Wednesday, 12 December 2012

Computer Aided Detection of Pulmonary Embolism (PE) in CTA Images

A dissertation submitted by

Yousef Ebrahimdoost

In partial fulfilment of the requirements for the degree of Doctor of Philosophy in

Computer Science

Kingston University London

2012

Director of Study: Professor Jamshid Dehmeshki

ABSTRACT

Pulmonary embolism (PE) is an obstruction within the pulmonary arterial tree and in the majority of cases arises from a thrombosis that has travelled to the lungs via the venous system. Pulmonary embolism (PE) is a fatal condition which affects all age groups and is the third most common cause of death in the US. Computed tomographic angiography (CTA) imaging has recently emerged as an accurate method in the diagnosis of pulmonary embolism. Each CTA scan contains hundreds of CT images, so the accuracy and efficiency of interpreting such a large image data set is complicated due to various PE look-alikes and human factors such as attention span and eye fatigue. Moreover, manual reading and interpreting a large number of slices is time consuming and it is difficult to find all the pulmonary embolisms (PE) in a data set. Consequently, it is highly desirable to have a computer aided detection (CAD) system to assist radiologists in detecting and characterizing emboli in an accurate, efficient and reproducible manner.

A computer aided detection (CAD) system for detection of pulmonary embolism is proposed in CTA images. Our approach is performed in three stages: firstly the pulmonary artery tree is extracted in the region of the lung and heart in order to reduce the search area (PE occurs inside the pulmonary artery) and aims to reduce the false detection rate. The pulmonary artery is separated from the surrounding organs by analyzing the second derivative of the Hessian matrix and then a hybrid method based on region growing and a new customized level set is used to extract the pulmonary artery

(PA). In the level set implementation algorithm, a new stopping criterion is applied, a consideration often neglected in many level set implementations. In the second stage, pulmonary embolism candidates are detected inside the segmented pulmonary artery, by an analysis of three dimensional features inside the segmented artery. PE detection in the pulmonary artery is implemented using five detectors. Each detector responds to different properties of PE. In the third stage, filtering is used to exclude false positive detections associated with the partial volume effect on the artery boundary, flow void, lymphoid tissue, noise and motion artifacts. Soft tissue between the bronchial wall and the pulmonary artery is a common cause of false positive detection in CAD systems. A new feature, based on location is used to reduce false positives caused by soft tissue.

The method was tested on 55 data scans (20 training data scans and 35 additional data scans for evaluation containing a total of 195 emboli). The system provided a segmentation of the PA up to the 6th division, which includes the sub-segmental level. Resulting performance gave 94% detection sensitivity with an average 4.1 false positive detections per scan. We demonstrated that the proposed CAD system can improve the performance of a radiologist, detecting 19 (11%) extra PE which were not annotated by the radiologist.

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENT

GLOSSARY

AA	Aortic artery
ABM	Area-based matching
AUC	Area under ROC curve
CAD	Computer aided detection
CFL	Courant-Friedrichs-Lewy
CT	Computed tomography.
CTA	Computed tomography angiography
CTPA	Computed tomography pulmonary angiography
DFB	Distance field boundary
DFS	Distance field source
DICOM	Digital imaging and communications in medicine
DT	Distance transforms.
DVT	Deep vein thrombosis
EM	Expectation maximization
FBM	Feature-based matching
FMM	Fast marching method
FN	False negative
FP	False positive
FROC	Free-response receiver operator characteristic
HCT	Helical computed tomography.
HU	Hounsfield units

IVC	Inferior vena cava
LA	Left atrium
LSM	Level set method
LV	Left ventricle
MDCT	Multi-detector computed tomography.
MR	Magnetic resonance
MRA	Magnetic resonance angiography
MRI	Magnetic resonance imaging
MSFM	Multi-stencil fast marching
NN	Neural network
PA	Pulmonary arteries
PDE	Partial differential equation
PE	Pulmonary embolism
PSF	Point-spread function
PV	Pulmonary vein(s)
PVE	Partial volume effect
RA	Right atrium
RMS	Root mean square
ROC	Receiver operator characteristic
RV	Right ventricle
SE	Structuring element
SI	Shape index
SVC	Superior vena cava
TN	True negative
TP	True positive
VED	Vessel enhancing diffusion
V/Q	Ventilation/Perfusion

LIST OF FIGURES

1. INTRODUCTION

1.1. Pulmonary Embolism

A Pulmonary Embolism (PE) is a sudden blockage in a lung artery (pulmonary artery) that happens when an embolus blocks blood flowing through an artery that feeds the lungs [6]. Typically, a blood clot first forms in an arm or leg and then travels to the lung. The mechanism of formation of blood clots in the vessels is identified as thrombosis which can be formed by a decrease in the velocity of blood [7], [8]. An embolus is the result of a breaking down of the thrombus which can travel and block the blood flow in vessels. Pulmonary embolism is the result of an embolus traveling through the inferior vena cava and the right heart and being trapped in the pulmonary arteries. There are several factors which can make someone more likely to develop a blood clot that can travel to the lung. Obesity, heart disease, travel (such as sitting in an airplane or a long car trip), pregnancy, cancer, etc, are the main causes of blood clots in lung. Some other factors such as smoking, hormone therapy, oral contraceptives, family members with thrombosis or embolism and cancer can be high risk factors for PEs [9]. The symptoms can be:

- Chest pain
- Increased or irregular heart beat
- Dizziness
- Cough or coughing up blood
- Difficulty catching breath
- Rapid breathing

PE is an extremely common and highly lethal condition, and is a leading cause of death in all age groups, though the risk of developing them is high in the elderly. In most cases, they can be identified as a complicated condition called deep vein thrombosis (DVT). In this situation, blood clots form in the deep veins of the body. These clots break free and then travel to the lung and finally block an artery. One in every 1,000 people in the UK develops DVT, and almost 10% of people with untreated DVT will develop a pulmonary embolism (PE).

(a)   (b)
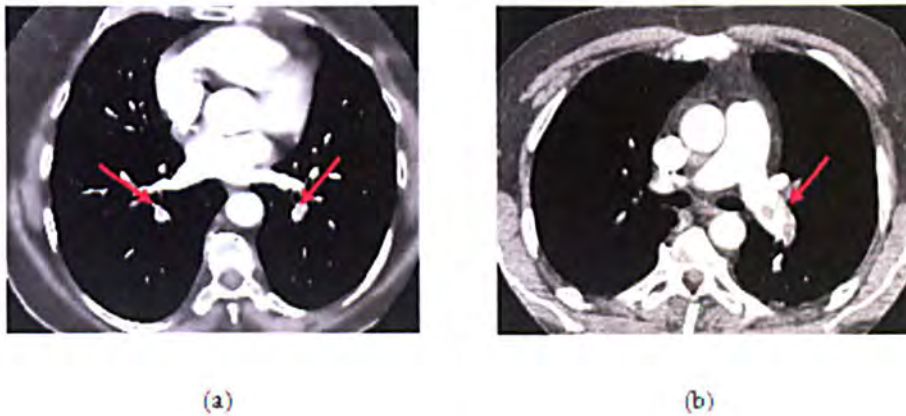
Figure 1-1: Pulmonary embolisms (arrowed) in different parts of the pulmonary artery :(a) Shows PEs in the left and the right segmental arteries. (b) Shows clots in the left main artery.

The third most common cause of death is in patients diagnosed with PE and the United States has a higher ranking in pulmonary embolism diseases, with more than 650,000 cases occurring each year [4]. Of these cases, 11% die in the first hour [1] and the untreated mortality rate of PE is estimated at 30% [2]. Thus, PE is a common disorder with a high morbidity and mortality. Therefore it is desirable to have an early and precise diagnosis of DVT and PE [3].

In patients with PE the symptoms are sometimes unspecific and in many patients are not present, so diagnosis is a challenging and difficult task, especially for those who have heart or lung disease. To find out the causes of PE symptoms, there are several clinical assessments or tests and image scanning methods to diagnose pulmonary embolisms.

1.2. Clinical Assessment and Test

Clinical assessment and clinical tests are the first and crucial step in the diagnosis of pulmonary embolisms [10], [11]. These assessments use signs and symptoms to estimate the probability of an embolism. The D-dimer test is a clinical test that measures the levels of a specific protein associated with the breakdown products of a blood clot. In the case of a normal test the probability of the existence of a pulmonary embolism is very low, but having high levels of the clot-dissolving substance D-dimer in the blood may increase the likelihood of blood clots. A disadvantage of this test is that the results of the test can be positive for many reasons, including injury, pregnancy, surgery and infection and hence it is not capable of specifically diagnosing an acute pulmonary embolism [12]. Ultrasound is the first imaging test in the diagnosis of blood clots and it is a fast and painless method which applies sound waves to image vessels in the leg. Although this test is fast, it is not very sensitive and it needs a person with high levels of experience to diagnose a blood clot [134].

Another diagnostic test of blood clots is venography which is rarely used for diagnosis. It creates X-ray images from veins and blood flow passing through them. Since veins are transparent in X-rays a special dye is injected into the blood to make it opaque [135].

Lung Scan:

Ventilation-perfusion scan (V/Q scan), pulmonary angiography, MR and CT are among the methods in the diagnosis of PE. The Ventilation-perfusion scan (V/Q scan) method applies a radioactive material to highlight ventilation (airflow) and flowing blood (perfusion) in the lung region. Inhaling a small amount of radioactive material and injecting it into a vein in the arm and then taking pictures of movement of air and blood flow, provide more accurate diagnosis of pulmonary embolisms.

Pulmonary angiography is a technique to see how blood flows through the vessels in the lung using a special contrast material (dye) and X-rays. In this procedure, a catheter is inserted inside the arteries; then dye is injected into the catheter and doctors can see live X-ray images which are taken from arteries. The dye helps doctors to diagnose a blockage inside the arteries. Recently, this method has been replaced by spiral CT angiography due to its cost and because it is invasive [13].

Magnetic resonance (MR) is another imaging procedure which can be used to examine pulmonary arteries in the lung and heart regions for detecting pulmonary embolisms [14]. In MRI, a powerful magnetic field, radio waves and a computer generate a detailed 3D image volume, which may be used with or without contrast material. Unlike the pulmonary angiography technique, MRI is a noninvasive imaging technique that produces detailed images of blood vessels and blood flow without inserting a catheter into the arteries. However, it is more expensive, has a lower spatial resolution than CT images and it is not without difficulty with respect to other imaging systems such as CT imaging [12].

X-ray computed tomography (CT), also known as computed axial tomography, is a medical imaging technique which uses X-rays to take multiple images and a computer to re-construct multidimensional views. CT generates a volume of data that can be manipulated by image analysis and has become an important tool and the gold standard in medical imaging to diagnose a large number of different diseases. CT has been identified as a noninvasive procedure for evaluating patients with suspicion of a pulmonary embolism over the last 15 years [17]. Compared with other techniques, this imaging technique is a simple, accurate and fast imaging technique [15], [16].

Recently, a new kind of CT scan called a spiral (or helical) CT scan is used to obtain cross-sectional pictures of the body. In this procedure, the X-ray machine rotates continuously around the body and during the spiral path, it generates cross-sectional pictures of the body. In comparing with a conventional CT scan, a spiral CT scan is faster and has higher resolution and can detect a smaller abnormal item. Since 1996, studies for evaluating spiral CT for the diagnosis of PE shows high sensitivities (53% to 92%) and high specificities (78% to 100%) [18], [19].

Computerized Tomographic Angiography (CTA image) is a CT scan of blood vessels which are injected by contrast material into the vascular system. CTA images have been established as the first imaging test for the diagnosis of pulmonary embolisms. In comparison with conventional pulmonary angiography, it has advantages such as low risk, direct imaging of blood clots and higher accuracy. In CTA images, details of the vascular anatomy can be rendered in three dimensions and the adjacent bony structures can be visualized. Due to the contrast material, the blood vessels can be seen as bright tubular structures (Figure. 1.2).
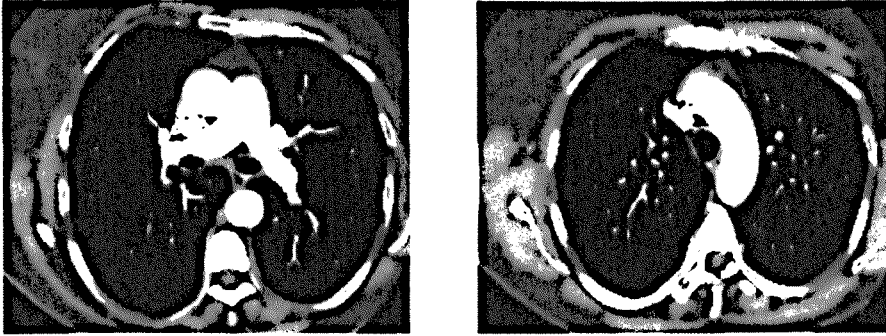
Figure 1-2: Axial CTA images show vascular vessels as bright tubular structures

A CT pulmonary angiogram (CTPA) is a medical diagnostic test that uses CT to create an image of the pulmonary arteries. Computed Tomographic Pulmonary Angiography has recently emerged as an accurate and alternative diagnosis tool for PEs [20], [21]. They can be used to perform accurately and quickly to portray vessels to the level of sixth-order of pulmonary artery branching [118]. In spite of some concerns about low sensitivity of CTPA, it is an attractive tool for clinicians because the results of CTPA are clear (positive or negative) and can be used for identifying alternative non-thrombotic causes of patients' symptoms in chest diseases [22]. Multi-row detector CTPA, which takes less time in image acquisition, has been evaluated to have high sensitivity (100%) and high specificity (89%) in the investigation of patients with suspected PEs in comparison with conventional pulmonary angiography [23].

Each CTPA study contains hundreds of CT slices, so the accuracy and efficiency of interpreting such a large image data set is complicated by various PE look-alikes and also limited by human factors, such as attention span and eye fatigue. Moreover, manually reading and interpreting a large number of slices is time consuming and it is difficult to find all the pulmonary embolisms (PE) in a data set [24]. Consequently, it is highly desirable to have a computer aided detection (CAD) system to assist radiologists in detecting and characterizing emboli in an accurate, efficient and reproducible manner.

1.3. Computer Aided Detection (CAD)

Computer aided detection (CAD) is a main technique in medical imaging which helps radiologists and clinicians to decrease observational oversights and false negative rates in interpreting medical images. The concept of CAD is different from "Computer aided diagnosis" which refers to software which analyzes diseases.

Several Computer Aided Detection (CAD) systems have been proposed to automatically detect pulmonary embolisms, most of which are based on prior pulmonary vessel segmentation and detection of the embolus and few system have been developed without using pulmonary vessel segmentation. Those systems, which use the pulmonary vessel as a search area, may increase false positive detection rates. Since some cases of false positives such as flow voids happen in veins, it is reasonable to limit the search area by separating pulmonary arteries from veins to reduce false positive detection. In fact the reduction of the search space contributes to both the efficiency of the detection process and reduction of false positives.

Our contribution is to present a CAD system for the automatic detection of pulmonary embolisms in CTA images which can help a radiologist to improve his/her job and increase the radiologist's certainty to detect PE. The capability of the proposed system is based on segmentation of the pulmonary artery to reduce the search area while aiming to reduce the false positive detection rate. Since some false positives are identified in other pulmonary vessels (such as veins), separation of the pulmonary artery from other vessels is likely to reduce false positive detections. In fact the reduction of the search space contributes to both the efficiency of the detection process and the reduction of false positives.

In the proposed CAD system: 1) Images are collected. 2) A nonlinear diffusion filter is used as a pre-processor on raw data. 3) The lung and heart region are extracted to reduce the search region for extracting the pulmonary artery (PA). 4) Before extracting the pulmonary artery a feature map based on eigenvalues of the Hessian matrix is created to remove the connectivity between the pulmonary artery and other organs that have the same intensity. 5) The pulmonary artery is extracted in two stages. In the first stage the major pulmonary artery is extracted in the heart region using a new customized level set method. In the level set method a new stopping criterion is used, a consideration which has usually been neglected in level set implementations. In the second stage, the peripheral arteries are separately segmented into left and right lung regions by a region-

growing algorithm. 6) Pulmonary embolism (PE) candidates are selected using five detectors that respond to different properties of the PE. 7) The three dimensional features (connected components) inside the segmented artery are analyzed to identify PEs and exclude false positives.

1.4. Outline of the Thesis

The organization of thesis is as follows. Chapter 2 gives some medical definitions and background of PE disease and a literature review of previous research in pulmonary embolism detection. Pulmonary artery segmentation is described in chapter 3. This chapter first reviews existing research in pulmonary artery segmentation and then describes a pre-processing algorithm based on non-linear diffusion filters, lung and heart segmentation algorithms, feature map extraction and finally pulmonary artery segmentation algorithms. Chapter 4 describes pulmonary embolism (PE) detection and removal of false detections. Discussion and experimental results and evaluation of PA segmentation and PE detection are presented in chapter 5 and finally conclusions are made in chapter 6. The literature review for each major component of the proposed CAD system is discussed within the related chapter.

2. MEDICAL DEFINITIONS AND LITERATURE REVIEW

This chapter provides the medical background associated with different kinds of pulmonary embolism, the causes of pulmonary embolisms, the pitfalls of detection and the medical imaging systems. It then provides a detailed description of previous research on CAD systems for detecting pulmonary embolisms. In the final section the proposed CAD system will be introduced.

2.1. Medical Machinery

2.1.1. CT Scan

A CT (computed tomography) scan (also known as CAT - Computer Axial Tomography), is a medical imaging procedure which applies a process of producing a two-dimensional image of a slice through a 3-dimensional object (Figure 2-1). The medical device which is used to generate CT scans utilizes X-rays and uses digital geometry processing to generate a 3D image of the inside of an object. The 2-dimensional X-ray images, which are taken by rotating around a single axis, are combined to generate a three dimensional (3D) image.
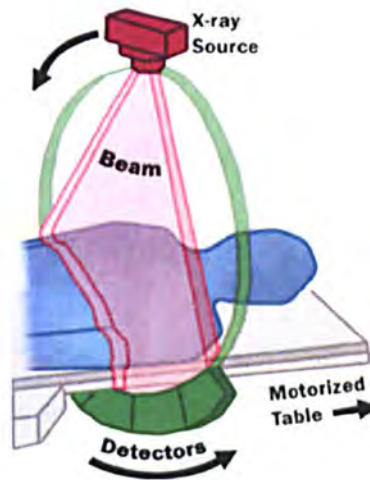
Figure 2-1: An X ray source sends out beams by rotating around a patient and their strength or attenuation is measured by a detector (taken from http://www.fda.gov/radiation-emittingproducts)

How a CT scanner works

A CT scanner is an X-ray machine from which several beams are simultaneously been sent through the human body from different angles. Inside the CT scanner there is an X-ray detector and the X-rays from the beams are detected by this detector after they have passed through the body and their strength or attenuation is measured (Figure 2-1). The absorption of the beams depends on the density of tissue: the less dense tissue such as the lungs, absorb less energy and beams passing these tissues will be stronger while beams that pass through denser tissue such as bone will be weaker.

At each orientation of beam sources around the body, the attenuation of the X-ray beam is measured and then by using a mathematical procedure of reconstruction (the Inverse Radon Transform) these projections are combined to recover the original radiographic densities at every location in the volume. The estimated radiographic densities are then computed for any volume intervals (voxels) and are expressed in Hounsfield units (HU). The Hounsfield unit transform is a scale transformation system in which the original

linear attenuation coefficient measurement is transformed into one in which the radio density of water in this system is identified as zero Hounsfield units (HU).

The Hounsfield unit scale is measured for different materials by:

$$HU = \frac{\lambda_x - \lambda_{water}}{\lambda_{water}} * 1000 \tag{1}$$

where, λ_x is the linear attenuation coefficient for material x. Table 2-1 shows the Radio-densities in Hounsfield units of various materials.

Table 2-1: Hounsfield unit's scale of various materials

Materials	HU
Air	-1000
Lung	-700
Soft Tissue	-300-100
Fat	-84
Water	0
Blood	30-45
Muscle	40
Bone	700-3000

2.1.2. MRI Scan

MRI (Magnetic Resonance Imaging) is a new imaging technique that uses magnetic and radio waves. The radio waves, which are many more times stronger than the magnetic

field of the earth (about 10,000 -30,000 times stronger), are sent through the patient's body. From measurements of the magnetic strength of the returned signal, a picture of the material densities can be created by the computer. In an MRI scanner the information of an object is usually extracted by changing the timing of the radio-wave pulses and objects surrounded by bones, usually give clearer pictures. This method of imaging device is particularly useful for imaging the brain and spinal cord diseases.

2.1.3. MRI Scanner Vs CT Scanner

MRI and CT scanners are the two most common medical imaging techniques to scan patients for determining the severity of certain diseases. Their application is based on parts of body that the radiologists or doctors need to see.

- CT scanners cannot show ligament or tendon disease; MRI scanners show more detail of these diseases.
- MRI scanners show better pictures than CT scanners for looking at the spinal cord.
- For visualizing lungs and organs in the chest cavity CT scans are better than MRI scans.
- Organ tear, broken bones, organ injury and broken vertebrae are better and more quickly seen on CT scans.
- CT scans are better for looking at cancer, bleeding in brain and inflammation of the lungs.

2.2. Pulmonary Embolism in CT Images

In CT images a pulmonary embolism can be identified as a dark region inside the segmented artery and clinical presentations are generally categorized into four different classes based on expected severity of the pulmonary artery occlusion. Each class is described in the section below.

2.2.1. Massive Pulmonary Embolism

Massive pulmonary embolism is usually caused by large emboli. It blocks a major part of the pulmonary artery and the results for the patients with this kind of PE are collapse and systemic shock. An example of this kind of PE can be large emboli blocking the main branches of the pulmonary artery, or a complete blockage of a left or right pulmonary artery. Massive PE usually causes hypotension and often results in significant mental status change for patients who are sweaty and pale. Figure 2-2 shows two massive pulmonary embolisms in CTA images of two patients.
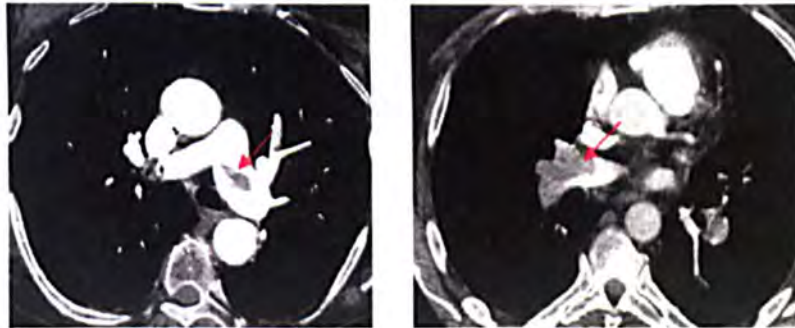
Figure 2-2: A large pulmonary embolus in the left and right branches of the pulmonary artery.

