed in this work, and would help to get closer toward making the improvement in automated medical image segmentation.

### *5.2.1 The Statistical Approach*

One hard question is that of whether the higher number of boundaries mean better results in medical image reconstruction. For further study, a different approach to the problem by using different filter to smooth the input image might be

helpful. The investigation of the statistical information for choosing parameters for more accurate threshold are also needed.

### *5.2.2 Further Testing of Other Phantoms*

Further testing of the hybrid techniques should be performed on other phantoms and human data sets. Although these techniques work for this phantom, they might not work as well in the real segmentation of a human data set. The phantom is fixed and mounted rigidly while scanning through the gantry. In practice, a patient's motion is one of the most common reasons that artifacts appear.

### *5.2.3 Acceleration of the Time Processing*

One of the improvement of three dimensional Canny algorithm would be to reduce the time of processing and use multiple processors or cluster to accelerate the whole process of segmentation for larger scan size. Baylor University has the High Performance and Research Computing Service that manages a Kodiak cluster with hundreds of computing nodes. It will be meaningful to continue developing this work using the support high performance system provided.

# BIBLIOGRAPHY


Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence* (6), 679–698.

Clarke, B. (2008). Normal bone anatomy and physiology. *Clinical journal of the American Society of Nephrology 3*(Supplement 3), S131–S139.

Fabijanska, A., M. Kuzanski, D. Sankowski, and L. Jackowska-Strumillo (2008). Application of image processing and analysis in selected industrial computer vision systems. In *Perspective Technologies and Methods in MEMS Design, 2008. MEMSTECH 2008. International Conference on*, pp. 27–31. IEEE.

Gonzalez, R. C. and R. E. Woods (2007). *Image processing*, Volume 2.

Gücük, A. and U. Üyetürk (2014). Usefulness of hounsfield unit and density in the assessment and treatment of urinary stones. *World journal of nephrology 3*(4), 282.

Hsieh, J. (2003). *Computed tomography: principles, design, artifacts, and recent advances*, Volume 114. SPIE press.

Lee, T. S. (2006). Software-based gradient nonlinearity distortion correction.

Padhye, A. S. (2013). Impact of varied low resolution phantoms on intensity modulated proton therapy dose distribution.

Pbroks13 (2007). Cross-section of bone.

Prince, J. L. and J. M. Links (2006). *Medical imaging signals and systems*. Pearson Prentice Hall Upper Saddle River, New Jersey.

Ramesh, J., K. Rangachar, and G. S. Brian (1995). Machine vision. *Chapter 4-Image filtering*.

Schulte, R. W., R. P. Levy, T. S. Lee, M. Neupane, F. Shihadeh, D. Slusarczyk, K. E. Schubert, and J. D. Slater (2005). A system for functional proton radiosurgery. In *Proceedings of the International Meeting for Brain Mapping and Surgical Treatment Planning*.

Schultze, B., P. Karbasi, V. Giacomelli, T. Plautz, K. E. Schubert, and R. W. Schulte (2015). Reconstructing highly accurate relative stopping powers in proton computed tomography. In *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2015 IEEE*, pp. 1–3. IEEE.

Sharma, N. and L. M. Aggarwal (2010). Automated medical image segmentation techniques. *Journal of medical physics/Association of Medical Physicists of India 35*(1), 3.

Walad, K. P. and J. Shetty (2014). Traffic light control system using image processing. *International Journal of Innovative Research in Computer and Communication Engineering 2*, 288–293.