2.2.2. Acute Pulmonary Infarction

Acute pulmonary infarction is a peripheral blockage in the pulmonary artery branches. Approximately 10% of patients who present with hemoptysis, acute pleuritic chest pain or shortness of breath may have this kind of occlusion and it is difficult to diagnose. Figure 2-3 shows acute pulmonary infarctions for two patients.
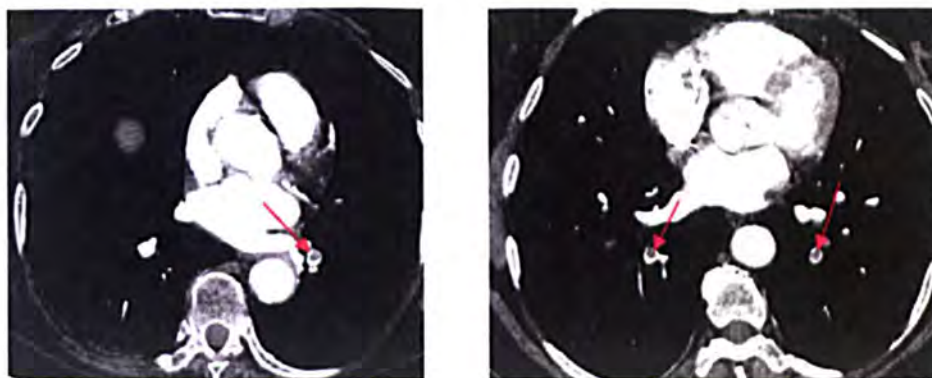
Figure 2-3: shows acute pulmonary emboli in the left and right peripheral pulmonary arteries

2.2.3. Acute embolism without infarction

This kind of PE has none of the specific symptoms as associated with massive PE, such as shortness of breath, cough, and pleuritic chest pain. Since lungs have a dual blood supply from pulmonary this kind of PE is unusual, occurring in approximately 10% of PE cases.

2.2.4. Multiple pulmonary embolisms (chronic and acute)

In general, pulmonary embolisms can be classified into two types: chronic and acute (Figure 2-4). Chronic pulmonary embolism, which is mainly a result of incomplete resolution of a pulmonary obstruction, may occur with previously documented history of pulmonary emboli over years. Patients with chronic pulmonary embolisms usually have hypertension and cor pulmonale, which is abnormal enlargement of the right side of the heart. The symptoms of chronic pulmonary embolism are not specific and are based on the development of pulmonary hypertension.

Unlike chronic pulmonary embolism, acute multiple PE may occur with no previously documented history of pulmonary emboli and have widespread obstruction of the pulmonary circulation with clots. Patients with acute multiple PE have a progressive shortness of breath and no continuous chest pain.

The shape of acute pulmonary embolism in CT images generally defines a hole or concavity in vessels, but the chronic emboli generally appear in vessel walls without any concavity in the vessels.
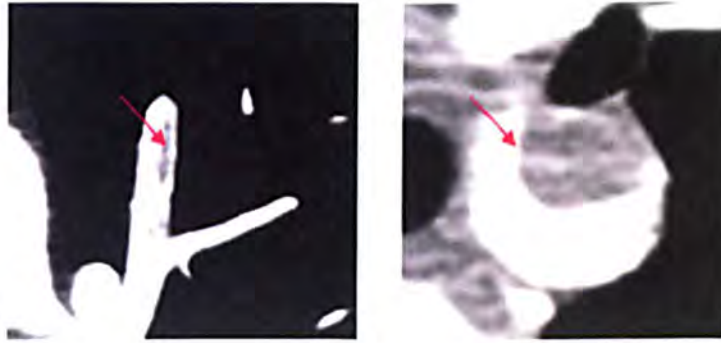
Figure 2-4: shows two different kinds of pulmonary embolism, chronic (left) and acute (right)

2.3. Pulmonary Embolism Pitfalls

The following circumstances are potential pitfalls to a reliable diagnosis:

1) Lymphoid tissue, which is located between a bronchus and an artery against the vessel wall and has the same intensity of PE, might be interpreted as a PE

2) The partial volume effect at the boundary of a vessel and bifurcations

3) Image noise associated with low dose and contrast

4) Streak artifacts made by beam hardening near the superior vena cava

5) A poorly enhanced vein. The scans are optimized for enhancement of the arteries because pulmonary embolism is arrested in the pulmonary arteries. The veins are often poorly filled with contrast material and this filling defect in the veins looks similar to a pulmonary embolism and might be misdiagnosed

6) Cardiac motion, which may make movement artifacts in CT images

7) Parenchyma disease causes a difference between intravascular and extra vascular contrast.

2.4. Causes for Pulmonary Embolism

The most common cause of pulmonary embolism is a blood clot which commonly originates in the deep veins of the legs. The clot is created when the blood flow slows or stops when a person stays in one position for a long time. Those patients who are recovering from major surgery are usually at risk. Sitting for long time periods without moving such in aeroplane journeys can also increase the risk of PE.

The sudden blocking of an artery of the lung (pulmonary embolism) is not the only cause of blood clots. Other effects that can also lead to pulmonary embolism are:

Fat: During bone surgery a long bone is usually fractured and fat from within the marrow of a broken bone escapes into the blood flow and forms an embolus.

Amniotic fluid which happens during a tumultuous childbirth and is forced into the pelvic veins can be a potential factor to form an embolus.

Air bubbles: Air bubbles can result in the formation of emboli if a catheter in one of the large central veins is mistakenly opened to air. During operations on veins and when a person is being resuscitated an embolus may also be formed by air bubbles.

A foreign substance: Foreign material such as talc injected by drug users can form an embolus and travel to the lungs.

Infected material: Infected material can also lead to the formation of emboli in the lung region associated with intravenous drug use, certain heart valve infections, and inflammation of a vein with blood clot formation and infection.

Cancer cells: Emboli might be created by the breakdown of a cancer cell which is released into the blood stream.

2.5. Lung and Heart Anatomy (Region of Interest for Searching PE)

The lung regions (left and right lungs) which reside between the abdomen and neck are essential organs of respiration in humans, providing a very large surface area to enable the exchange of oxygen and carbon dioxide between the body and the environment. The lung region provides oxygen for the bloodstream via airways which comprise the nose, the larynx, the pharynx, the trachea, the bronchi and the bronchioles. The heart region is located between left and right lung in the middle of the chest, behind and slightly to the left of the breastbone. The pulmonary artery (PA) carries the deoxygenated blood from the heart ventricle to left and right lungs for oxygenation. The oxygenated blood travels from lungs back to the heart via the pulmonary veins (PV) and then to rest of the body. The cardiovascular system is comprised of the following components for blood circulation:

- **Superior Vena Cava:** Superior vena cava caries the oxygen-poor blood from the upper part of the body into the right atrium.

- **Inferior Vena Cava:** Inferior vena cava carries the oxygen-poor blood from lower part of body into the right atrium.

- **Right Atrium:** The oxygen-poor blood returning from body is collected by the right atrium and then is forced into the right ventricle.

- **Right Ventricle:** Right ventricle collect the oxygen-poor blood from right atrium and forces it into lung region.

- **Pulmonary arteries:** Pulmonary arteries (left and right pulmonary artery) carry the blood from right ventricle to left and right lungs to pick up oxygen.

- **Pulmonary Veins:** Pulmonary veins (left and right) carry the oxygenated blood from lung region to left atrium.

- **Left Atrium:** The oxygen-rich blood returning from lung is collected by left atrium and then is forced into the left ventricle.

- **Left Ventricle:** Left ventricle is a strongest chamber in heart region to push the blood via aortic artery to other parts of body.

- **Aortic Artery:** The aortic artery (descending and ascending) carry the oxygen-rich blood form heart region to other parts of body.

Figure 2-5 shows a simplified cardiovascular system and Figure 2-6 is an illustration of the anatomy of the heart showing three different CTA cross-sectional slices.
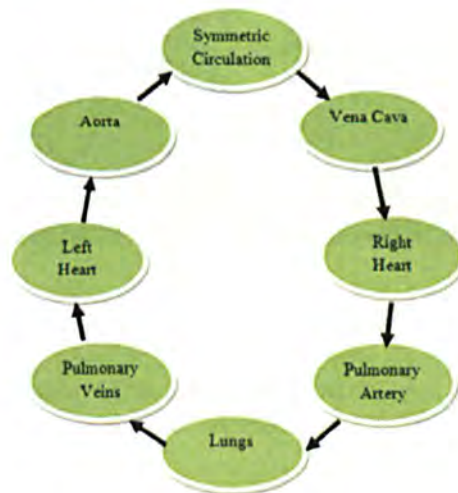
Figure 2-5: shows the simplified cardiovascular system

(a)

(b)
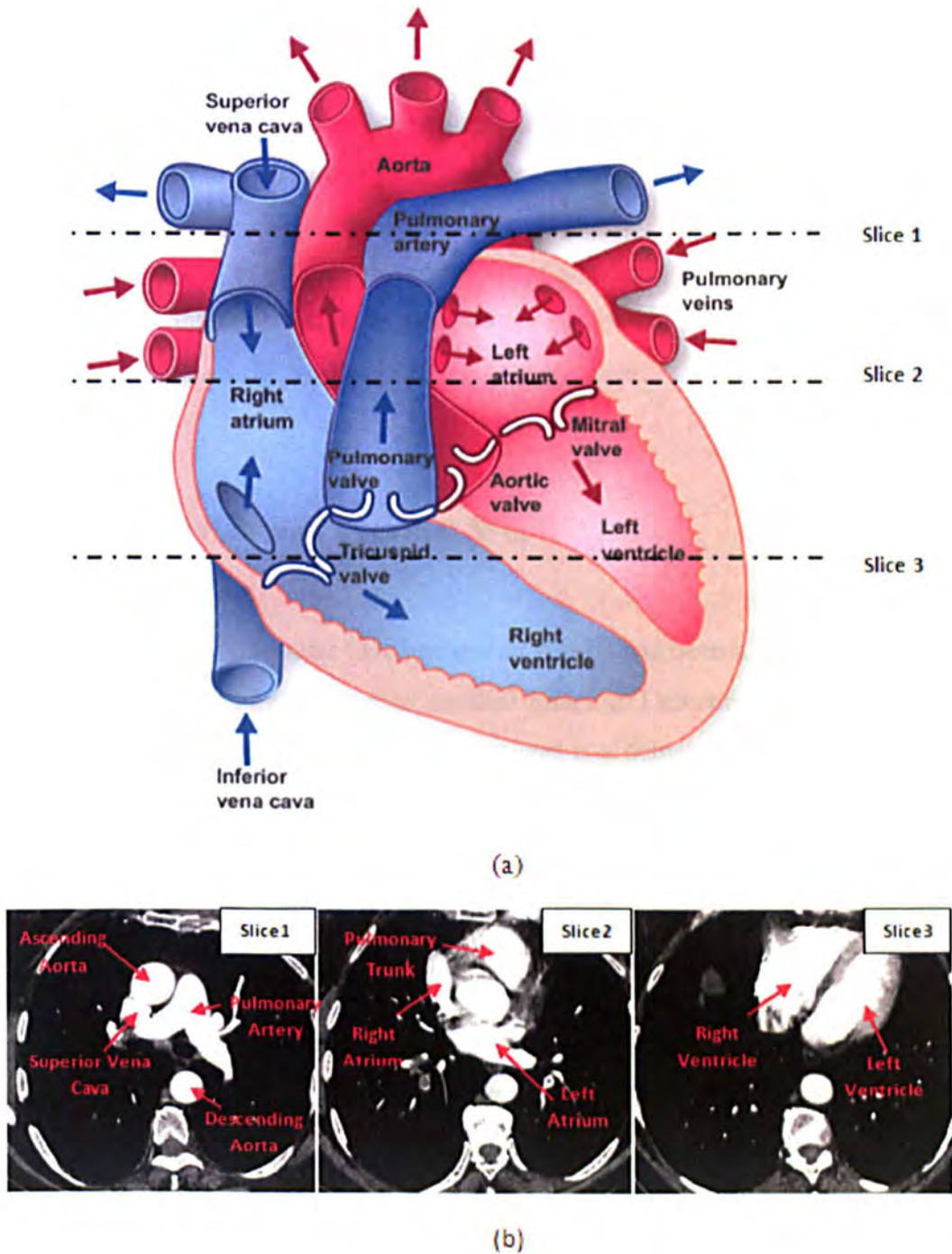
Figure 2-6: (a) shows the anatomy of the heart region. (b) Shows the cross section in three different CTA slices

2.6. Protocol of DICOM

One of the initial steps in medical image processing is to read data from stored CT image formats. CT images are stored in files with a DICOM format. DICOM, which stands for digital imaging and communication in medicine, defines a procedure of communicating and comparing different types of digital medical imaging devices or software. It is a common standard format for storing image data captured by medical imaging devices such as CT and MRI. It is also an extensive set of standards that support tasks such as image printing and transmission. It is possible for equipment to communicate remotely through a network or a media by using the DICOM standard.

2.7. Literature Review and Previous Work

Over last three decades, computer aided detection (CAD) has been a primary tool for the automatic detection of many diseases in medical imaging. Detection of abnormalities in the chest area across a large number of CTA images is a demanding job for doctors and radiologists. Because of the highly challenging nature of this task in interpreting pulmonary disease, a image analysis research has focused on developing methods to detect pulmonary diseases. Pulmonary embolism (PE) is a common and a fatal pulmonary disease in many countries and across all age groups, and several CAD systems have been developed which automatically help the radiologists detect pulmonary embolism (PE). The performance of a CAD system will depend strongly on the characteristics of a PE, such as their size distribution, the diameter of the artery being blocked, whether patients have other pulmonary diseases, and the quality of the CT scan which can be degraded by motion artefacts and incomplete circulation of injected contrast material. This review will summarise the studies that have been reported on the automatic detection of PE.

Automatic detection of PE is a more challenging task than for other pulmonary diseases because the pulmonary arterial system comprises an extensive network containing a wide range of vessel diameters. Y. Masutani et al [25] proposed a fully automated method for computerized detection of pulmonary embolism in CTA images based on volumetric image analysis. They first extracted pulmonary vessels and then analyzed several three

dimensional features inside the extracted vessels. The features that were analysed are based on local contrast using mathematical morphology to derive measures of the degree of curvi-linearity (based on second derivatives), vascular size and geometric features such as volume and length. They tested their CAD system on 19 clinical data sets (11 positive and 8 normal) and achieved 7.7 false positives per data set at 100% sensitivity, and 2.6 false positives at 85% sensitivity, when the PE volume was between 16 and 64 mm^3. They tested on only 11 data sets containing 21 PE samples, which is too small to cover the variety of different PEs.

Zhou et al [26] developed a preliminary investigation of computer aided detection of pulmonary embolisms in CTPA images. They extracted the vessels using an adaptive three-dimensional (3D) voxel clustering method based on expectation-maximization (EM) and detected true PE by applying a rule-based false-positive (FP) reduction method. In this preliminary study, they considered 14 patients with positive PE. In the cases without lung disease, their method detected 92.0% of proximal emboli and 77.8% of sub-segmental PE, with an average of 18.3 false positives per patient while, in the cases with extensive lung disease, 66.7% and 40.0% of the PEs were detected with an average of 11.4 false positives per case.

In a second study, Zhou et al [27] collected more data sets for improving the detection accuracy for PE and evaluating the performance of CAD systems by using 43 CTPA images. They applied 3D multiscale filters to enhance vascular structures including the vessel bifurcations and suppress non-vessel structures such as the lymphoid tissue surrounding the vessels and then they used a hierarchical EM estimation to segment the vessels by collecting the high response voxels at each scale. They searched for suspicious PE areas using a second adaptive multiscale EM estimation and using rule-based false positive (FP) reduction method, the true PEs based on the features of PE and vessels were identified. An experienced chest radiologist identified 435 PEs in artery branches (172 sub-segmental and 263 proximal to sub-segmental). These PE locations were identified as the "gold standard" in their CAD system. They achieved 81% and 78% sensitivity (proximal PEs) and 79% and 73% sensitivity (sub-segmental PEs) with 33 and 24 false positive detections per case.

Das et al [28] developed a CAD system and evaluated it's performance for automated detection of peripheral pulmonary embolisms in 33 consecutive multi-detector row CT

pulmonary angiography images. Their system identified a list of candidates and analyzed them according to shape, density, size and texture. The system showed 88% (164/186) sensitivity for the detection of segmental pulmonary embolisms and 78% (94/120) for the detection of sub-segmental pulmonary embolisms with 4 false positive detections per case. In a different study, Das et al [32] used 45 cases of which 29 cases had 213 PE in all vessels. They achieved 82% sensitivity with 3 false positive detections per case.

Jeudy et al [29] developed a CAD system using 22 data sets containing 251 PE (188 in segmental arteries and 63 in sub-segmental arteries) as a reference standard. Their system achieved a sensitivity of 80% for the segmental and 76% for sub-segmental with 1.8 false positive detections per dataset.

Schoepf et al [30] evaluated a CAD system for automated detection of pulmonary embolisms on multi-detector row CT studies of varying diagnostic quality in 23 patients (13 female, mean age 52) with PE and of 13 patients (all female, mean age 49) without PE. Their system identified correctly all 23 patients with PE and showed 92% (119/130) sensitivity for the detection of segmental pulmonary embolism and 90% (92/107) for sub-segmental pulmonary embolisms with 4.8 false positive detections per case. Maizlin et al [31] evaluated the same CAD system using large datasets (104 CTPA images) with 45 PE as standard references in 15 of the patients. The system identified 18 central and segmental PE and 8 sub-segmental in 8 patients, while it missed 14 proximal and 5 sub-segmental PE in 7 patients. Their system showed low sensitivity (57.8%) with 0.93 false positive detection rates per case.

Buhmann et al [33] developed a prototype computer-aided diagnosis (CAD) system. They applied artificial intelligence techniques for the detection of pulmonary embolism (PE). Forty multi-detector row computed tomography dataset which had good image quality were marked for suspicion of PE by six general radiologists using commercially available lung evaluation software. The 212 PE were identified as the "gold standard" by a radiologist. Of these, 65 (31%) were centrally located (in pulmonary trances, lobar, main, first order segmental arteries) and 147 (69%) were peripherally located (in the higher order segmental and sub-segmental arteries). Their CAD showed 74% sensitivity of

detection for central PE and 82% for peripheral PE with 3.85 false positive detections per case.

H. Bouma et al [34] proposed a new system for automatic detection of pulmonary embolisms in CT images. Their system consists of candidate detection, feature computation and classification. The candidate detection process focuses on inclusion of PEs and the exclusion of false positives. The feature computation is based on shape, intensity and size of PE and the shape and location of the pulmonary vascular tree. They also used several classifiers to optimize the performance of their system. The system was trained on 38 and evaluated on 19 other data sets. Their system showed 63% sensitivity on evaluation data sets with 4.9 false positive detections per data set, which enabled the radiologists to improve the number of PE detections by 22%.

Most CAD systems are based on prior pulmonary vessel segmentation and a few approaches have been established without using pulmonary vessel segmentation to detect the embolus. J Liang and J Bi [35] presented a fast effective toboggan-based approach for automated PE detection in CTPA images without extraction of pulmonary vessels. They achieved 80% sensitivity at 4 false positive detections per case.

The performance of all reviewed CAD systems depend on the characteristics of the pulmonary embolism (PE) such as it's size distribution, the diameter of the artery being blocked and the percentage of blockage related to a patients' conditions, such as whether they have other pulmonary diseases or not, and the quality of CT images. Those systems that used a large dataset, reported a high false positive rate and low sensitivity of detection. Since CAD systems for PE detection usually spend more time classifying the large number of initially detected suspicious regions as the true positive (TP) and false positive (FP) detections, but most of them generate a high rate of false positive detections.

To keep CAD systems as an effective tool to help radiologists for detecting PE, it is desirable to reduce false positive detection rates. Reducing the search area in the detection process can reduce the false positive detection rate in CAD systems. In fact, the reduction of search space contributed to both efficiency of the detection process and reduction of false positives.

Previous systems have used all the pulmonary vessels (i.e. all the vessels in heart and lung region). As a search area and as a consequence the false detection rate is high. Since

pulmonary embolism occurs only within the pulmonary artery, the proposed CAD system uses a small search area (pulmonary artery) as a search area aiming to reduce the rate of false detection.

2.8. The Proposed CAD System

Fully automatic computer aided detection (CAD) of pulmonary embolism is developed that aims to reduce the false positive detection rate by restricting the search area and increasing the sensitivity of detection by analyzing different 3D features. The proposed CAD system consists of three main steps:

- Segmentation (Pulmonary Artery)
- Detection (Pulmonary Embolism Candidates)
- Removal (False Detection)

The search area for pulmonary embolism is the pulmonary artery wherein the blood clot occurs, so in the first step the pulmonary artery is extracted (Chapter 3). Segmentation of the pulmonary artery (PA) is the most challenging task of the project as it has anatomically an elongated shape and is the same intensity with nearby vessels, such as the superior vena cava, the aorta and veins in the lung and heart region. Considering the difficulty of pulmonary artery extraction, an efficient algorithm for segmenting the pulmonary artery (PA) tree in three dimensional CTA images has been proposed.

In this algorithm, to reduce the search area the lung region (left and right) are first segmented from the raw data and the heart region is extracted by selecting the region between the lungs. A pre-processing algorithm based on the eigenvalues of the Hessian matrix is used to remove the connectivity between the pulmonary artery and nearby pulmonary organs. Before extracting the pulmonary artery, the superior vena cava (SVC) is extracted and removed from the heart region using an up-to-down region growing algorithm to facilitate the segmentation of pulmonary artery. The pulmonary artery tree is separately extracted in the heart and lung region. The major artery is first segmented by a region growing method initialized by a seed point which is automatically selected within the pulmonary artery trunk in the heart region, and then the segmentation result is refined

using a new customized 3D level set algorithm. The level set algorithm applies the output of region grower as the initial contour. The peripheral artery is segmented in left and right lung regions using a distinct region growing algorithm after enhancement by a vessel enhancing filter.

Pulmonary embolism (PE) can be identified as a dark region inside the segmented artery in the contrast enhanced CT images. The second step of the proposed CAD system uses five detectors based on intensity, geometric analysis, a morphological operator (top-hat transform defined on page 86), analysis of the perpendicular plane of the vessel centerline followed by ellipse fitting and a search for the disconnected part of the vessel to identify voxels inside the segmented artery as PE candidates. These five detectors are designed to respond to different properties of PE (Chapter 4).

In the third step, the PE candidates are filtered by extracting and analyzing different features to distinguish true PE from look-alikes (i.e. removal of false detections). Figure2-7 shows the flow chart for the major component of the proposed CAD system and the box numbering refers to the related chapters describing the major components.
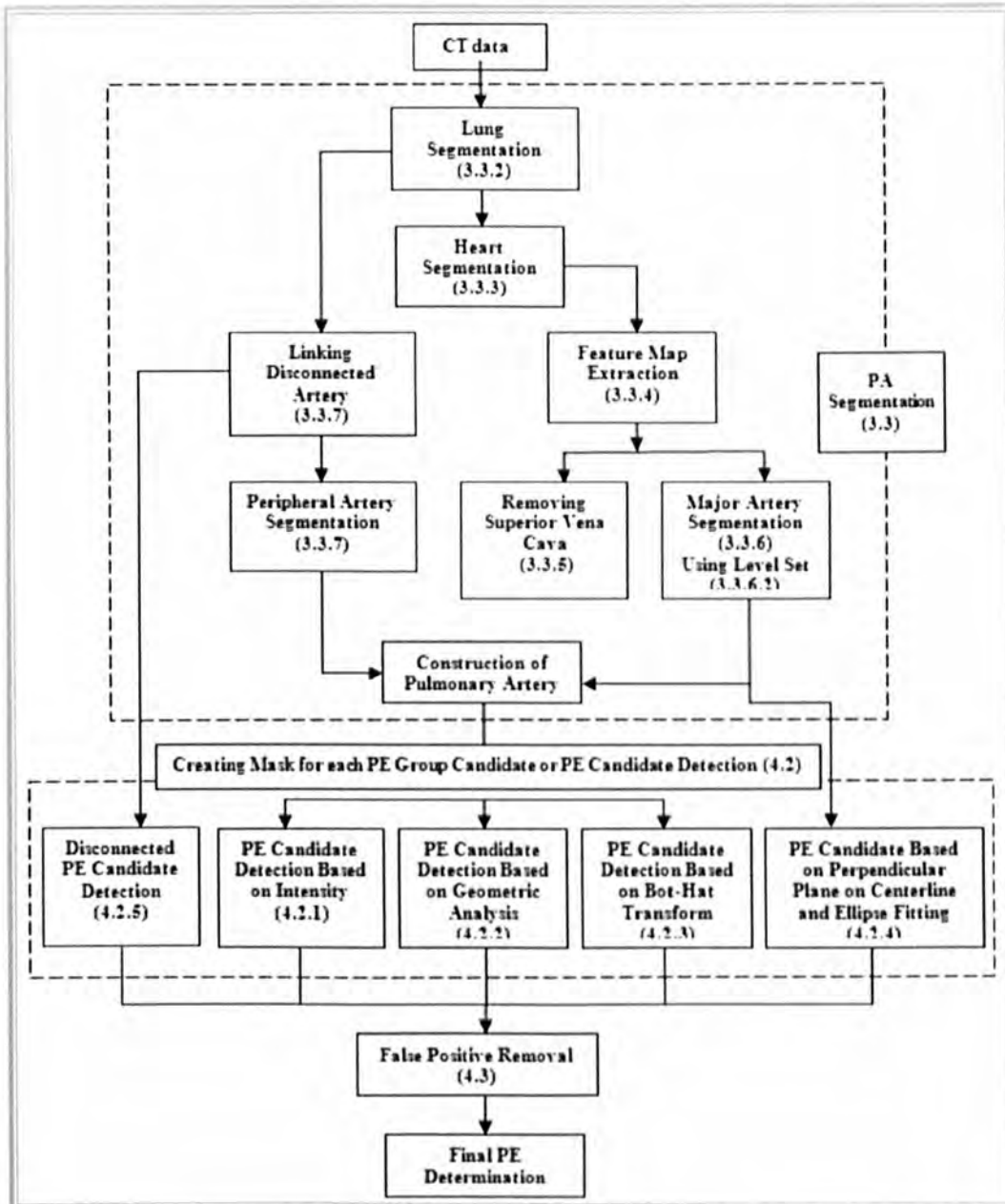
Figure 2-7: Flowchart of the proposed CAD system, which consists of PA segmentation, PE detection and false positive reduction.

3. PULMONARY ARTERY (PA) SEGMENTATION

3.1. Introduction

Pulmonary artery tree segmentation is an essential requirement for some medical applications, such as detection and visualization of pulmonary embolism (PE). Moreover, segmentation of the pulmonary artery tree is a very complicated and challenging task as it has multiple connections with other anatomical structures such as the pulmonary veins, the superior vena cava and the aorta, which can interfere with the segmentation of pulmonary artery boundaries. In this chapter the proposed algorithm for pulmonary artery (PA) segmentation in CTA images is described.

In this algorithm, first, the raw data is smoothed using a non-linear diffusion filter. Next, to reduce the search area, the lung regions are segmented and the heart region is extracted by selecting the region between the lungs. A pre-processing algorithm based on the eigenvalues of the Hessian matrix is used to remove connectivity between the pulmonary artery and other nearby pulmonary organs. The pulmonary artery tree is extracted in two steps. We first extract the major artery in the heart region using a region growing method initialized by a seed point which is automatically selected within the pulmonary artery trunk in the heart region and then refine the segmentation using a customized 3D level set

algorithm, using the output of a region grower as the initial contour. A new stopping criterion for the proposed level set algorithm is used an issue often neglected in many level set implementations. In the second step, the peripheral artery is segmented in left and right lung regions using a distinct region growing algorithm after enhancement by a vessel enhancing filter.

The organization of this chapter is as follows. In section 3.2 vessel segmentation approaches are reviewed and then the pulmonary artery segmentation algorithm will be discussed (section 3.3).

3.2. Vessel segmentation

Vessel segmentation is a key component of automated detection of disease in medical image processing. It is dependent on image modality, type of application and based on whether the method of segmentation is automatic or semi-automatic or it can be performed manually. Manual segmentation is usually accurate but is impractical for large datasets because it is tedious and time consuming. Automatic segmentation methods can be useful for clinical applications if they have: 1) the ability to segment like an expert; 2) exhibit excellent performance for diverse datasets; and 3) have a reasonable processing speed. There are different models and approaches in vessel segmentation and the main approaches are:

- Pattern recognition techniques
- Knowledge-based approaches
- Neural Network-based techniques
- Model-based methods

3.2.1. Pattern Recognition Techniques

The pattern recognition method is categorised into a region growing approach, mathematical morphology schemes, differential geometry-based approaches, a multi-scale method and skeleton-based approaches. This method incorporates automatic detection and classification of features or objects.

In the region growing method, an image is segmented by incrementally adding pixels to a region starting from some seed points based on special criteria. The criteria can be

similarity in pixel values or adjacency in terms of distance space [36]. Mathematical morphology schemes are based on the shape or form of an object using structure elements. Donizeli [37] used a hybrid method based on region growing and mathematical morphology schemes to segment large vessels. The mathematical morphology "top-hat" first was used to extract line-like structure and then by using a binary region growing technique, they removed background noise artefacts. In differential geometry-based approaches an image is treated as a hyper-surface and features are extracted by using the curvature and crest line of surface. Multi-scale image segmentation is a process by which an image will be segmented based on image resolution. In this method, large vessels are extracted from a low resolution image and a very small blood vessel is segmented at a high resolution level. In the skeleton-based method which is a known centreline detection technique, first the centreline of blood vessel is extracted and then by connecting these centrelines, blood vessel tree is created. Tozaki et al. [38] applied a centreline detection method to extract blood vessel and bronchus in lung region from CT images. They first segmented the image using a thresholding method and then separated the bronchus from the blood vessels based on anatomical characters and finally by using a 3D thinning algorithm the vessel centrelines were extracted.

3.2.2. Knowledge-Based Approaches

In general, in a knowledge-based system, an intelligence decision with justification is made using artificial intelligence tools based on posterior knowledge. In image segmentation tasks especially in vessel extraction, a knowledge-based algorithm applies the posterior knowledge to lead the segmentation process and to visualize the vessel structure. Several works have been developed to segment and analyse the vessel tree using a knowledge-based algorithm. Smets et al [112] delineated the blood vessel using a knowledge-based approach by encoding the general knowledge about blood vessels such as knowing high intensity for centreline of vessel, high intensity region are surrounded by parallel edges, etc. More related works can be found in [113], [114], [115].

3.2.3. Neural Network-Based Techniques

The Neural network (NN) has been developed in a wide range of applications especially in the medical image processing field. This method is applied as a classifier when the system is trained with a group of medical images and the trained system segments the target image. The neural network method is a very attractive method as it is easy to learn the system. That means that by choosing a good training set, including all possible features and objects, the system learns the boundary of classification in its feature space. Several works have been performed in the image segmentation field by applying the neural network method [116], [117].

3.2.4. Model-Based Methods

The main logical idea behind these methods is analyzing the image in a top-down fashion. Since the structures of interest or organs have a repetitive form of geometry, it is possible to find a probabilistic model for explaining the variation of the shape of the organ as constraints. In model based methods, these constraints are imposed as prior knowledge in the image segmentation process. These models are categorised in four approaches: 1) template matching, 2) deformable model, 3) parametric models, 4) generalized cylinders.

In image processing, the template matching algorithm is a technique for extracting parts of the image which match with a template image. This technique can be classified in two main groups, area-based matching (ABM) and feature-based matching (FBM). In feature-based matching algorithm those features are extracted which appear in the template pattern and are stored based on template information. While, in area-based matching techniques, every possible pixel/voxel position in the image (search image) are compared with the information of template data by least square or correlation methods.

The template matching algorithm is very capable in applications in which the search area and the image patterns are previously known. For example it can be used to control the quality of manufactured products [108] and in robotic science to navigate a robot [109]. In medical image processing, some works have been performed to extract the tubular shapes such as blood vessels by using the template matching algorithm [110], [111].

Deformable models (active contour models) describe curves or surfaces moving under internal and external forces and have been widely applied in medical image segmentation. These models are divided into two main categories: parametric deformable models, in which the curves and surfaces are explicitly represented in parametric form, and geometric deformable models in which a curve is implicitly represented as the zero level set of a higher dimensional constant function. The level set and fast marching methods are two main numerical approaches for evolving the curve (2D) and surface (3D) which have been used in a wide range of applications such as computer animation, some tasks in image processing fields such as optimal path planning [62], [63], shape representation [64], [65], registration [66], [67] finding distance fields from one and more points in computer vision [68] and segmentation [69]. In the proposed CAD system a geometric deformable model (level set algorithm) is used to segment the pulmonary artery, and a fast marching algorithm to extract the skeleton of segmented pulmonary artery. In the next sections these two models will be described in more detail.

3.2.4.1. Level Set Algorithm

Over the past decades, a large number of works on geometric deformable models in the sense of level set implementations have been proposed in the image processing field especially for image segmentation. Osher and Sethian [39] first introduced the level set method to capture the moving front. The main idea in level sets is to evolve a higher dimensional function, for which it's zero-level set shows the location of the propagating contour. In comparison with other deformable models, the level set method has more advantages: it can handle merging or splitting of the evolving contour, and numerical stability of the solution (the final result of evolution) is independent of initialization. In a level set algorithm, the evolving contour is the zero level set of a higher dimension of a constant function, which results in a complex computation. To remove this difficulty a narrow band level set method was introduced by Sethian.

In the level set formulation, letting $C\ (t)$:

$$C(t) = \{X | \phi(t, X) = 0\} \quad ; \quad X = (x_1, x_2, ..., x_n) \tag{1}$$

be the zero level set of function $\phi(t, X)$, then the evolution equation of the level set function ϕ can be written in the following partial differential equation:

$$\phi_t + F|\nabla\phi| = 0, \quad \phi(X, 0) = \phi_0(X) \tag{2}$$

where F is called the speed function of level set equation [39] and plays the most important role in the level set method which moves the contour in the normal direction and depends on the image data (external energy) and the level set function ϕ (internal energy). $\phi_0(X)$, is the initial level set function. The level set algorithm method is categorized in two main groups, edged-based and region based method.

Edge-Based Model

In the edge-based algorithm image gradient is considered as a constraint to stop the evolving contours/ surfaces on the boundary of the desired objects. In these models a general edge detector is defined by a positive and decreasing function g:

$$\lim_{x\to\infty} g(x) = 0 \ , \ g(0) = 1 \tag{3}$$

For example it can be as:

$$g(|\nabla u_0(x, y, z)| = \frac{1}{1+|\nabla G(x,y,z)*u_0(x,y,z)|^2} \tag{4}$$

where $G * u_0$, a smoother version of u_0, is the convolution of the image u_0 with the Gaussian filter:

$$G(x, y, z) = \frac{1}{\sqrt{\sigma}} e^{-|x^2+y^2+z^2|/4\sigma} \tag{5}$$

In edge-based models, the evolving contour relies on image gradient, so in practice the gradient of an image cannot be zero on the object boundary and it may cross the object boundary where the edge is weak. The popular edge-based level set models are introduced in [52], [70], [71], [72], [73].

Region-based model

Region-based models use the intensity of the region of interest by creating certain region descriptors as the main factor to move the interface. Unlike edge-based models, region-based models do not use the image gradient and are capable of segmenting objects with weak boundaries and the initial contours can be initialized anywhere in the image. The most popular region-based models are proposed in [74-80]. The region based models have often been inspired from the well-known region-based model proposed by Mumford and Shah [81]. This model is used for an image segmentation task by smoothing the image and setting a set of discontinuities points for representing the object boundaries. Chan and Vese proposed a new model for active contours to detect objects in an image, based on techniques of curve evolution, using the Mumford–Shah functional for segmentation. Their model can detect objects whose boundaries are weak and not necessarily defined by a gradient. The method is derived by minimization of an energy based-segmentation. They introduced an energy functional $F(c_1, c_2, C)$ defined by:

$$F(c_1, c_2, C) = \mu \cdot Length(C) + v \cdot Area\big(inside(C)\big) + \lambda_1 \int |u_0(x, y) - c_1|^2 dxdy +$$
$$\lambda_2 \int |u_0(x, y) - c_2|^2 dxdy \qquad (6)$$

where $\lambda_1 > 0$, $\lambda_2 > 0, \mu \geq 0, v \geq 0$ are constants and the first and second integral are taken inside and outside the closed curve C, respectively. By minimizing the above functional, they introduced the following region-based level set algorithm to segment an object in an image.

$$\frac{\partial \varphi}{\partial t} = \delta_\epsilon(\varphi) \left(div \left(\frac{\nabla \varphi}{|\nabla \varphi|} \right) - v - \lambda_1 (u_0 - c_1)^2 - \lambda_2 (u_0 - c_2)^2 \right)$$
$$\varphi(0, x, y) = \varphi_0(x, y) \qquad (7)$$

Where, c_1, c_2 are the mean intensity inside and outside the evolution curve respectively and $\varphi_0(x, y)$ is the initial contour. This model relies on intensity homogeneity. In fact the intensity of regions to be segmented must be statistically homogeneous (roughly a constant) in each region [82]. But intensity inhomogeneities are an issue which often happens in real world images and interfere with image segmentation. In order to overcome these difficulties Li at el [82] proposed a new region-based active contour

model that relies on intensity information in local regions at a controllable scale. In their method, first a region scalable fitting energy is defined which is based on the contour and two fitting functions. These two fitting functions locally approximate the image intensities on the two sides of the contour. The region scalable fitting energy is then incorporated into a variational level set formulation with a level set regularization term defined as:

$$\frac{\partial \varphi}{\partial t} = -\delta_\epsilon(\varphi)(\lambda_1 e_1 - \lambda_2 e_2) + v\, \delta_\epsilon(\varphi)\, div\left(\frac{\nabla\varphi}{|\nabla\varphi|}\right) + \mu\left(\nabla^2\varphi - div\left(\frac{\nabla\varphi}{|\nabla\varphi|}\right)\right) \tag{8}$$

where, δ_ϵ is Dirac delta function and e_1 and e_2 are two functions defined by:

$$e_i(x) = \int K_\sigma(y-x)|I(x) - f_i(y)|^2 dy \quad i = 1,2$$

$$f_i(x) = \frac{K_\sigma(x) * (M_i^\epsilon(\varphi(x))I(x))}{K_\sigma(x) * (M_i^\epsilon(\varphi(x)))}$$

$$K_\sigma(x) = \frac{1}{(2\pi)^{n/2}\sigma^2} e^{-|x|^2/2\sigma^2} \tag{9}$$

$$M_1^\epsilon(\varphi) = H_\epsilon(\varphi) \qquad M_2^\epsilon(\varphi) = 1 - H_\epsilon(\varphi)$$

$$H_\epsilon(x) = 1/2\left(1 + \frac{2}{\pi}\arctan(\frac{x}{\epsilon})\right)$$

To segment the pulmonary artery, we developed a new customized edge-based level set algorithm with a new stopping criterion which is described in more detailed in section 3.3.2.5-7.

3.2.4.2. Fast marching Method

The fast marching method is identified as a propagating interfaces algorithm which was proposed by Sethian in [83], [84]. If the front Γ is considered to propagate in one direction (inward or outward), then it has a monotonically decreasing or increasing propagation based on the sign of the speed function F. The position of the front is tracked by computing the arrival time $T(x, y, z)$ when it crosses the point $A(x, y, z)$. The motion

of the front Γ with speed function F which can be just positive or negative is defined as follows:

$$|\nabla T(x,y,z)|F = 1$$

$$T(\Gamma_0) = 0 \tag{10}$$

The arrival time of the initial position of the propagating front is considered as zero values. When the speed function F is only based on the position $A(x,y,z)$, then the above equation reduces to an equation known as the Eikonal equation. The Eikonal equation is a non-linear partial differential equation and is of the following form:

$$|u(X)| = F(X) \quad X \in \Omega \tag{11}$$

Where u is zero on the boundary of Ω (open set in $\mathbb{R}^n$).

In general, based on selecting different speed conditions in the propagating or tracking algorithms, the level set and fast marching method can be categorized in the initial value problem and boundary value problem respectively.

Boundary Value Problem	Initial Value problem
$\lvert\nabla T(x,y,z)\rvert F = 1$	$\phi_t + F\lvert\nabla\phi\rvert = 0$
$\Gamma(t) = \{(x,y,z)\lvert T(x,y,z) = t\}$	$\Gamma(t) = \{(x, y, z)\lvert\phi(x, y, z, t) = 0\}$
$F > 0$ or $F < 0$	F : Arbitrary

Both the boundary value problem and the initial value problem are special cases of a partial differential equation (PDE) known as Hamilton-Jacobi equation (HJ):

$$\alpha u_t + H(Du, x) = 0 \tag{12}$$

Where, $H(Du, x) = F|\nabla u| - (1 - \alpha)$, and α is either zero or one. Du represents the partials of u in each variable. By selecting $\alpha = 1$ and then $H = F|\nabla u|$ the level set is derived from the Hamilton-Jacobi equation and the fast marching algorithm is derived by choosing $\alpha = 0$.

Implementation of Fast Marching Algorithm

The fast marching algorithm is a procedure to find the arrival time $T(x, y, z)$ of the Eikonal equation:

$$|\nabla T(x, y, z)|F = 1 \tag{13}$$

at each point (x, y, z) given its neighbours as shown in Figure3-1.
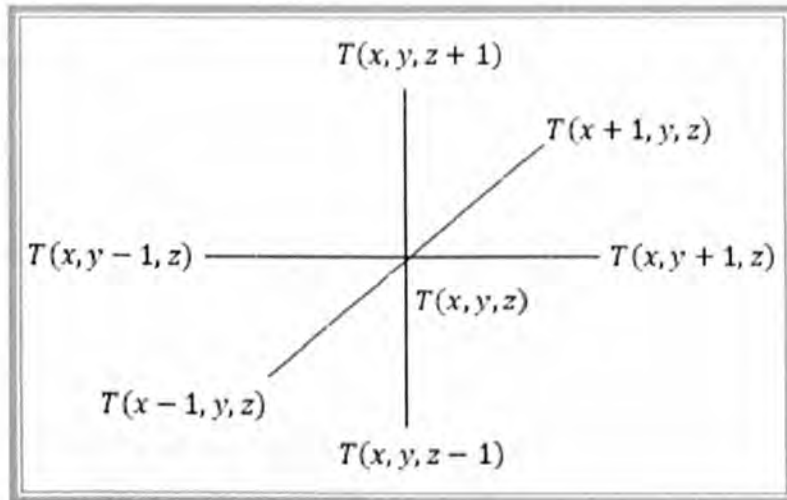
Figure 3-1 shows six neighbours of point $T(x, y, z)$

The Fast marching method (FFM) is appropriate for the Cartesian domain [85], [86] and triangulated surfaces [87], [88]. In this section, a numerical implementation of the fast marching algorithm in the Cartesian domain is described.

Sethian and Osher in [89] proposed the following discretizations of $|\nabla T|$ which gives a weak solution for the Eikonal equation.

$$Max\left(D_{ijk}^{-x}T,0\right)^2 + Min\left(D_{ijk}^{+x}T,0\right)^2 +$$

$$Max\left(D_{ijk}^{-y}T,0\right)^2 + Min\left(D_{ijk}^{+y}T,0\right)^2 +$$

$$Max(D_{ijk}^{-z}T,0)^2 + Min(D_{ijk}^{+z}T,0)^2 = \frac{1}{F_{ijk}^2} \qquad (14)$$

In order to find an easier but less accurate solution, the following upwind scheme was applied to discrete $|\nabla T|$. The upwind scheme refers to the direction of information propagation and it is defined in [90], [91].

$$Max\left(D_{ijk}^{-x}T,-D_{ijk}^{+x}T\right)^2 + Max\left(D_{ijk}^{-y}T,-D_{ijk}^{+y}T\right)^2 + Max\left(D_{ijk}^{-z}T,-D_{ijk}^{+z}T\right)^2 = \frac{1}{F_{ijk}^2} \quad (15)$$

Where, D_{ijk}^- and D_{ijk}^+ are the backward and forward finite difference schemes given by:

$$D_{ijk}^{-x}T = \frac{T_{i,j,k} - T_{i-1,j,k}}{\Delta x} \quad , \quad D_{ijk}^{-y}T = \frac{T_{i,j,k} - T_{i,j-1,k}}{\Delta y} \quad , \quad D_{ijk}^{-z}T = \frac{T_{i,j,k} - T_{i,j,k-1}}{\Delta z}$$

and

$$D_{ijk}^{+x}T = \frac{T_{i+1,j,k} - T_{i,j,k}}{\Delta x} \quad , \quad D_{ijk}^{+y}T = \frac{T_{i,j+1,k} - T_{i,j,k}}{\Delta y} \quad , \quad D_{ijk}^{+z}T = \frac{T_{i,j,k+1} - T_{i,j,k}}{\Delta z}$$

Where, Δx, Δy, Δz are the grid spacing in x, y, z directions, respectively. By substituting the above backward and forward finite difference scheme into equation (15) the following discretization scheme is derived.

$$\begin{pmatrix} Max \left(\frac{T_{i,j,k}-T_{i-1,j,k}}{\Delta x}, \frac{T_{i,j,k}-T_{i+1,j,k}}{\Delta x}, 0\right)^2 + \\ Max \left(\frac{T_{i,j,k}-T_{i,j-1,k}}{\Delta y}, \frac{T_{i,j,k}-T_{i,j+1,k}}{\Delta y}, 0\right)^2 + \\ Max \left(\frac{T_{i,j,k}-T_{i,j,k-1}}{\Delta z}, \frac{T_{i,j,k}-T_{i,j,k+1}}{\Delta z}, 0\right)^2 \end{pmatrix} = \frac{1}{F_{ijk}^2} \qquad (16)$$

By choosing:

$$T = T_{i,j,k} \qquad , \qquad T_x = Min(T_{i-1,j,k}, T_{i+1,j,k})$$

$$T_y = Min(T_{i,j-1,k}, T_{i,j+1,k}) \qquad , \qquad T_z = Min(T_{i,j,k-1}, T_{i,j,k+1})$$

The equation (16) is simplified as follows:

$$Max \left(\frac{T-T_x}{\Delta x}, 0\right)^2 + Max \left(\frac{T-T_y}{\Delta y}, 0\right)^2 + Max \left(\frac{T-T_z}{\Delta z}, 0\right)^2 = \frac{1}{F_{ijk}^2} \qquad (17)$$

For solving the Elkonal equation several algorithms have been proposed such as the iterative method which was proposed by Rouy in [90] and the single pass method or fast marching update scheme which was proposed independently by Sethian [86], Tsitisklis [92], and Helmsen [93].

The fast marching method for solving the Eikonal equation explicitly maintains a thin zone (narrow band) of candidate points that separates the grid points of known solutions from the unknown solutions. The main idea behind the fast marching method is that it establishes an order for selecting of grid points based on criteria and the arrival time $T(x, y, z)$ at any grid points depends on the smaller value of the neighbours. In fact, the fast marching method is a front propagating algorithm in one direction from smaller values for arrival time $T(x, y, z)$ to larger ones.

Fig 3.2 shows how the fast marching update scheme works in 2D discrete grid points. The algorithm starts with an initial grid point where the value of arrival time T is known and the value of F is considered to be known at every point. Four neighbours of the initial

point then will be updated by applying the upwind finite difference equation (15). The one out of four neighbours which has the smallest value of T is chosen as a known point and its four neighbours (six neighbours in 3D) are updated. This scheme moves on through the entire domain successively.

(a) (b) (c)

(d) (e) (f)
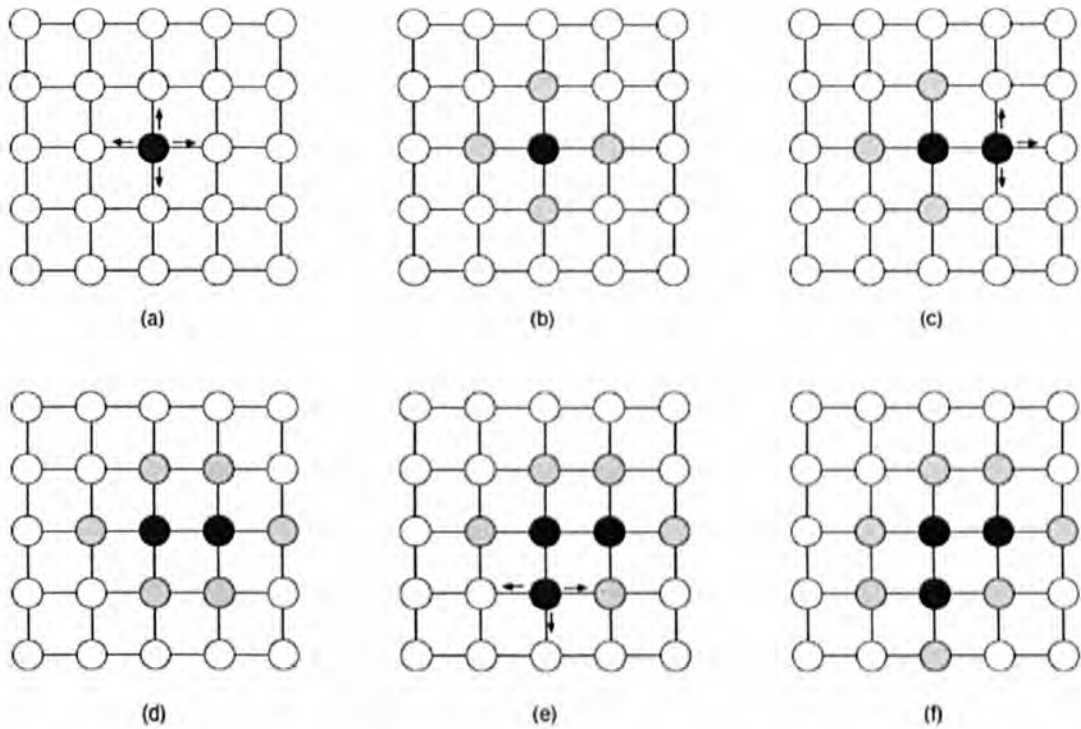
Figure 3-2 shows the fast marching propagation scheme in 2D. (a) Shows an initial point when T value is known. (b) Show four neighbours for initial point. (c) The smallest value of T is selected. (d) The front propagates. (e) and (f) The algorithm continues.

The fast marching algorithm is applied on grid points which are classified in three different types of points.

- **Alive points** or known points (black points in fig3.2) which are points where the value of T are known

- **Trial points** (gray points in fig3.2) these points are around the alive points and are known narrow band points. In the propagation process the narrow band points are updated to alive points while the narrow band expands.

- **Far away points** (white points in fig3.2) are all the points that have not yet been visited and these points convert to narrow band points during the propagation.

The fast marching algorithm is described in more detail in Appendix B.

High Accuracy Fast Marching Method

In the previous section, for solving the Eikonal equation using the fast marching method, the gradient ∇T was approximated by a first order finite difference scheme. In the high accuracy fast marching method, to approximate the gradient ∇T the second order finite difference scheme is used. The second order backward finite difference schemes to approximate the first derivatives are given by [91]:

$$D_{ijk}^{-x}T = \frac{3T_{i,j,k} - 4T_{i-1,j,k} + T_{i-2,j,k}}{2\Delta x}$$

$$D_{ijk}^{-y}T = \frac{3T_{i,j,k} - 4T_{i,j-1,k} + T_{i,j-2,k}}{2\Delta y}$$

$$D_{ijk}^{-z}T = \frac{3T_{i,j,k} - 4T_{i,j,k-1} + T_{i,j,k-2}}{2\Delta z}$$

and the second order forward finite difference scheme is given by:

$$D_{ijk}^{+x}T = -\frac{3T_{i,j,k} - 4T_{i+1,j,k} + T_{i+2,j,k}}{2\Delta x}$$

$$D_{ijk}^{+y}T = -\frac{3T_{i,j,k} - 4T_{i,j+1,k} + T_{i,j+2,k}}{2\Delta y}$$

$$D_{ijk}^{+z}T = -\frac{3T_{i,j,k} - 4T_{i,j,k+1} + T_{i,j,k+2}}{2\Delta z}$$

By substituting the above second order forward and backward finite difference scheme in equation (15) the following equation is derived.

$$\left(\begin{array}{l} Max\left(\frac{3T_{i,j,k}-4T_{i-1,j,k}+T_{i-2,j,k}}{2\Delta x},\frac{3T_{i,j,k}-4T_{i+1,j,k}+T_{i+2,j,k}}{2\Delta x},0\right)^2 + \\ Max\left(\frac{3T_{i,j,k}-4T_{i,j-1,k}+T_{i,j-2,k}}{2\Delta y},\frac{3T_{i,j,k}-4T_{i,j+1,k}+T_{i,j+2,k}}{2\Delta y},0\right)^2 + \\ Max\left(\frac{3T_{i,j,k}-4T_{i,j,k-1}+T_{i,j,k-2}}{2\Delta z},\frac{3T_{i,j,k}-4T_{i,j,k+1}+T_{i,j,k+2}}{2\Delta z},0\right)^2 \end{array}\right) = \frac{1}{F_{ijk}^2} \qquad (18)$$

By choosing:

$$T = T_{i,j,k}$$

$$T_x = Min(4T_{i-1,j,k} - T_{i-2,j,k}, 4T_{i+1,j,k} - T_{i+2,j,k})$$

$$T_y = Min(4T_{i,j-1,k} - T_{i,j-2,k}, 4T_{i,j+1,k} - T_{i,j+2,k})$$

$$T_z = Min(4T_{i,j,k-1} - T_{i,j,k-2}, 4T_{i,j,k+1} - T_{i,j,k+2})$$

The equation (18) is simplified as follows:

$$Max\left(\frac{3T-T_x}{2\Delta x},0\right)^2 + Max\left(\frac{3T-T_y}{2\Delta y},0\right)^2 + Max\left(\frac{3T-T_z}{2\Delta z},0\right)^2 = \frac{1}{F_{ijk}^2} \qquad (19)$$

Based on the high accuracy fast marching method, the value of T must be greater than T_x, T_y, T_z so, the equation (19) can be simplified as follows:

$$\left(\frac{3T-T_x}{2\Delta x}\right)^2 + \left(\frac{3T-T_y}{2\Delta y}\right)^2 + \left(\frac{3T-T_z}{2\Delta z}\right)^2 = \frac{1}{F_{ijk}^2} \qquad (20)$$

Multi-Stencil Fast Marching Method

The Multi-Stencil Fast Marching Method (MFMM), which was proposed in [100], is an improved version of the fast marching algorithm which has very accurate results for solving the Eikonal equation in Cartesian domains. This method calculates the solution at each grid point by solving the Eikonal equation along several stencils. These stencils which are centred at each grid point should cover all the neighbours of each grid point and if they are not aligned with the natural coordinate system the Eikonal equation must be derived by using directional derivatives and is solved by applying a higher order finite difference scheme. In 2-dimensional space all 8neighbours of each grid point are covered by two stencils whereas in 3-dimensional space all 26neighbours are covered by six

stencils. By computing the arrival time at each grid point by solving the Eikonal equation along each stencil, the solution that satisfies the upwind condition is picked.

For example in 3D for all stencils to find the first order approximation of directional derivatives the following equation is solved.

$$\sum_{i=1}^{3} Max \left(\frac{T(X)-T_i}{\|X-X_i\|}, 0 \right)^2 = \frac{1}{F^2(X)} \tag{21}$$

Where

$$T(X) = T_{i,j,k} \qquad , \qquad T_1 = Min(T_{i-1,j,k}, T_{i+1,j,k})$$

$$T_2 = Min(T_{i,j-1,k}, T_{i,j+1,k}) \qquad , \qquad T_3 = Min(T_{i,j,k-1}, T_{i,j,k+1})$$

And X_i is an alive (known) point at which the T_i is minimum.

Upwind Condition

By solving the quadratic equation, in the case of two solutions the minimum one does not satisfy the equation (21), so is rejected. The other solution $T(X)$ must be checked with the three adjacent neighbours T_1, T_2 and T_3. If the value of $T(X)$ is greater than the three neighbours the solution is accepted. Otherwise, $T(X)$ must be checked with the values of two remaining adjacent points. If it is greater than the values of two points, then the quadratic equation is solved based on their values and the maximum solution is selected. Otherwise the solution would be:

$$min \left(T_i + \frac{\|X-X_i\|}{F(X)} \right) \quad i = 1,2,3 \tag{22}$$

Distance Field

A distance field in 3D or 2D is the smallest distance from each point to the given set of objects. Each object can be a representation of data on a voxel grid or as an explicit representation of a surface. The distances between the point and an object can be

computed by using different metrics, such as the Euclidean or max-norm distance. A sign can be assigned to the distance field (signed distance field), if the given objects are closed or orientable. Distance fields are used in many applications such as computer graphics, computer vision and medical imaging especially in extracting the skeleton of objects.

The fast marching method is one of most capable algorithms to compute the distance field at Euclidean metric if the front evolves at a unit speed. By choosing $F = 1$ in fast marching algorithm, the arrival time at each point shows the Euclidean distance. This distance field can be the minimum distance of each point to object boundary (DFB) or a minimum distance to a known source point (DFS). In Figure3-3 the computed distance field is visualized for different 2D shapes using the high accuracy fast marching method. The first column shows the distance field from the known source point (DFS) while the second column shows the object boundary (DFB).

(a)   (b)
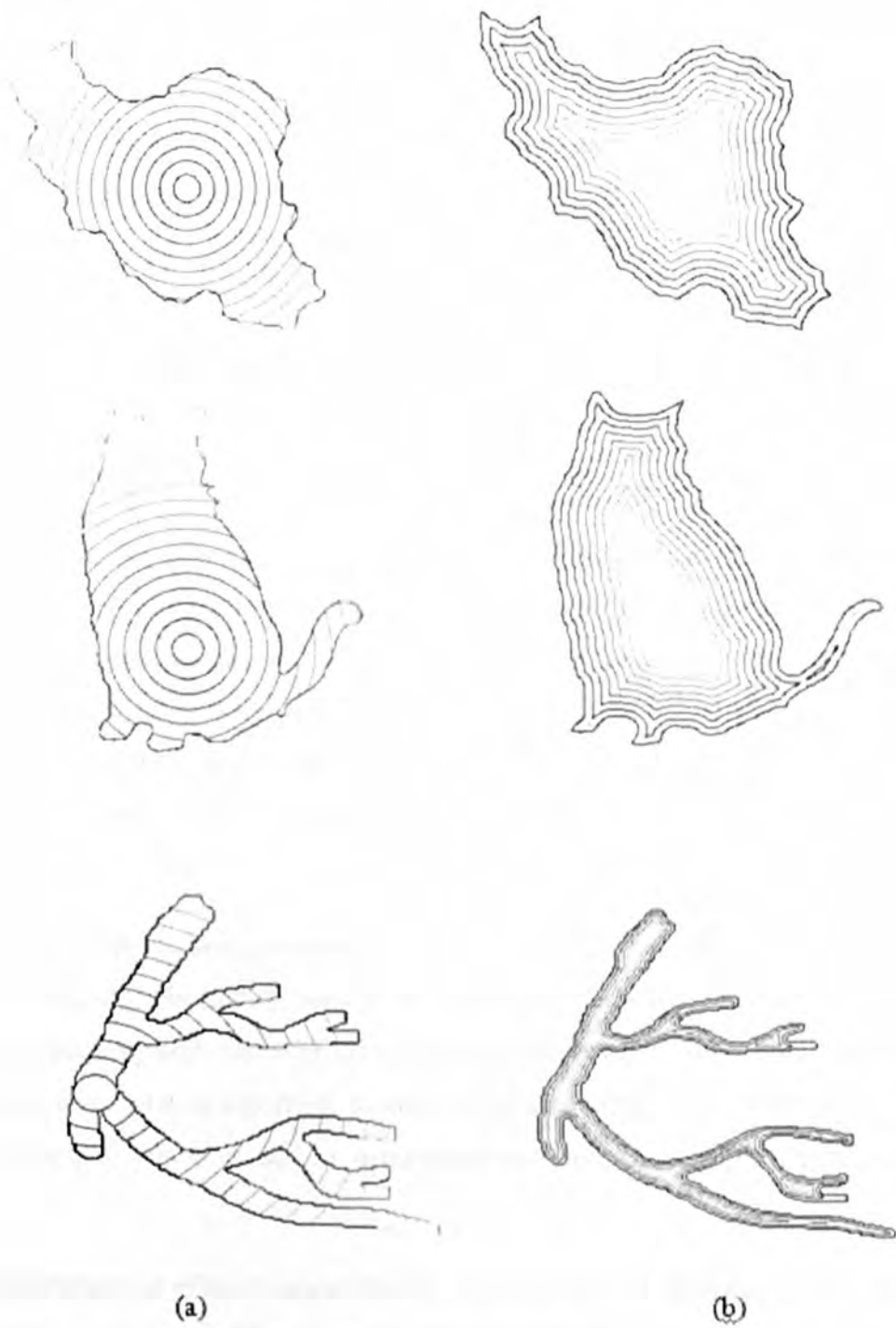
Figure 3-3 shows the computed distance field for different 2D shapes using high accuracy fast marching method. (a) Shows the distance field from the known source point. (b) Shows the distance field from the object boundary.

3.3. Pulmonary artery (PA) segmentation algorithm

Pulmonary artery tree segmentation is an essential requirement for some medical applications, such as detection and visualization of pulmonary embolisms (PEs). Moreover, as discussed in the previous chapter, segmentation of the pulmonary artery tree is complicated because it has multiple connections with other anatomical structures such as the pulmonary veins which can interfere with the segmentation of pulmonary artery boundaries. Many methods have been developed to perform the segmentation of an elongated shape such a pulmonary artery. In [40-42] a tubular enhancement filter has been employed using a combination of the eigenvalues of a Hessian matrix followed by thresholding. Zahlten used a wave propagation method which proceeds by recursively marching through 3D neighbours of a seed voxel using image intensity, also based on thresholding. This algorithm creates a bifurcation graph by looking at the connectivity of the wavefront [43]. The method reported by Sebbe et al [44] uses slice marching and fast marching algorithms, which is sensitive to irregular vessel boundaries but is less sensitive than wave-front based methods. In general, most previous attempts to segment the pulmonary artery can be categorized into region growing methods [123-125], enhancing the vessels with Hessian matrix [126-128] and fuzzy connectivity methods [129].

In all previous methods, seed points for segmentation algorithms were manually selected. Therefore, automatically finding seed points allows the algorithm to be included in fully automatic pulmonary embolism detection systems. Moreover, in most previous proposed methods, the segmentation algorithm sometimes fails by merging the pulmonary artery to an adjacent organ: this is caused by acquisition resolution or noise and partial volume effect.

In this contribution, an efficient algorithm for segmenting the Pulmonary Artery (PA) tree in 3D pulmonary Computed Tomography Angiography (CTA) images is presented. The pulmonary artery tree is extracted in two stages. We first extract the major artery in the heart region using a region growing method initialized by a seed point which is automatically selected within the pulmonary artery trunk. The segmentation result of the major artery is refined using a customized 3D level set algorithm. The major artery is extracted after the following steps:

1) Enhancing the image by a non-linear diffusion filter.

2) Extracting lung and heart region as a search area for pulmonary artery segmentation.

3) Removing the connectivity between the artery and other organs using a feature map based on eigenvalues of the Hessian matrix.

4) Removing the superior vena cava

5) Segmenting the pulmonary artery using a customized level set method starting from the initialized contour inside the pulmonary trunk.

In the second stage, the peripheral pulmonary arterial network is extracted from the lung regions using a region growing algorithm, starting from seed points which are end points of left and right artery and are automatically selected. The following sections provide a detailed description of all steps for extracting the pulmonary artery.

3.3.1. Diffusion Filter

In medical image segmentation, especially in analysing a feature of an image, two requirements are necessary, smoothing of homogeneous regions and preserving the edges or location of boundaries. To achieving these two, a few works have been developed based on convolution of an image by a Gaussian filter. Although the Gaussian filter is a simple and effective tool to remove image noise, it also smoothes the boundaries of objects. P. Perona and J. Malik [45] proposed a non-linear diffusion model called anisotropic diffusion model to smooth an image whilst preserving the object boundaries. Anisotropic diffusion is a process in image processing to decrease image noise without removing parts of the image which are important for interpretation of images such as edges and lines. In fact in image processing, the anisotropic diffusion filter is a powerful tool to enhance and restore an image which is based on a PDE of heat transfer:

$$\frac{\partial I}{\partial t} = div(c(\|\nabla I\|)^2 \nabla I) \tag{23}$$

where, $c(\|\nabla I\|)^2$ is diffusion coefficient which is usually based on the image gradient to control the diffusion rate in order to preserve the edge of image and it can be defined as:

$$c(\|\nabla I\|)^2 = \frac{1}{1+\frac{\|\nabla I\|^2}{k^2}} \quad \text{or} \quad c(\|\nabla I\|)^2 = e^{-\left(\frac{\|\nabla I\|}{k}\right)^2} \tag{24}$$

where k is a gradient modulus threshold that controls the conduction. In other words, k is a manually-selected contrast constant that controls the spatial extent of the smoothing.

Figure 3-4 shows the graphs of two diffusion coefficients with different values for k.
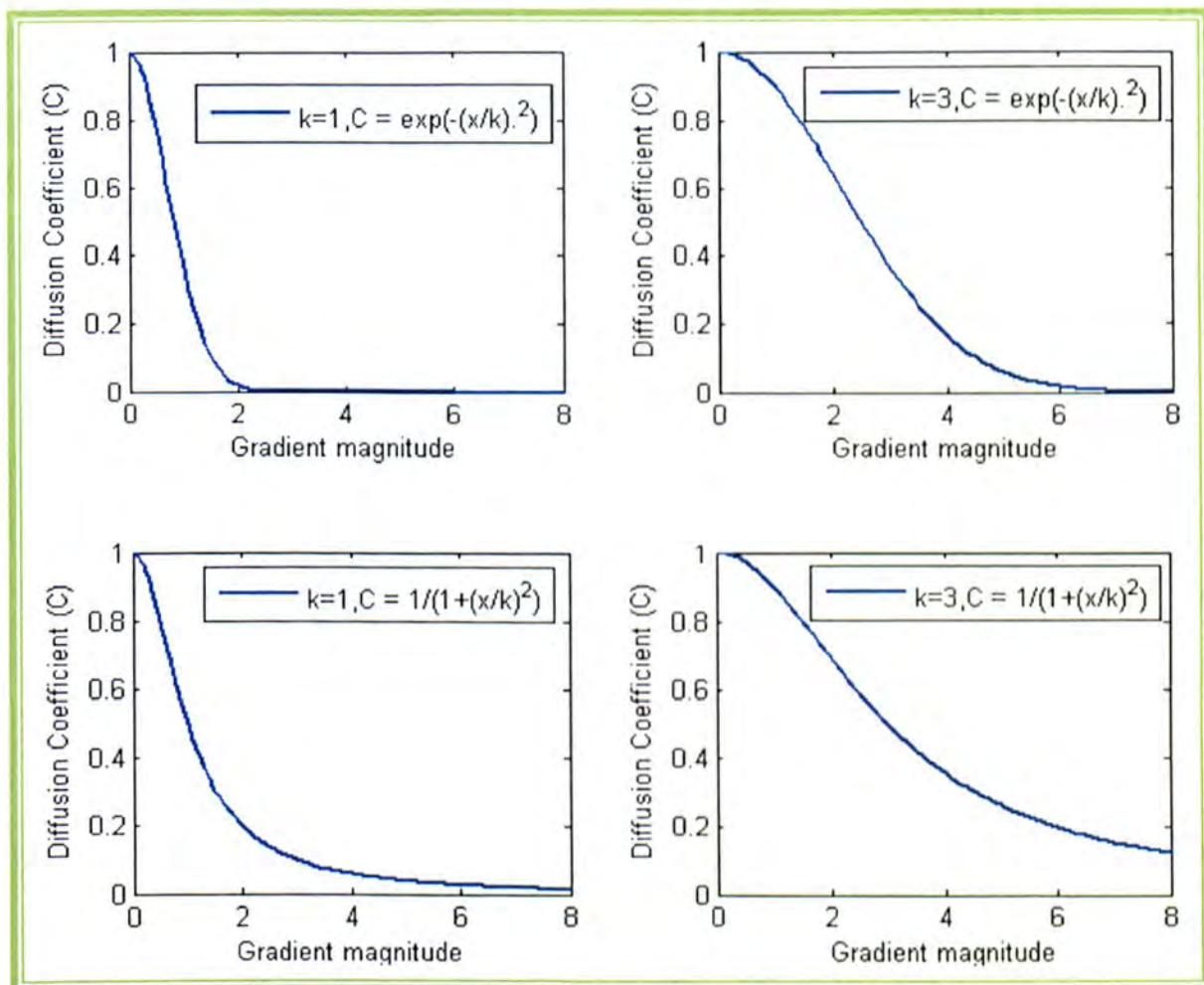
Figure 3-4 shows the graph of two different diffusion coefficient with different value for **k**.

In the proposed PA segmentation algorithm, the raw data is first smoothed using the non-linear diffusion filter proposed by P. Perona and J. Malik [45] by selecting the diffusion coefficient $c(\|\nabla I\|)^2$ as:

$$c(\|\nabla I\|)^2 = e^{-\left(\frac{|\nabla I|}{k}\right)^2}$$

3.3.2. Lung segmentation

Lung segmentation is a first step for quantification of most pulmonary diseases in pulmonary CT images such as analyzing lung parenchyma density [46], analyzing the mechanism of diaphragm [47] and airway analysis [48]. Accurate lung segmentation from high resolution CT images is a complicated task because of various detailed tracheal structures and complex lung anatomy. In pulmonary artery segmentation, lung segmentation is a crucial pre-processing step to reduce the search area. Several works have been developed to segment the lung region in CT images. In the proposed CAD system, an efficient algorithm is used to segment the lung region. This algorithm is a fully automatic technique to segment the lungs in CT images which is based on three main steps. In the first step the lungs region is extracted using an optimized thresholding method and then left and right lungs are separated by identifying the anterior and posterior junctions. The left and right lungs are the search area for segmenting the peripheral artery. In the final step, the boundary of the lungs region will be smoothed. Figure3-5 shows the diagram of the lung segmentation algorithm.
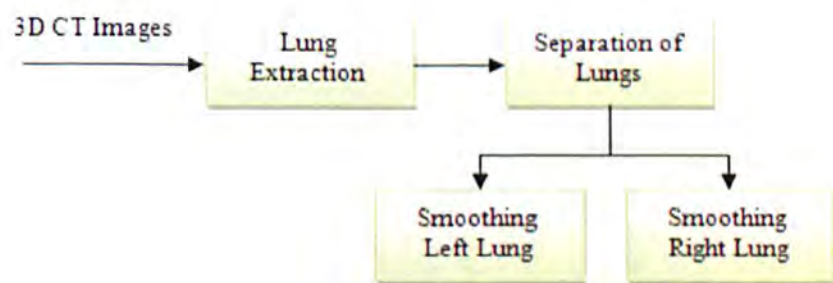
Figure 3-5 shows the diagram of lung segmentation algorithm

3.3.2.1. Lung extraction

The main aim of lung extraction is to separate the voxels of lung tissue from the voxels corresponding to the surrounding area using gray level thresholing. Instead of using a fixed gray level threshold, an optimized thresholding which allows the system to

accommodate the small variation in tissue density and uses an iterative method. In this method the lung region (body region) is separated from the background (non-body region) using a threshold T^k which is calculated iteratively [49]:

$$T^{k+1} = \frac{L_b^k + L_n^k}{2} \tag{25}$$

where L_b^k and L_n^k are the mean gray value of the body and non-body voxels after separating the body from non-body region with threshold T^k, respectively. Using an initial threshold T^0, the procedure (25) updates the thresholds until the condition $T^k = T^{k+1}$ holds. The initial threshold T^0 is chosen based on the intensity value for pure air (-1000 HU). The optimized threshold allows the system to adapt any small changes in tissue intensity.

After thresholding the data using optimized threshold, a 3D connected component labelling is used to extract the big component which is related to lung region (Figure3-6). The extracted big component is contains of the left and right lungs, trachea and airways. In order to perform quantitative analysis of lung tissue and facilitate left and right lung separation, the trachea must be separated from the extracted big component. Before separating the trachea, the data is rearranged to keep a big component for optimizing memory of system by creating a convex box containing the big component.

Figure 3-6 shows the result of lung extraction using 3D connected component labelling.

The trachea is segmented using an up to down region growing algorithm, starting from a seed point which is automatically selected from the top slices by searching for a circular connected component close to the middle of the slices. The result of trachea segmentation is subtracted from the result of lung extraction to remove the trachea.

To find seed points for an up-to-down region growing algorithm, 2D connected component labelling is used from the top slices after thresholding based on CT number for pure air (-1000 HU). Some properties of extracted components (objects) are measured such as: 1) Eccentricity which is a parameter which measures deviation of a conic section from being circular. The value is between 0 and 1. The ellipse whose eccentricity is 0 is actually a circle, while an ellipse whose eccentricity is 1 is a line segment. 2) Area which is the actual number of pixels in the region. 3) Solidity which is the proportion of the pixels in the convex hull. The closer 2D connected component to middle of top slice whose eccentricity is less than 0.60, the area is more than 150 and its solidity is more than 0.40 is selected as the seed region for the up-to-down region growing algorithm. The criterion of the growing in the proposed algorithm is the similarity of voxel values in which voxels with CT number between -900 HU and -1100 HU are collected as the trachea region. The trachea is then separated from the extracted big component to achieve lung region. Figure 3-7 and Figure 3-8 show the results of trachea and lung extraction.

(a)

(b)

(c)

(d)
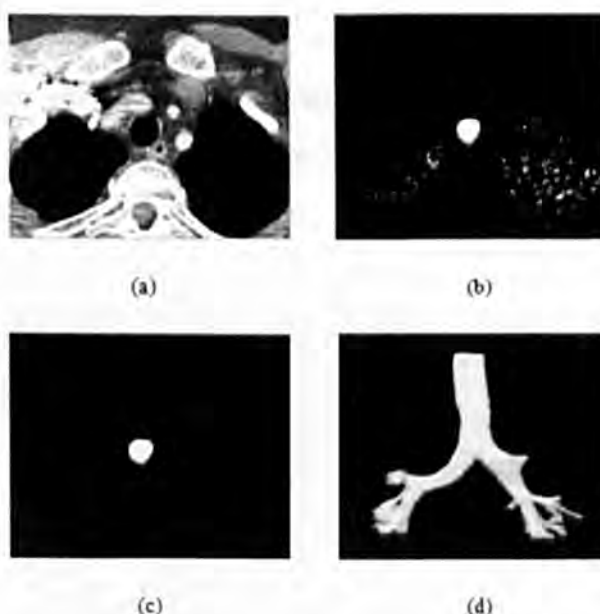
Figure3-7 shows the results of trachea extraction algorithm. (a) Original top slice which is automatically selected to find seed points for up-to-down region growing algorithm. (b) Binary mask after gray level

thresholding with CT number between -900 HU and -1100 HU. (c) Seed points detected for trachea extraction. (d) 3D rendering for 3D trachea extraction results.

Figure 3-8 shows the result of lung extraction. The trachea and main bronchus were removed from the lung region

3.3.2.2. Left and Right Lung Separation

In cross section CT images, the posterior and the anterior junctions between left and right lungs are sometimes thin and low contrast so in the lung extraction process applying gray level thresholding may not separate the left lung from the right lung near to these junctions. To separate the left and right lung, first those slices that contain a single connected lung component are identified and then the location of the junction's lines is recognized. Unlike the proposed method in [50] which uses a 3D morphological operator to separate the left and right lungs which is a time consuming method, the proposed method searches for junction positions in 2D slices. The separation process is only applied on those slices which have a big and single connected lung component. The location of the junction lines changes through the dataset due to the smooth pulmonary anatomy. To find the position of the junction line, first the number of pixels is counted at each column and the minimum number corresponds to the position of a junction line (Figure 3-9).
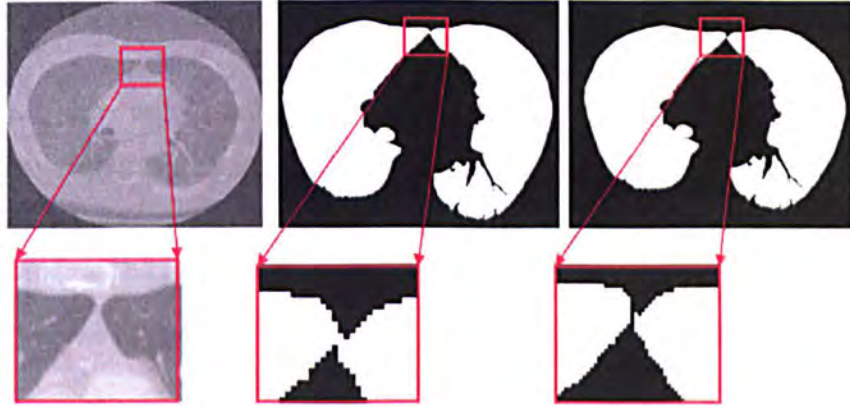
Figure 3-9 shows the original image, connectivity between left and right lungs after gray level thresholding and separated left and right lungs.

3.3.2.3. Smoothing

In CT pulmonary images, the CT number used for extracting the lung region is between -1100 HU and -900 HU and CT number for vessel in lung region is between 100 HU and 300 HU. So in the lung extraction process, some vessels may be excluded from the lung region near the mediastinum and make empty spaces in the lung region. To fill the empty spaces caused by blood vessels and smooth the lung boundaries, a mathematical morphology operator is used. This mathematical morphology operator is a dilation operator followed by an erosion operator. The erosion operator is chosen to be slightly larger than the dilation operator to prevent connectivity between right and left lungs. The structure element applied for the morphology operator is a disk with radius of 12mm for the erosion operator and 10 mm for the dilation operator. Figure 3.10 shows the result of a smoothing step in the lung segmentation algorithm.

Figure 3-10: shows the results of the smoothing step in lung segmentation algorithm. (a) Original image. (b) Binary image after applying the optimized gray level threshold which shows voids caused by blood vessels. (c) The smoothed binary image

3.3.3. Heart Segmentation

Since the pulmonary trunk, left artery and right artery are in the region between the left and right lungs called the mediastinum (the region which contains the heart), so the heart region should be segmented as a search area for the pulmonary artery segmentation algorithm. In the proposed CAD system heart segmentation is a crucial task to find seed points for pulmonary artery segmentation algorithm automatically.

An efficient algorithm is used to extract the region between the left and right lungs as the heart region automatically. In this algorithm, the binary lung region is the input data and the heart region will be output of the system and is based on three main steps: 1) Extracting four sharp points of left and right lungs at each slice. 2) Separating the edge between up and bottom sharp points on the left and right lung in each slice. 3) Connecting the two top and two bottom sharp points. Figure3-11 shows the design of the heart segmentation algorithm.
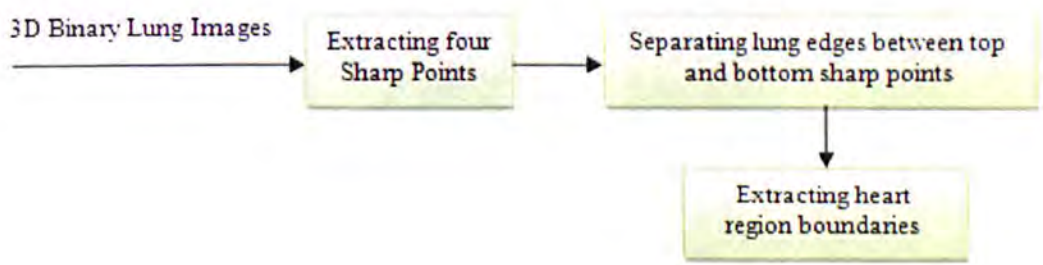
Figure 3-11 shows the diagram of the heart segmentation algorithm

Heart segmentation algorithm

Heart segmentation is based on three main steps. In the first step, the four sharp points are identified at each 2D binary lung slice. To extract four sharp points: First, the two large components are selected at each slice which is related to left and right lungs using a 2D connected component labelling algorithm. Second, by tracing object boundaries the exterior boundaries of the left and right lung are selected. Third, the key points of object boundaries are identified. The key points are identified by computing and comparing the distance between the exterior boundary points and the line which connects the top and

Figure 3-13 shows the results of lung and heart segmentation.

(a)

(b)

(c)

(d)   (e)
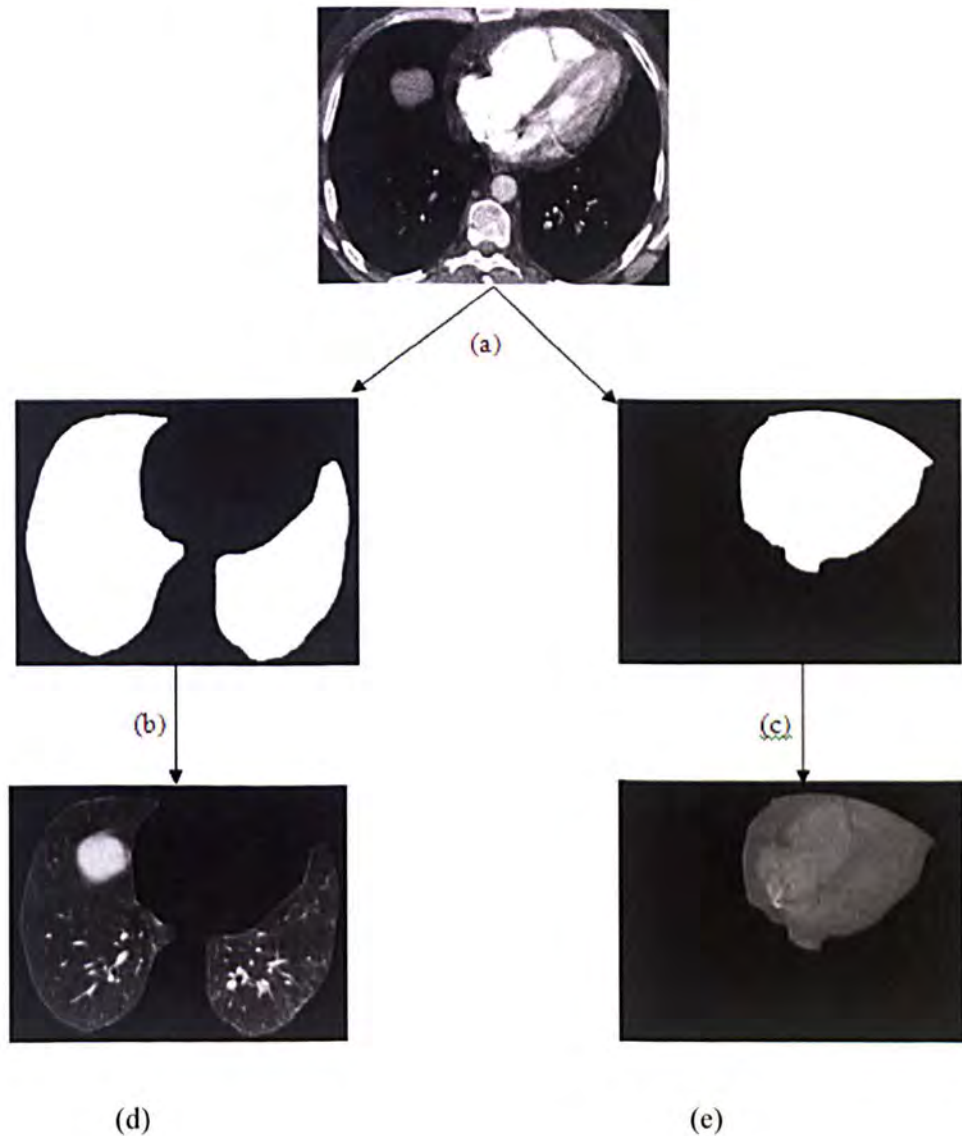
Figure 3-13: shows the lung and heart segmentation results. (a) Original image. (b) Mask of lung segmentation result. (c) Mask of heart segmentation result. (d) The extracted lung region from original image. (e) The extracted heart region from original image.

3.3.4. Feature map

The main interference in segmentation of the pulmonary artery in the heart region comes from those structures which also have high intensity and are close to the pulmonary artery such as the pulmonary vein, the superior vena cava and the aorta. Since the boundaries between different high-valued organs always have a relatively lower intensity value but low contrast, so in all segmentation algorithms based on their CT number, these two high-

valued organs are connected. Figure3-14 shows an original image with connectivity between the pulmonary artery trunk, superior vena cava and aorta.

(a)   (b)
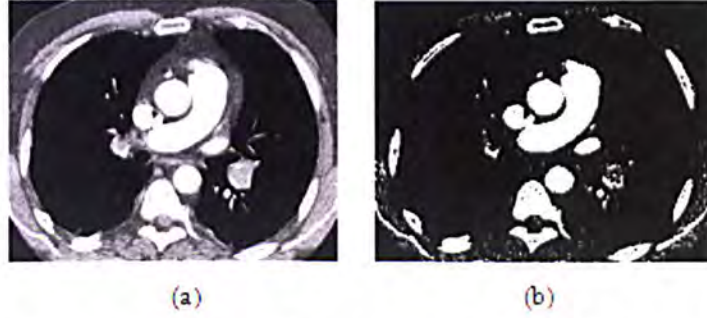
Figure 3-14: shows connectivity between the pulmonary artery and the adjacent organs. (a) Original image. (b) The binary image after adaptive thresholding that shows connectivity between pulmonary trunk and aorta and superior vena cava.

A feature map is constructed to enhance the boundaries between two different organs. The Hessian matrix and its eigenvalues are employed in the proposed pulmonary artery segmentation algorithm as a pre-processing step to construct a feature map. The feature map utilizes the 2nd-order derivatives of the image intensity after smoothing (using a 3D Gaussian kernel) at multiple scales to identify boundaries between two bright structures.

Hessian matrix

To analyze the local behaviour of a 2D/3D image I, the Taylor expansion of image I in the neighbourhood of a point x_0 is usually used.

$$I(x_0 + \Delta x) = I(x_0) + \Delta x^T \nabla I(x_0) + \Delta x^T H(x_0) \Delta x \tag{26}$$

This expansion approximates the structure of an image up to second order where, $\nabla I(x_0)$ is the gradient vector of image I at point x_0 and H is the Hessian matrix:

$$H(x_0) = (\nabla . \nabla^T) I(x_0) = \begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{bmatrix} . \begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \end{bmatrix} I(x_0) = \begin{bmatrix} \frac{\partial^2 I(x_0)}{\partial x^2} & \frac{\partial^2 I(x_0)}{\partial x \partial y} & \frac{\partial^2 I(x_0)}{\partial x \partial z} \\ \frac{\partial^2 I(x_0)}{\partial y \partial x} & \frac{\partial^2 I(x_0)}{\partial y^2} & \frac{\partial^2 I(x_0)}{\partial y \partial z} \\ \frac{\partial^2 I(x_0)}{\partial z \partial x} & \frac{\partial^2 I(x_0)}{\partial z \partial y} & \frac{\partial^2 I(x_0)}{\partial z^2} \end{bmatrix} \tag{27}$$

The Hessian matrix, which describes second order gray level intensity variation, has been widely applied to enhance the image and detection of specific structures such as

vasculature in medical image analysis [56]. Eigen-value and eigenvector analysis of the Hessian matrix is used to characterize the local structure of each point in 2D and 3D images.

To remove the connectivity between the pulmonary artery and other tubular structures, such as the pulmonary vein and the aorta, a feature map is constructed by a boundary enhancement filter:

$$w(p) = \text{Max }_{\sigma_{min} \leq \sigma \leq \sigma_{max}, 1 \leq k \leq 3} \; (\sigma^r * \lambda_{\sigma,k}(p)) \tag{28}$$

where, $\lambda_{\sigma,k}(p)$ denotes the k-th eigenvalues of the Hessian matrix $\mathcal{H}_\sigma(p)$ of image $I_\sigma(p)$ at scale σ. Parameter r introduced in [51] is the r-normalized derivative for a normalized second derivative. In the proposed CAD system, a feature map (mask) is constructed based on a predefined threshold value T_w applied on $w(p)$ to remove the connectivity between the pulmonary artery and other organs. Figure3-15 shows results of applying the feature map filter to CTA images using $\sigma = 2.5$, 3.0 and 3.5 (σ is selected based on the size of pulmonary artery), $T_w = 50$. The parameter r is chosen experimentally as $r = 1.5$.
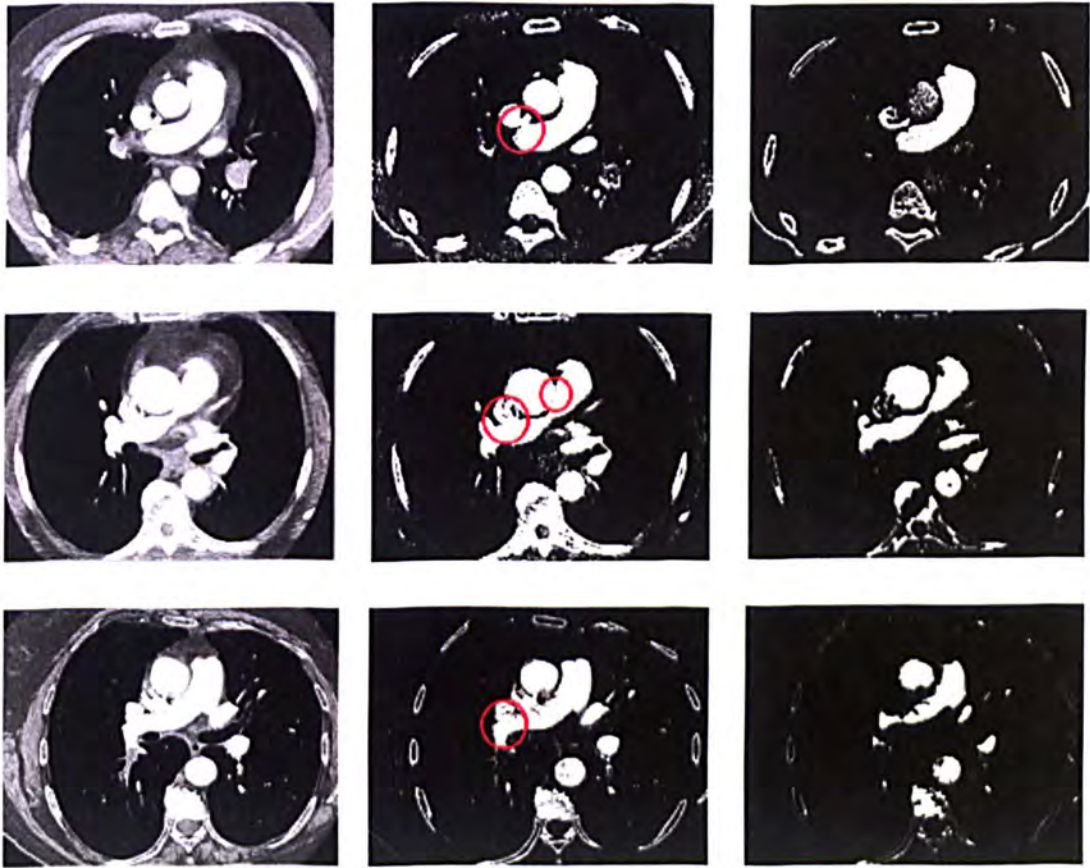
Figure 3-15: Shows removal of the connectivity between the pulmonary artery and the surrounding organs. The first column shows the three original CTA images; the second column shows the results of thresholding; the third column shows the effect of using the feature map to remove the connectivity between the pulmonary artery and the adjacent organs.

3.3.5. Removing Superior Vena Cava (SVC)

Anatomically, the superior vena cava is located in the middle of the chest and is surrounded by the trachea, aorta, thymus, pulmonary artery and right lung bronchus. It empties into the right atrium and then the right ventricle (Figure 3-16). It returns the blood to the heart region from the neck, both upper limbs and the head.
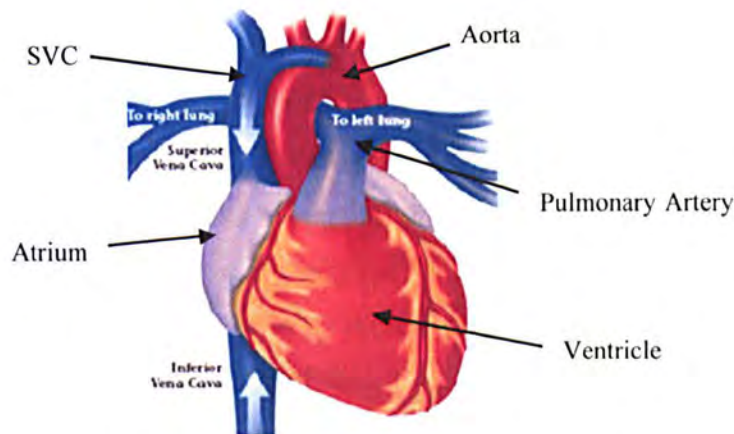
Figure 3-16 shows the middle region of the chest which consists of trachea, aorta, thymus, pulmonary artery, atrium and ventricle

In a CTA image the CT value of the superior vena cava is the same as or slightly more than the CT value for the pulmonary artery and aorta and it may have connectivity with pulmonary artery and the aorta (Figure3-17).
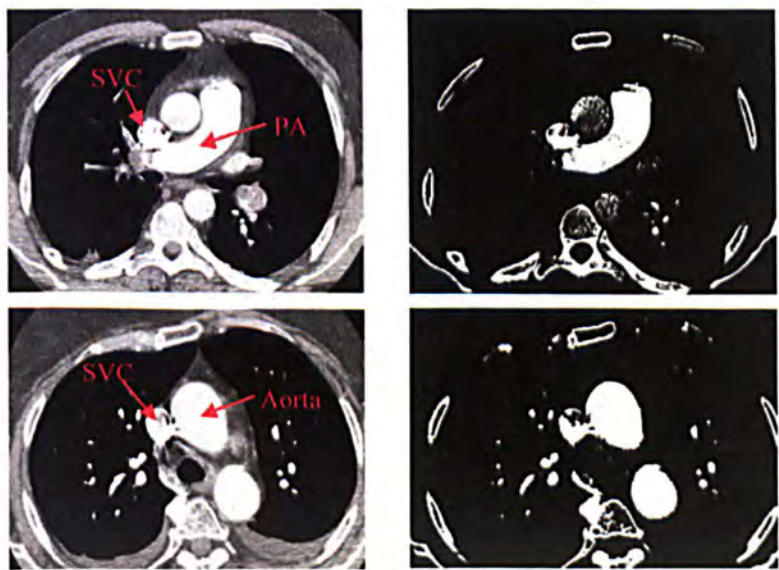
Figure 3-17 shows the connectivity between the superior vena cava and the other organs. First row shows the connectivity with the pulmonary artery and the second row shows the connectivity with the ascending aorta.

The superior vena cava interferes with the segmentation and separation of the pulmonary artery. In order to facilitate the segmentation of the pulmonary artery, the superior vena cava is extracted using a region growing algorithm starting from a seed point which is automatically selected in top slices in the heart region. Since the contrast material is injected through the superior vena cava to the pulmonary vessels in CTA images, the CT value of the superior vena cava is slightly larger than the CT value of other vessels especially in the top of aortic arc region. Therefore the seed points can be identified after applying a threshold determined by the CT number for superior vena cava (400 HU), followed by 2D connected component labelling. These seed points are related to the largest component. Figure 3-18 shows the results of segmentation of the superior vena cava. The left column shows the original 2D top slices of the heart region in CTA images, the middle column shows the selected seed points and the right column presents the superior vena cava segmentation results using a top-down region growing algorithm.

Figure 3-18: shows the superior vena cava segmentation results. Left column: Original 2D top slices of heart region for looking for seed region. Middle column: The seed region which was selected automatically in top slices in the heart region. Right column: Superior vena cava segmentation results.

3.3.6. Major Artery Segmentation

The pulmonary artery segmentation algorithm is performed in two steps: major and peripheral artery segmentation. The major artery is extracted by a new customized level set algorithm in the heart region by using an initial surface which is automatically selected inside the pulmonary artery region as an initial counter. The proposed level set algorithm is an edge-based model which relies on an edge function depending on the image gradient to stop curve evolution. In noisy images, using a strong isotropic smoothing Gaussian filter smoothes the edges and makes the curve pass the boundary of the object. To smooth the homogeneous region and preserve the edges or locations of boundaries, an anisotropic diffusion filter proposed in [45] is applied for pre-processing. Figure 3-19 shows the diagram of a major artery segmentation algorithm.
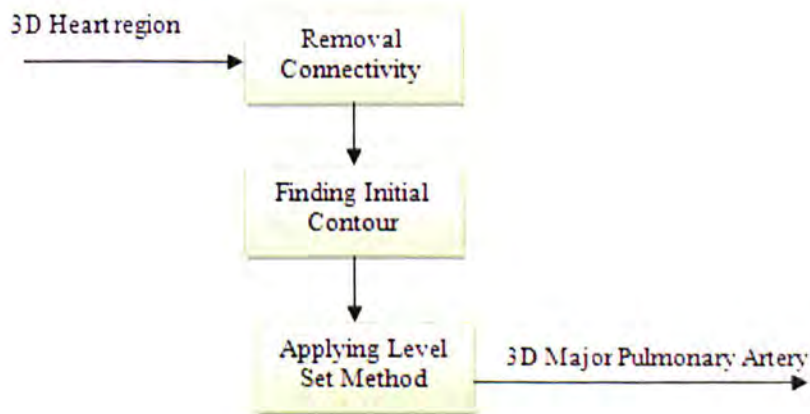
Figure 3-19 shows the diagram of the major artery segmentation algorithm.

3.3.6.1. Initial Contour

The initial contour for the 3D level set algorithm is selected within the pulmonary artery. First, the initial segmentation of the artery is taken by thresholding with an optimized threshold. This optimized threshold is automatically selected by analyzing an intensity histogram of the heart region. Based on anatomical analysis, there are two main groups in the heart region: 1) The pulmonary artery, superior vena cava, right atrium and ventricle which have the highest CT value, and 2) The left atrium and ventricle, aorta and pulmonary vein which have a lower average CT value than the pulmonary artery. Figure 3-20 (first row) shows the distribution of CT values in the heart region for two different datasets, from which the first peak can be associated with the CT value of group (2) and

the next peak can be associated to group (1). In order to find an optimized threshold, the distribution CT value profile of the heart region in an image is analyzed. The two biggest maximum peaks of profile are identified and the minimum peak between these two is selected as an optimized threshold. Hence a suitable threshold for segmentation of the pulmonary artery T_{pa} for the first dataset was automatically set as a value $T_{pa} = 1201$ (Figure3-20 left profile of second row) and for the second dataset the optimized threshold was computed as a value $T_{pa} = 1135$ (Figure.3-20 right profile of second row).
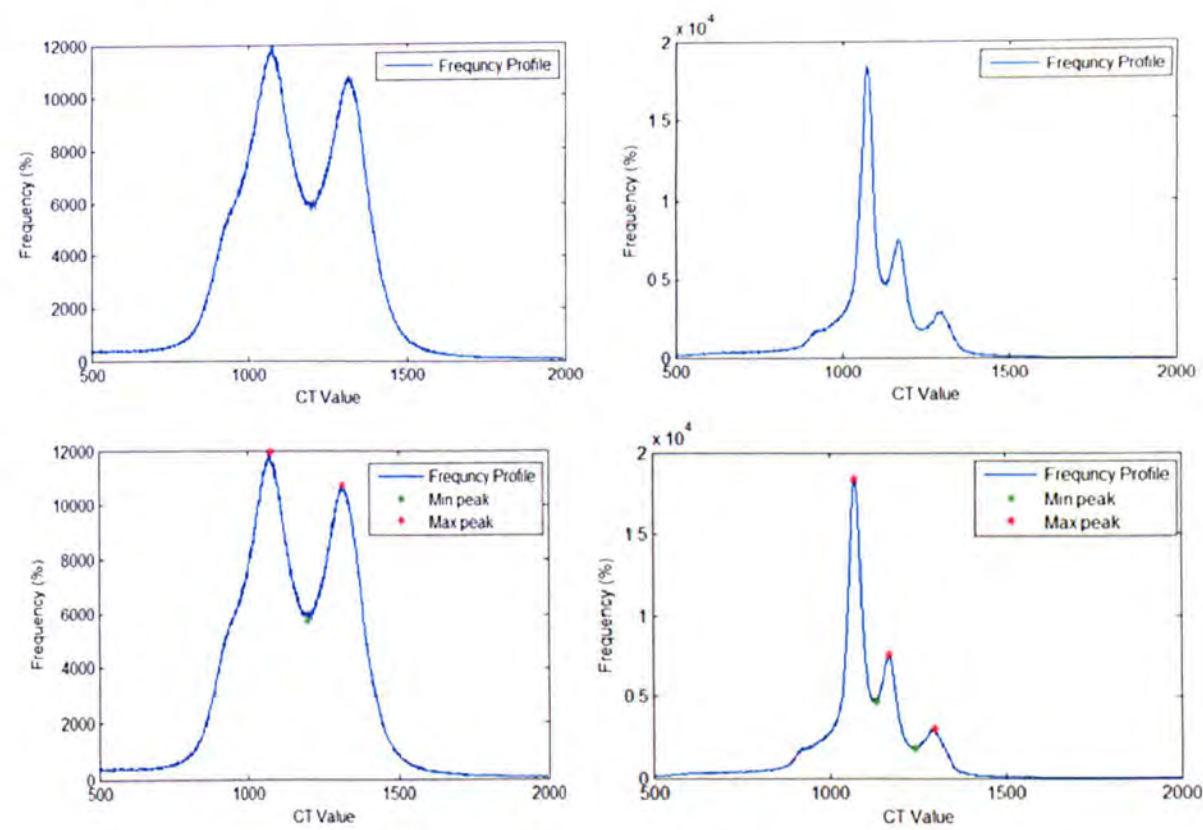
Figure 3-20 First row show the distribution of CT values profile in the heart region for two different datasets and second row shows the computed maximum and minimum peaks. The minimum peak between the two biggest maximum peaks of the profile is considered as an optimized threshold.

After initial pulmonary artery segmentation using an optimized threshold, the feature map filter is applied to remove the connectivity between the artery and other organs. In some datasets the feature map filter may not completely remove the connectivity, so to ensure that there are no more connections between the pulmonary artery and other organs, a feature map filter following a morphological erosion operator is applied. Figure 3-21 shows the results of removing connectivity between the pulmonary artery and other

organs (superior vena cava and aorta) using the feature map filter followed by a morphological erosion operator.

(a)   (b)   (c)

Figure 3-21 shows the removal of connectivity between the right artery and superior vena cava using the proposed feature map and erosion operator. (a) Shows the connectivity between the superior vena cava and the right artery (arrowed). (b) Feature map implementation result. (c) Shows the completed removal of connectivity after using the erosion operator.

After removing the connectivity, a region growing algorithm is applied starting from a seed point which is automatically selected inside the artery by searching for the largest connected component in the heart region. The biggest connected component is selected using a 2D connected component labelling algorithm. Figure 3-22 shows automatic seed points selection results on a different dataset. The left column shows the original CTA images and the middle column presents the thresholding results on the original data using an adaptive threshold to extract an initial contour and the right column shows the selected seed points.

Figure3-22: shows selecting seed point inside the pulmonary artery. Left: Original row data. Middle: thresholding results using an adaptive threshold. Right: The selected seed points inside the pulmonary artery in the heart region.

The result of the region growing algorithm is considered as an initial contour for the level set algorithm to obtain a desirable segmentation of the pulmonary artery. Figure 3-23 shows the different 3D initial contours which are automatically selected inside the pulmonary artery in the heart region.

Figure 3-23 shows different 3D initial contours selected inside pulmonary artery.

3.3.6.2. Proposed 3D Level Set Method

As Figure3-15 shows, the vessel size can be reduced after applying the feature map filter on CT scans and it can be the reason that a CAD system misses some potential PEs in those trunk and main arteries with reduced size. In order to enhance the segmentation result a new customized level set algorithm is applied using the initial contour described in section3.3.6.1. The new customized level set algorithm is inspired from a geometric active contour model proposed by Caselles in [52] to improve the segmentation results derived from region growing algorithms. This model which overcomes the limitation of classic snakes is obtained by minimization of the functional:

$$E(C) = \alpha \int_0^1 |C'(q)|^2 dq + \lambda \int_0^1 g\big(|\nabla I(C(q))|\big)^2 dq \qquad (29)$$

$$= \int_0^1 \Big(E_{int}(C(q)) + E_{ext}(C(q)) \Big) dq$$

where $C(q): [0\ \ 1] \to \mathbb{R}^2$ is a parameterized curve, $I: \mathbb{R}^2 \to \mathbb{R}$ is an image in which the object boundary will be detected and α and λ are real constants. By minimizing the functional $E(C)$ using Euler-Lagrange the following curve evolution equation is derived.

$$\frac{\partial C(t)}{\partial t} = g(I)\kappa \overrightarrow{\mathcal{N}} - (\nabla g . \overrightarrow{\mathcal{N}})\overrightarrow{\mathcal{N}} \qquad (30)$$

where κ is the Euclidean curvature and $\overrightarrow{\mathcal{N}}$ is the unit inward normal. Assuming that the curve C is a level set of a function u (C coincides with the set of points $= constant$), then u is an implicit representation of curve C and is topology free so that different topologies of the zero level set do not imply different topologies of u. So, this implicit representation of curve C can be appropriate for curve evolution applications in the

framework of the level set method introduced by Osher and Sethian [39]. The level set framework of the model by adding a constant v which controls the speed of evolution and adding the term $\nabla g \cdot \nabla u$ which attracts the curve to the boundary of objects is:

$$\frac{\partial u}{\partial t} = |\nabla u| div\left(g(I) \frac{\nabla u}{|\nabla u|}\right) + vg(I)|\nabla u| \tag{31}$$

This equation is equivalent to:

$$\frac{\partial u}{\partial t} = g(v + \kappa)|\nabla u| + \nabla u \cdot \nabla g \tag{32}$$

$$\kappa = div(\frac{\nabla u}{|\nabla u|})$$

That means that the level sets move according to:

$$C_t = g(I)(v + \kappa)\vec{\mathcal{N}} - (\nabla g \cdot \vec{\mathcal{N}})\vec{\mathcal{N}} \tag{33}$$

The advantage of this model is that the model uses fewer parameters. The parameter v is the only parameter which is used in the model and it is possible to choose $v = 0$ (no constant velocity), and the model still converges but in a slower motion. Since this model evolves an initial curve according to the boundary attraction term towards one direction (inwards or outwards), therefore, in order to achieve a proper result it needs a specific initialization step. The initial curve should be completely interior or exterior to the real object boundary. As described in the initial contour section, the initial contour is automatically selected inside the real object boundary (inside the pulmonary artery), so, inspired by this model and using a new customized external energy, the following edge-based level set algorithm is used to segment the pulmonary artery:

$$\Phi_t = |\nabla\Phi|\left(div\left(\text{Fext} \frac{\nabla\Phi}{|\nabla\Phi|}\right) + v\text{Fext}\right) \quad \text{in } [0, \infty[\times R^3$$

$$\Phi(x,y,z,0) = \Phi_0(x,y,z) \quad \text{in } R^3 \tag{34}$$

where Φ_0 is the initial level set function (in our application the result of initial major artery segmentation described in section 3.3.6.1 is considered as the initial level set function), $v \geq 0$ is constant and F_{ext} is an external energy function.

We proposed a new external energy which is composed of two energy functions:

$$F_{ext} = F_{edge} \times F_{intensity} \qquad (35)$$

F_{edge} is based on edge strength which can be a positive and decreasing function that depends on the gradient of the image u_0 .

$$F_{edge}(x, y, z) = \frac{1}{1 + |\nabla G(x,y,z) * u_0(x,y,z)|^2} \qquad (36)$$

or

$$F_{edge}(x, y, z) = e^{-|\nabla I(x,y,z)|}$$

where $G * u_0$, a smoother version of u_0, is the convolution of the image u_0 with the Gaussian filter.

$F_{intensity}$ is based on intensity values which prevent front propagation to the background where the gradient of the boundary of object is weak.

$$F_{intensity} = \begin{cases} \exp(-1/2 \times (\frac{I(x,y,y)-c1}{s})^m) & I(x,y,z) \leq c1 \\ 1 & c1 < I(x,y,z) < c2 \\ \exp(-1/2 \times (\frac{I(x,y,y)-c2}{s})^m) & I(x,y,z) \geq c2 \end{cases} \qquad (37)$$

where c1 and c2 are defined as the minimum and maximum CT values of the pulmonary artery (in our application c1=150 HU and c2=350 HU) and s, m are tuning parameters of the function which are empirically set to 50 and 2 respectively.

Figure 3-24 shows the graph of external energy based on intensity function. This function is based on prior intensity information of arterial tree.
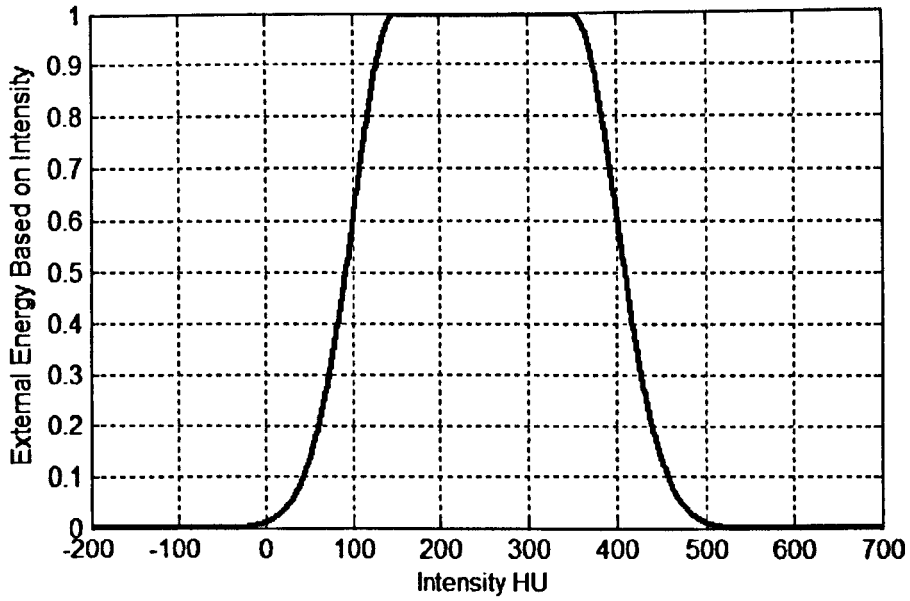
Figure 3-24 Function of intensity based on external energy (c1=150, c2=350, s=50 and m=2)

In the implementation of the level set algorithm, it is numerically crucial to keep periodically the evolving level set function close to a signed distant function as the level set function Φ can develop a very sharp and flat shape during the evolution [52],[53]. A re-initialization process has been used to keep the level set function as a signed distance function in the level set implementation algorithm. A common re-initialization method is to solve the following equation proposed by Sussman, Smereka, and Osher [54]

$$\frac{\partial \varphi}{\partial t} = S(\varphi_0)(1 - |\nabla \varphi|) \tag{38}$$

Where, φ_0 is a function that is to be re-initialized, and $S(\varphi_0)$ is the sign function given by [54].

$$S(\varphi_0) = \frac{\varphi_0}{\sqrt{\varphi_0^2 + \epsilon^2}} \tag{39}$$

Given any initial data for φ, and by solving the equation (38) to the steady state, the new values for φ have the property that $|\nabla \varphi| = 1$, since convergence occurs when the right-hand side of (38) is zero. The sign function controls the flow of information: for negative values, information flows one way, and for positive φ the information flows other way [55].

3.3.6.3. Level Set Stopping Criterion

Level set algorithms commonly choose an arbitrary stopping criterion for the evolution process based on a fixed (and usually large) number of iterations, M. This can lead to problems with stability and convergence of the evolving function. Firstly we may achieve the desired segmentation before the iteration process ends, hence performing much unnecessary computation and increasing the algorithm's execution time. These unnecessary iterations may also result in the evolving curve crossing the object boundary, especially at weak boundaries. Secondly, the speed function, which forces the embedded curve into a steady-state, is leaky, so it can produce unsatisfactory segmentation results after much iteration [131]. As a consequence we should wish to choose a smaller value for the stopping criterion. We propose a new stopping criterion which uses a measure of the mean curve energy, which forces the curve toward the object boundaries, to halt the evolution process at an optimum iteration

$$E = \frac{\oint_{\Omega} F d\delta}{\oint_{\Omega} d\delta} \tag{40}$$

$$\Omega = \{(x, y, z) | -\beta \le \Phi(x, y, z) \le \beta\}$$

where Ω is a narrow band of the evolving surface and β is the width of narrow band. We first apply the method with relatively large iteration value (M) which is determined by the individual application. During the process, the mean energy values of the N last iterations are kept as a profile. The process terminates where the variation of this profile falls to zero.

Figure 3-25a shows an example of the mean energy plotted against iteration number. Figure 3-25b is the contour associated with the best iteration number to segment the pulmonary artery trunk is $N_0 \approx 130$. Using a large iteration number (M $\approx$ 300) results in the evolving curve growing beyond the boundary of the artery (Figure 3-25-c).
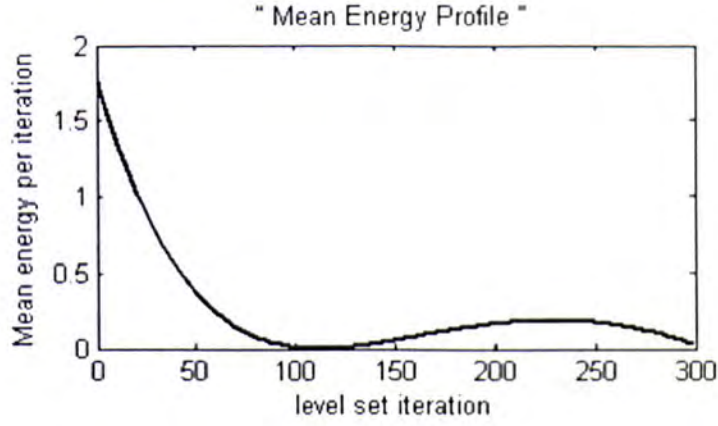
(a)

(b)   (c)

Figure 3-25 (a) Shows mean curve energy against iteration number for pulmonary trunk segmentation, (b) segmentation after 130 iterations (c) segmentation after **M**=300 iterations. The curve has crossed the boundary of the pulmonary trunk due to the large number of iterations.

3.3.6.4. Numerical Implementation of the Proposed Level Set Method

In order to reduce the running time of the process, a narrow band level set method has been used. The idea of narrow band level sets was introduced by Adalsteinsson and Sethien to solve the level set Eq.34 for a narrow band β pixel/voxel wide around the interface [68]. The values of Φ within the narrow band are updated during the computation.

In the level set Eq.34 all spatial derivatives $\frac{\partial \Phi}{\partial x}$, $\frac{\partial \Phi}{\partial y}$ and $\frac{\partial \Phi}{\partial z}$ are approximated by the central difference, and the temporal partial derivative $\frac{\partial \Phi}{\partial t}$ is approximated by the forward

difference. The approximation of Eq.34 by a finite difference scheme can be simply written as:

$$\Phi_{i,j,k}^{m+1} = \Phi_{i,j,k}^{m} + \delta t H(\Phi_{i,j,k}^{m}) \tag{41}$$

Where, $H(\Phi_{i,j,k}^{m})$ is the approximation of the right hand side of Eq.34 by the central spatial difference scheme and δt is the time step. More details of the numerical implementation of the proposed level set algorithm are given in Appendix B.

Stability and CFL condition

The value of the time step has been arbitrarily assigned in most level set implementations. The larger the time step, the higher the speed of the curve or surface evolution. For small values of the time step, the convergence is slow but the result is satisfactory, while large values of the time step, give a speed-up of the evolution but lead to poor results. Therefore, there is a stability issue which depends on the selected time step.

By solving the proposed partial differential equation (level set equation) numerically using a finite difference scheme, some approximation errors are presented at each iteration. These approximation errors keep accumulating over time and lead to instability of the solution. This instability is a local issue and even for an accurate numerical solution, this instability may occur. The global stability issue states that the numerical range of dependence must always contain the theoretical range of dependence of solution. Otherwise, no matter how accurate the solution, the error will always be magnified at each step and will lead to instability.

The application of the above principle to level sets leads to the Courant-Friedrichs-Lewy (CFL) condition [55] as:

$$F_{max} \cdot \delta t \leq \min(h_x, h_y, h_z) \tag{42}$$

Where, h_x, h_y and h_z are the grid spacing in x, y and z direction respectively and F_{max} is the maximum absolute evolving speed of all the points on grid. This means that the evolving contour cannot cross more than one grid at each time step.

Figure 3-26-a, b and c shows the results of major pulmonary artery segmentation results applied to different datasets. The left columns show different slices of three original datasets. The middle columns show the selected initial contours inside the pulmonary artery. The right columns represent the segmentation results of the major pulmonary artery using the proposed 3D level set algorithm.
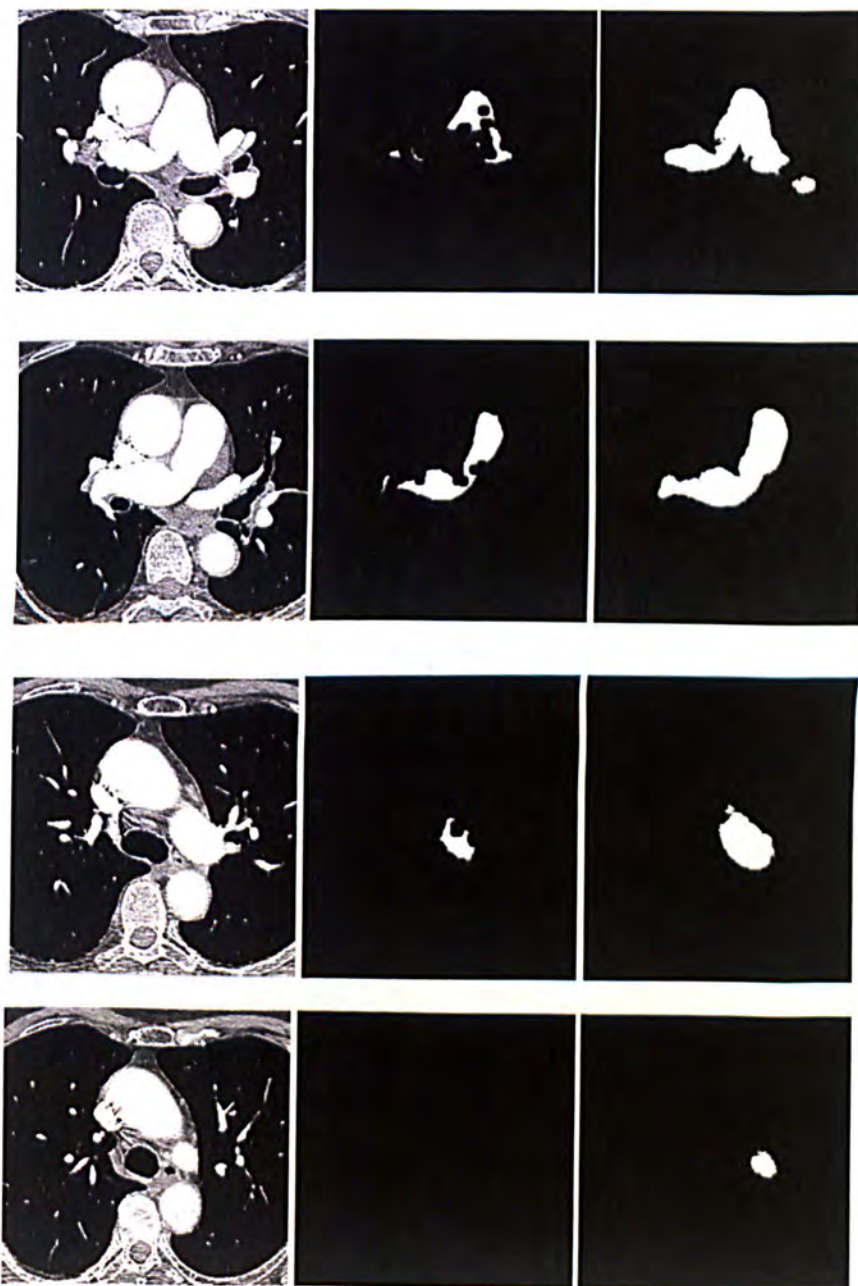
Figure 3-26-a: (a) Segmentation result using 3D level set method. Left: Different slices of the original dataset. Middle: Initial contours selected inside the pulmonary artery. Right: Segmentation results.
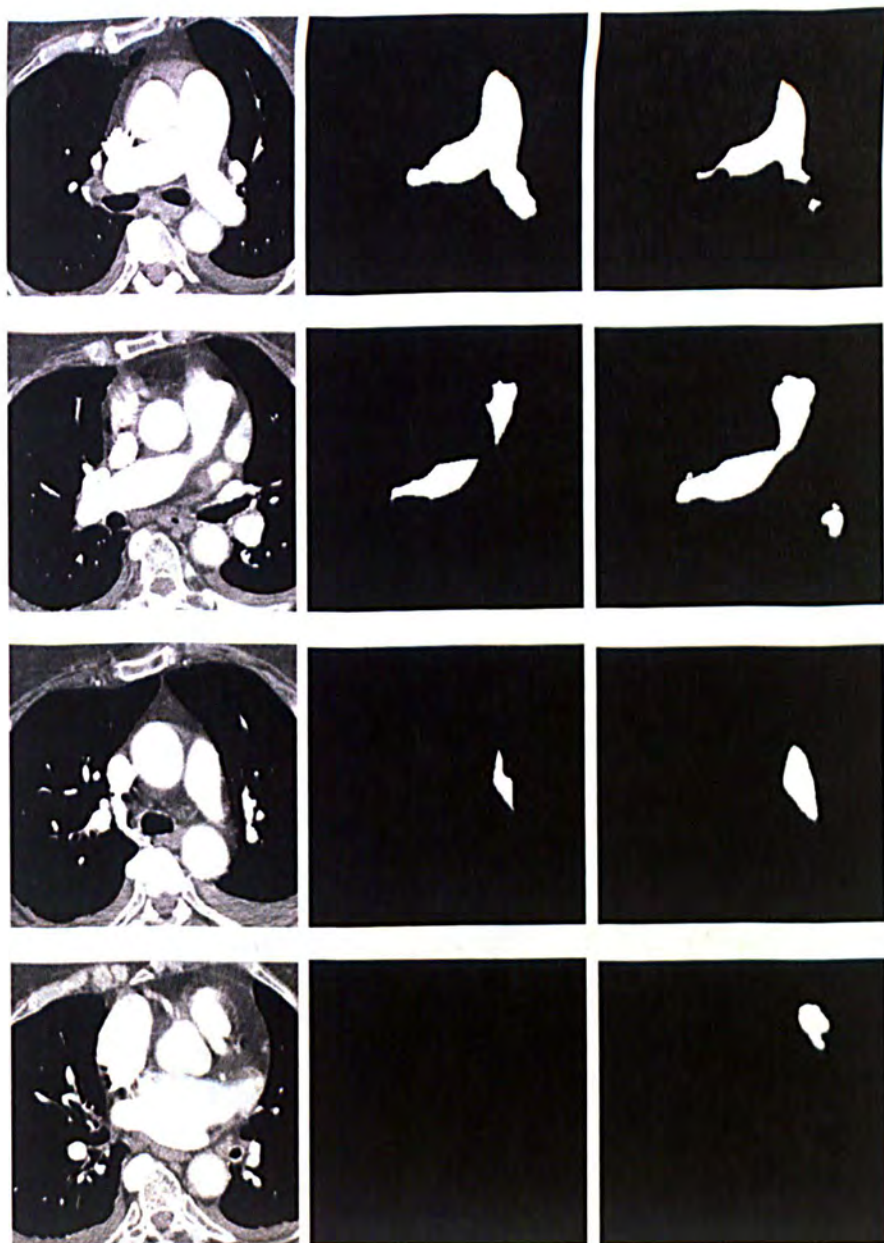
Figure3-26-b. Segmentation result using the 3D level set method. Left: Different slices of the original dataset. Middle: Initial contours selected inside the pulmonary artery. Right: Segmentation results.
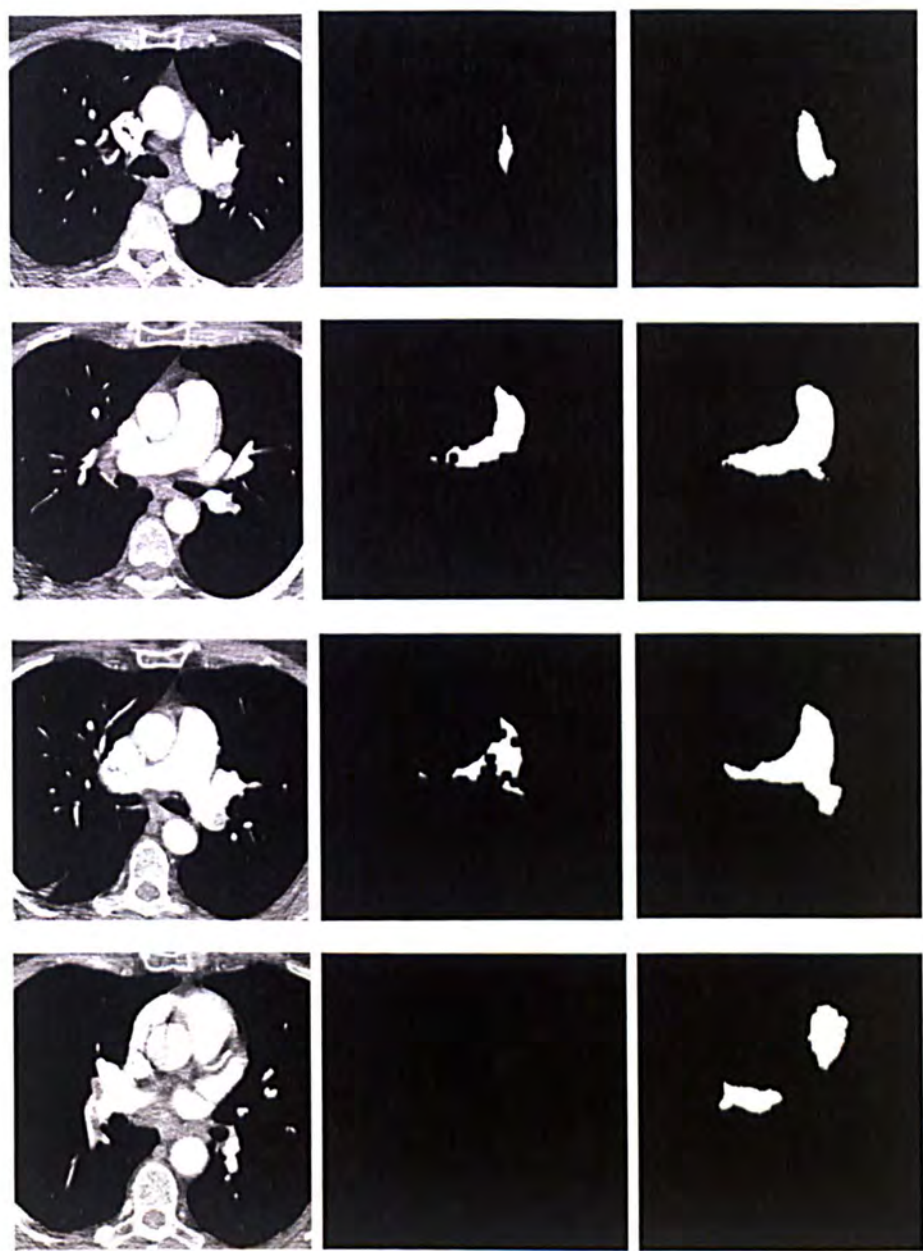
Figure3-26-c. Segmentation result using the 3D level set method. Left: Different slices of the original dataset. Middle: Initial contours selected inside the pulmonary artery. Right: Segmentation results.

Figure3-27 shows the 3D segmentation results of the major pulmonary artery in the heart region using the proposed level set algorithm.

Figure 3-27 Segmentation results of the major pulmonary artery using the 3D proposed level set algorithm for different datasets.

3.3.7. Peripheral Artery Segmentation

The peripheral pulmonary arterial network is separately extracted from the lung regions (left and right) using a region growing algorithm, starting from seed points which are end points of the left and the right artery and are automatically selected. The seed points are selected by applying an "AND" operator between the mask resulting from major artery segmentation in the heart region and the convex hull box containing the lung region. A pre-processing is first performed to enhance vascular structure to improve vessel visualization in volume rendering techniques and to facilitate the task of segmentation [56]. Due to the point-spread function (PSF) the high threshold value which is used for major artery segmentation cannot be useful for periphery artery segmentation and so, the threshold value for the peripheral artery is selected slightly smaller than that used in the heart region to extract the major artery. In the peripheral artery of patients with PE, the blockage can be so large that the artery is disconnected from the rest of the vessel network. So before applying the region growing algorithm, a tracking algorithm is used to connect the peripheral artery to the unconnected artery. This tracking algorithm is based on three steps, first segmentation of the peripheral artery (collecting voxels with CT number for the pulmonary artery), second the segmentation of the unconnected artery caused by a PE (collecting voxels with CT number for an embolus) and applying an OR operation to connect the segmentation artery and the dilated unconnected artery. The segmentation of unconnected artery can be used for PE candidate detection. Figure3-28 shows the diagram of the peripheral artery in lung region.
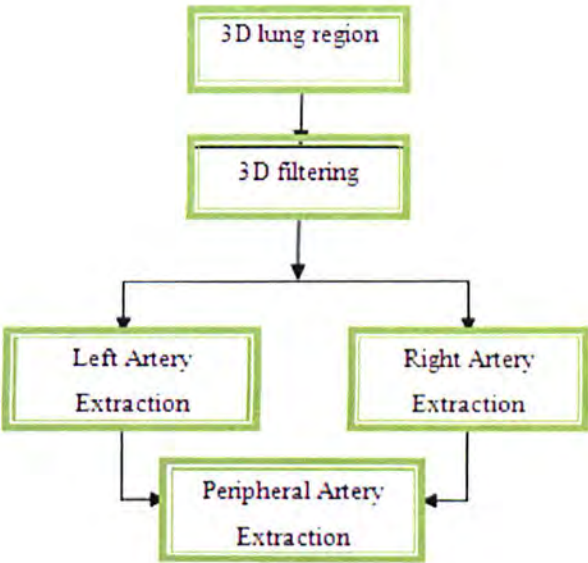
Figure 3-28 shows the diagram of peripheral artery segmentation algorithm.

3.3.7.1. Vessel Enhancing

Accurate quantification and segmentation of vascular structures such as the peripheral pulmonary artery (segmental and sub-segmental) is crucial in diagnosis, detection and treatment of vascular diseases. Vessel enhancement in medical images is needed as a pre-processing procedure to improve and facilitate the task of segmentation of vascular structures. Before extracting the peripheral artery in lung regions a 3D enhancing filter is first created to enhance vascular structures. Then the left and right extracted vessels are filtered using this filter to collect voxels related to vascular structures.

A few works have been developed to enhance the vascular structure. Frangi proposed a vessel enhancement filter by introducing a vesselness measurement based on geometrical analysis of all eigenvalues of the Hessian matrix [56]. Manniesing in [57] developed a method to enhance vascular structures which is known as a VED filter (Vessel Enhancing Diffusion). This filter is a combination of a smoothing vessel filter which is based on analysing eigenvectors and eigenvalues of the Hessian matrix and a non-linear anisotropic diffusion procedure. In fact, the VED is an enhancing diffusion filter, which is utilized in the framework of scale space theory to guide a diffusion process and measures the vesselness based on eigenvalues of the Hessian matrix.

We applied the following Frangi filter to enhance and collect voxels related to the vessel structure in the left and right lung regions.

$$V(\sigma) = \begin{cases} 0 & if \ \lambda_2 \geq 0 \ or \ \lambda_3 \geq 0 \\ \left(1 - e^{-\frac{R_A^2}{2\alpha^2}}\right).e^{-\frac{R_B^2}{2\beta^2}}\left(1 - e^{-\frac{S^2}{2\gamma^2}}\right) & otherwise \end{cases} \tag{43}$$

With range $V_s \in [0 \ 1]$, and the sensitivity of the filter is controlled by thresholds and constants α, β and γ and the ratios R_A , R_B and S are defined as follows:

$$R_A = \frac{|\lambda_2|}{|\lambda_3|} \ , \qquad R_B = \frac{|\lambda_1|}{\sqrt{|\lambda_2\lambda_3|}} \qquad , \qquad S = \sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}$$

in which R_A differentiates plate structures from line-like structures, R_B measures deviation from a blob shape structure and S differentiates between background and foreground (vessel). The vesselness measure is computed at different scales σ and has a maximum value when the scale σ is matched with the size of the vessel to be detected.

$$V = \max_{\sigma_{min} \leq \sigma \leq \sigma_{max}} V(\sigma) \tag{44}$$

where, the vessel structures are expected to be found between two maximum and minimum scales σ_{min} and σ_{max}.

Figure 3-29 shows application of the filter to different data set. The parameters α and β are considered fixed values ($\alpha = \beta = 0.5$) and the value of γ which depends on the gray level range on an input data set is also set at $\gamma = 1150$ in all results of the proposed work. The left column in Figure 3-29 shows the original data set and the middle column shows the results of peripheral artery segmentation algorithm using a region growing algorithm and the right column presents the results of applying the filter on segmented vessels. It can be seen that in the middle column the pulmonary artery is connected to the other organs, such as airways, via soft tissue between the pulmonary artery and trachea. But as shown in the right column, these soft tissues which are related to non-vascular structure were removed by applying the proposed filter. The removal of soft tissues is crucial to reduce the false positive detection rate (chapter 4).
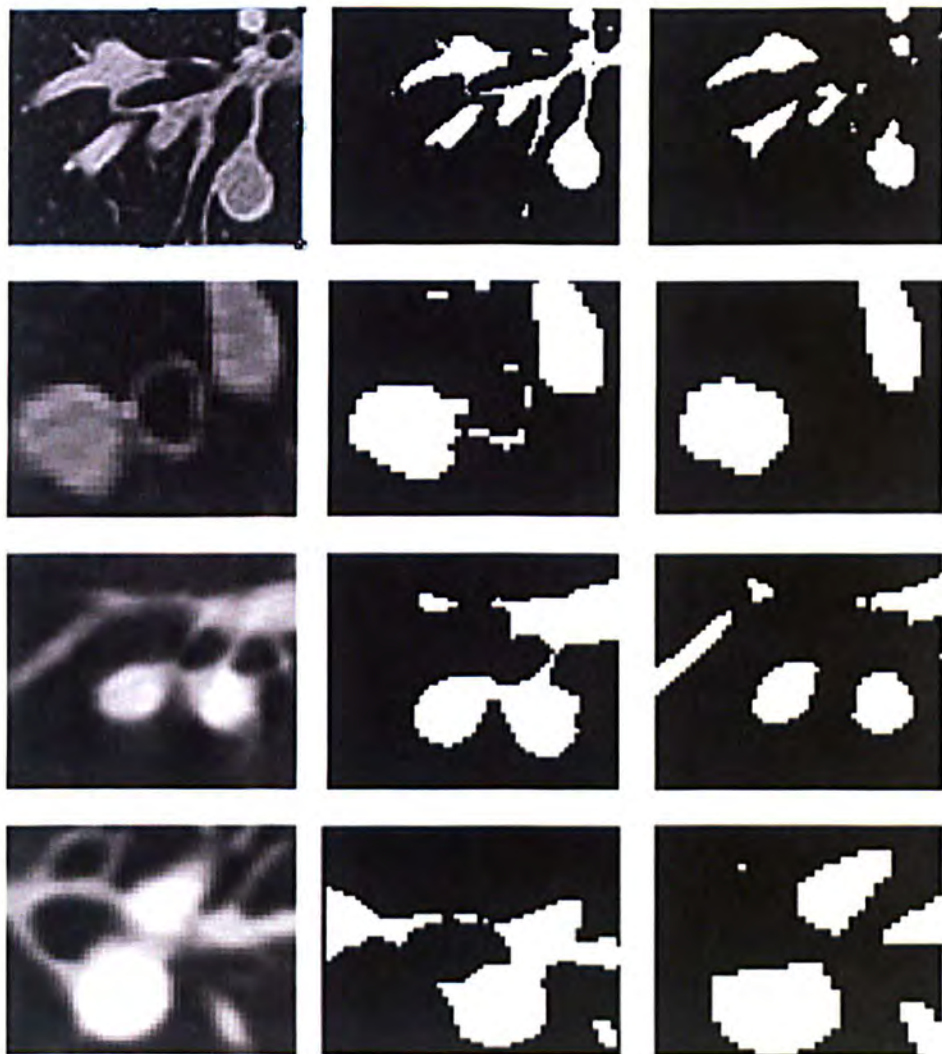
Figure 3-29 Left column shows 2D contrast CT images of peripheral vessel in the lung region. Middle column shows the peripheral artery segmentation results using a region growing algorithm and the right column presents the results using the enhancement filter.

3.3.7.2. Left and Right Peripheral Artery Segmentation

The left and right peripheral arteries are automatically and separately segmented in the left and right lung regions using a distinct region growing algorithm. The region growing algorithm is started from seed points which are automatically selected in the left and right arteries. The seed points are selected by applying an "AND" operator between the mask resulting from major artery segmentation in the heart region and the convex hull box containing the lung region. The region growing algorithm collects all enhanced voxels related to pulmonary artery tree in the lung regions. Figure 3-30 shows the result of the peripheral artery segmentation algorithm.
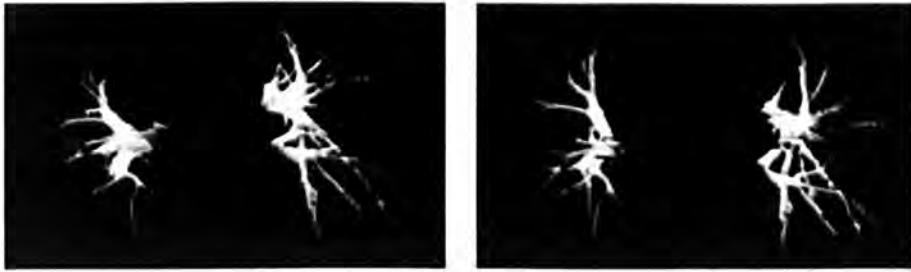
Figure 3-30 shows periphery artery segmentation results in different datasets.

The major pulmonary artery, the peripheral artery and the path to unconnected arteries are all combined with an OR operation as a search area for pulmonary embolism candidate detection.

Figure3-31 shows the final result of pulmonary artery tree segmentation (Figure3-31-b the major artery in heart region, (c) the peripheral artery in the left lung region, (d) the peripheral artery in the right region). Based on the new stopping criterion the optimum iteration number is $N_0 \approx 80$ where the variation of the mean energy profile falls to zero (Figure 3-31-a).

(a)

(b)

(c)

(d)
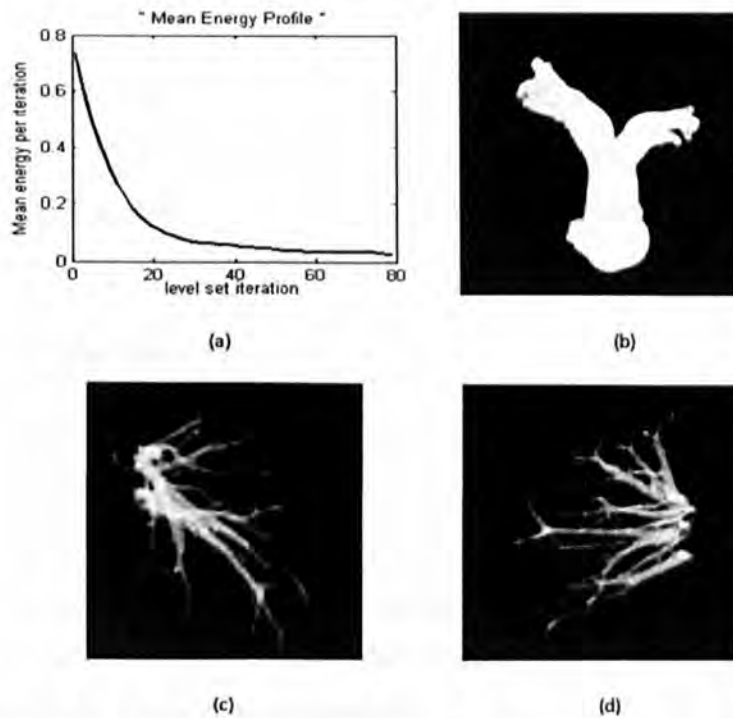
Figure 3-31 : (a) Shows the mean energy profile of the level set algorithm, (b) the completed segmentation of the major pulmonary artery in the heart region, (c) the peripheral artery in the left lung region, (d) the peripheral artery in the right region.

3.4. Summary

In this chapter one of the major components of the proposed CAD system (PA segmentation) has been demonstrated. To segment the pulmonary artery which is the search area for pulmonary embolism (PE) a fully automatic algorithm has been developed and the major components of PA segmentation algorithm are:

Lung and Heart Segmentation

The lung and heart region were extracted from the raw data to reduce the search area for segmenting the pulmonary artery. A fully automatic technique has been proposed to segment the lung region in CT images which is based on three main steps. In first step lungs region is extracted using an optimized thresholding method from CT images and then the left and right lungs are separated by identifying anterior and posterior junctions and finally the boundary of the lungs region will be smoothed. The validation of the proposed method is evaluated in chapter 5. The heart boundaries have been identified by extracting and connecting four sharp points of the left and right lungs.

Feature Map Extraction

The main issue to segment the pulmonary artery tree is connectivity between the pulmonary artery and the surrounding structure such as aorta, superior vena cava and vein. A feature map was first constructed using the eigenvalue of the Hessian matrix to remove the connectivity.

Pulmonary Artery Segmentation

The pulmonary artery tree has been separately extracted in the heart and lung regions. The major artery was first segmented by a region growing method initialized by a seed point which was automatically selected within the pulmonary artery trunk in the heart region, and then the segmentation result was refined using a new customized 3D level set algorithm. The level set algorithm applies the output of the region grower as the initial contour. The peripheral artery was segmented in the left and right lung regions using a region growing algorithm after enhancement by a vessel enhancing filter.

4. PULMONARY EMBOLISM DETECTION

4.1. Introduction

The first step to detect the pulmonary embolism is pulmonary artery segmentation which reduces the search area in order to reduce false positive detections. The second step is searching for a pulmonary embolism candidate inside the segmented artery and finally removing false positive detections by analysing the features of candidates. Some works based on feature extraction have been developed to detect true pulmonary embolisms and remove false positive detections. Masutani et al. [25] used feature based on the intensity, local contrast, volume and length of PEs to detect true PEs and removing false positive detections. Zhou [26], [27] proposed a CAD system to detect PEs by applying features based on intensity, edge strength, length and volume. Pichon [58] used features based on the intensity and size. He highlighted potential PEs on a 3D representation of the vessels by projecting the intensities inside the vessel and selecting first quartile of intensities between the medial axis and the surface of vessel. H. Bouma et al [34] proposed a CAD

system for the classification of pulmonary embolisms inside the segmented vessels using new features such as location and shape of vessel.

In all previous proposed CAD systems all the pulmonary vessels were selected as search areas to detect PE, which increases false positive detection rate. Moreover in the segmentation process, the regular shapes of vessels are not completely extracted and emboli that are connected to the vessel wall can be missed (see Figure 4-1) during the segmentation process. So by extracting the pulmonary artery as the only place to find the pulmonary embolism, and by analyzing the shape of segmented artery and the shape of lumen as a further feature of PE to detect the missed PE, we improve the sensitivity of detection. We also proposed a new feature which is the location of PE candidate to remove false positive detection. Most false positive detections come from soft tissues which are located between a bronchus and an artery, against the vessel wall and have the same intensity of PE and might be interpreted as PEs. By tracking the segmented trachea and its branches (bronchus), false positive detections derived from soft tissues are removed.
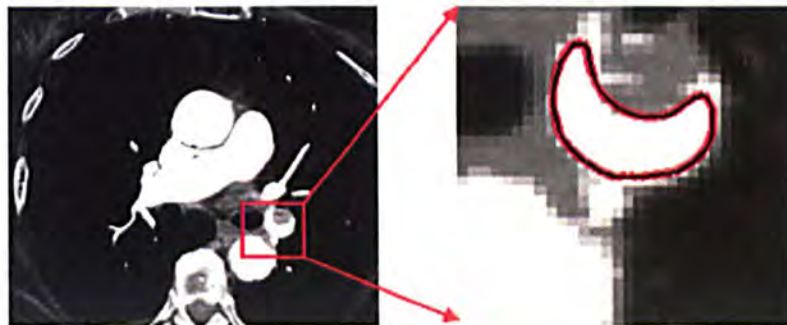
Figure 4-1 shows the segmentation results for the pulmonary artery. The acute pulmonary embolism was missed in the segmentation process.

In the contrast enhanced CT images, pulmonary embolism (PE) can be identified as a dark region inside the segmented artery. However, other tissues or regions can have similar intensities. These are associated with the partial volume effect (voxels on the boundary of the segmented artery), parenchyma disease, streak artefact radiating from the superior vena cava, noise and lymphoid tissue (see Figure 4-2). To reliably detect the PE a group of voxels inside the segmented artery are extracted as an initial candidate for pulmonary embolism and subjected to the following analysis to distinguish true PE's from look-alikes (false positives).

(a)   (b)   (c)
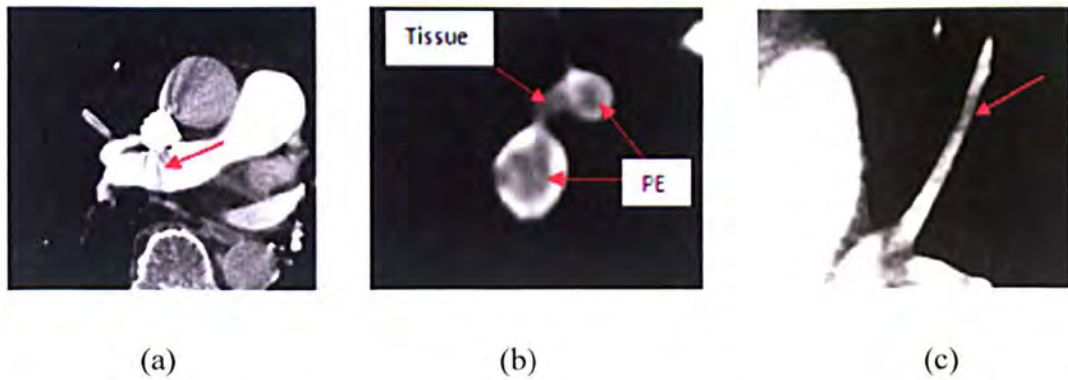
Figure 4-2: shows dark regions which are not caused by pulmonary embolisms but have the same appearance. (a) Artefacts radiating from the superior vena cava. (b) Lymphoid tissue. (c) Flow void

This chapter consists of two main sections, 1) PE candidate detection section which investigates five different PE detectors to find voxels that may be related to PE. These five different PE detectors respond to different characteristics of Pulmonary embolism and 2) False positive reduction, which is based on analysis of extracted features to remove false positives.

4.2. PE Candidate Detection

The aim of PE candidate detection is to find voxels inside the segmented artery which have the same intensity as PEs. Five different detectors are used to reject false positives or accept true PE's. The five PE detectors respond to the different properties of PEs. Four of the detectors are based on intensity search voxels inside segmented artery and the other is based on analysing the shape of a segmented artery to find PE candidate voxels outside the artery.

4.2.1. First PE Detector (Based on Intensity)

The first PE detector is based on CT value. A PE inside the segmented artery is darker than the contrast-enhanced blood, so using an optimal threshold the dark objects can be discriminated from the bright objects. This detector collects voxels as initial PE candidates if the original CT values of the voxels are inside the range between -25 Hounsfield units (HU) and 75 HU (see section5.3.3). Figure4-3 shows the results of PE candidate detection using first the PE detector.

The results of the first PE candidate detector show that this detector does not respond to all types of PE. As Figure4-3 shows the small PE inside the pulmonary artery (indicated by the red circle) was not detected by this detector.

Figure 4-3 shows the result of the first PE detector. Left: Original image in which the un-detected PE is arrowed by red circle. Right: The binary mask created by the first PE detector.

However, this detector cannot distinguish between embolism and lymphoid tissue or soft tissues which are located between the pulmonary artery and airway system. Figure4-4 shows the results of the first PE detector in which the detector cannot discriminate between embolism and soft tissue.

(a)   (b)   (c)

Figure4-4: shows applying first PE candidate detector based on CT value. (a) The original image which shows pulmonary embolism inside pulmonary artery and soft tissue between pulmonary artery and main branch of bronchus. (b) The mask derived from pulmonary artery segmentation results. (c) The result of PE candidate detector which cannot discriminate between embolism and soft tissue.

4.2.2. Second PE Detector (Based on Geometric Analysis)

As the result of first PE detector showed, it does not respond to some characteristics of PE. For detecting all potential PEs, other detectors are examined. The second detector is based on differential geometry and uses eigenvalues of the Hessian matrix to detect a dark region in the segmented artery [25]. The second derivatives of multi-scaled image at scales between δ_{min} and δ_{max} are computed in three orthogonal directions which have the directions of maximum second derivative values. Since in an enhanced CT image a pulmonary embolism has a CT value less than its background, the values of the second derivatives are positive in the orthogonal directions. So, if the eigenvalues are sorted by decreasing magnitude ($|\lambda_1| > |\lambda_2| > |\lambda_3|$) then a dark spot can be detected by a positive value of the first eigenvalue λ_1 (Figure. 4-5).
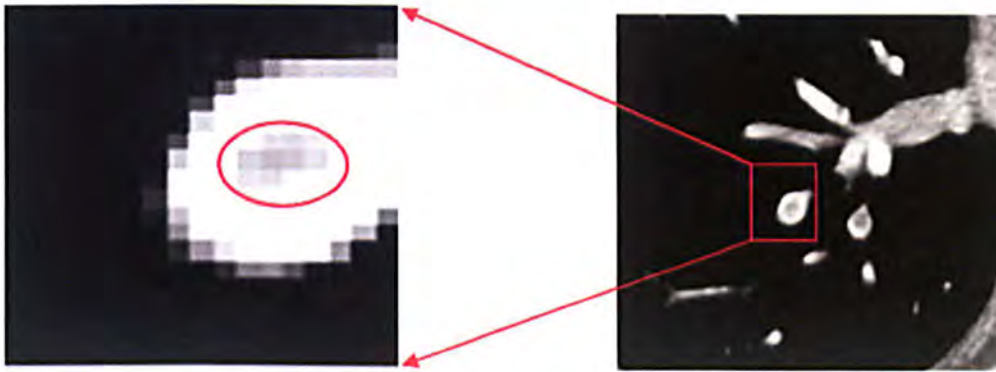
-83.5367	-134.7267	-150.9208	-147.1207	-137.8511	-125.9204
-131.3239	-135.9728	-108.4126	-84.8553	-76.1775	-81.3435
-141.2881	-90.0965	-44.3794	-28.1444	-28.7862	-52.2649
-112.2902	-26.3698	24.6454	30.5725	19.2918	-45.2494
-74.0878	27.6435	38.9811	34.8827	30.3940	-51.8968
-60.1247	29.9614	24.6409	23.6912	21.2454	-65.3091
-73.8201	-21.1435	14.0866	-14.0048	-31.1917	-89.6663
-109.3547	-68.6476	-41.5470	-34.2958	-67.6573	-136.3869
-148.0075	-129.0235	-109.0358	-109.5449	-141.4370	-162.6978
-137.1005	-149.8246	-154.1481	-157.5354	-146.9865	-96.9911

Figure 4-5 Shows the result of PE detector using eigenvalues of the Hessian matrix. Positive first eigenvalues (numbers inside the red closed curve) identify a PE candidate.

This detector responds to detect dark areas (potential PEs) which are surrounded by bright lumen. Figure 4-6 shows the results of applying the second detector, locating a candidate PE which was not detected by the first detector.

Figure 4-6 shows the result of the second detector which was missed by the first detector

4.2.3. Third PE Detector (top-hat Operator)

The third PE detector utilizes the morphological black top-hat operator [59] or local contrast [25]. The black top-hat transform is a gray level mathematical morphology operator used to separate voxels, which can be as a PE voxels, whilst removing noise. The top-hat transform consists of two operations, first a gray level morphological closing and then subtraction of the original image.

top-hat transform = gray level morphological closing operator – original image

The gray level morphological closing operator removes dark structure and the subtraction operator enhances the dark region. Figure4-7 shows the results of applying the top-hat operator on two different datasets. The dark regions are extracted when the top-hat transform is applied inside the segmented artery. Since the boundary of the segmented artery has a lower intensity than the centre of the artery so the dark boundary of segmented artery is also extracted by applying the top-hat transform. In order to reduce false positive detections, the boundaries of the segmented artery are removed by a morphological erosion operator.
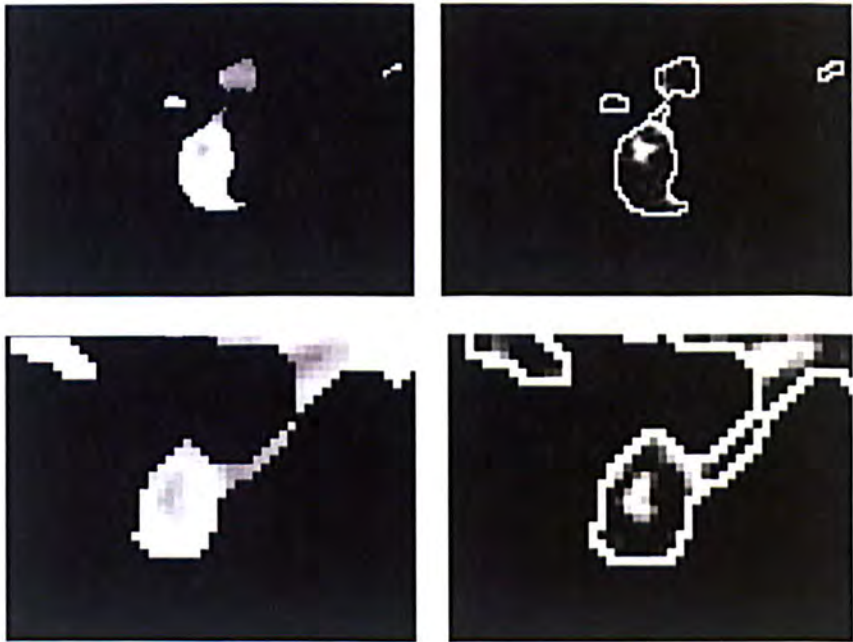
Figure 4-7: shows the results of using 3D top-hat morphological operator on the original dataset. The left column shows the original dataset and the right column represents the results of top-hat operator. As the results show, the dark area inside the artery were enhanced and the dark boundary segmented artery is also extracted which can be removed by an erosion operator.

Figure 4-8 shows the results of applying of the three first detectors. Figure4-8a shows the original dataset with a PE and Figure4-8b represents the enlarged image of the small PE inside the artery which was missed by the first detector (Figure4-8c). Figure4-8d and Figure4-8e show the detection results used by the second and third PE detectors, respectively, and finally Figure4-8f is a mask generated by a logical OR of three detectors.
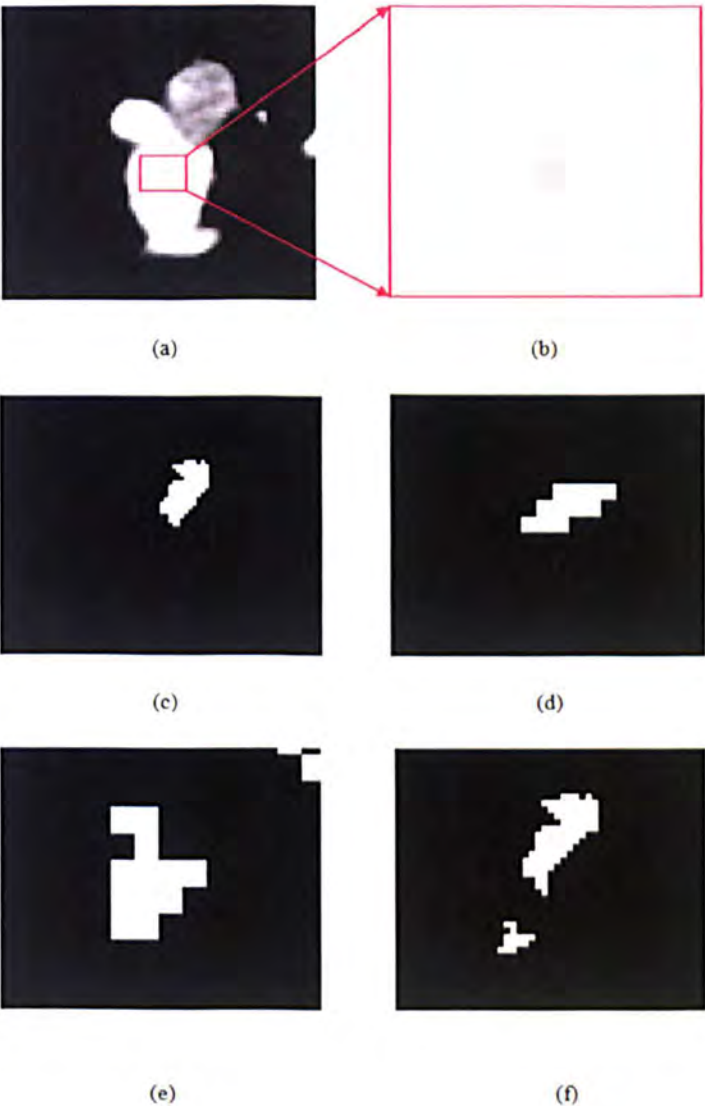
Figure 4-8: shows the results of PE detection by the first three detectors. (a) The original dataset with PE. (b) The enlarged picture of a small PE inside the pulmonary artery. (c) Detection result by the first detector which missed to detect the small PE and just detected the big one. (d) the mask resulted from the second PE detector. (e) The detection result using the top-hat morphological operator and (f) represents the mask generated by a logical OR of three PE detectors.

4.2.4. Fourth PE Detector (Based on Perpendicular Plane on Centreline and Shape Index)

The fourth PE detector uses shape analysis and is specialized for detecting an embolus connected to the vessel wall which may be missed by the previous detectors. The first three detectors search for a PE inside the segmented artery but in the segmentation

process, the regular shape of the artery is not completely extracted and emboli that are connected to the vessel wall can be missed (see Figure 4-9). In these cases we analyze the shape of the segmented artery and the shape of the lumen as a further feature of PE candidate detection.

(a)   (b)   (c)
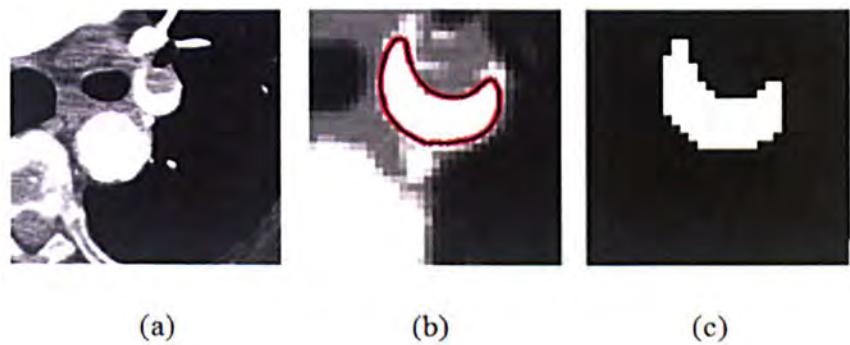
Figure 4-9 (a) Original image of vessel with PE (b) segmentation result (c binary mask of segmentation has an irregular shape which results in the embolism being overlooked.

The first shape feature based on the shape of the segmented artery, measures the eccentricity of the arterial cross-section. A healthy pulmonary artery has a circular cross-section perpendicular to the centreline of the pulmonary artery, but an embolus can result in this appearing non-circular. The non-circular shape of a cross-section of an artery can be found from the eccentricity of the shape. The eccentricity is a parameter which measures deviation of a conic section from being circular, and takes a value between 1 (a line segment) and 0 (a circle). Cross-sections with an eccentricity less than 0.5 are selected for further analysis. The convex hull points of non-circular cross-section are extracted and used to estimate a best-fitting ellipse. Subtracting the non-circular artery from the ellipse after thresholding (between -25 HU and 75 HU) will be selected as seed points for a region growing algorithm to connect PE candidate voxels in 3D CT images. Figure 4-10 shows the flow diagram of applying the fourth PE detector.
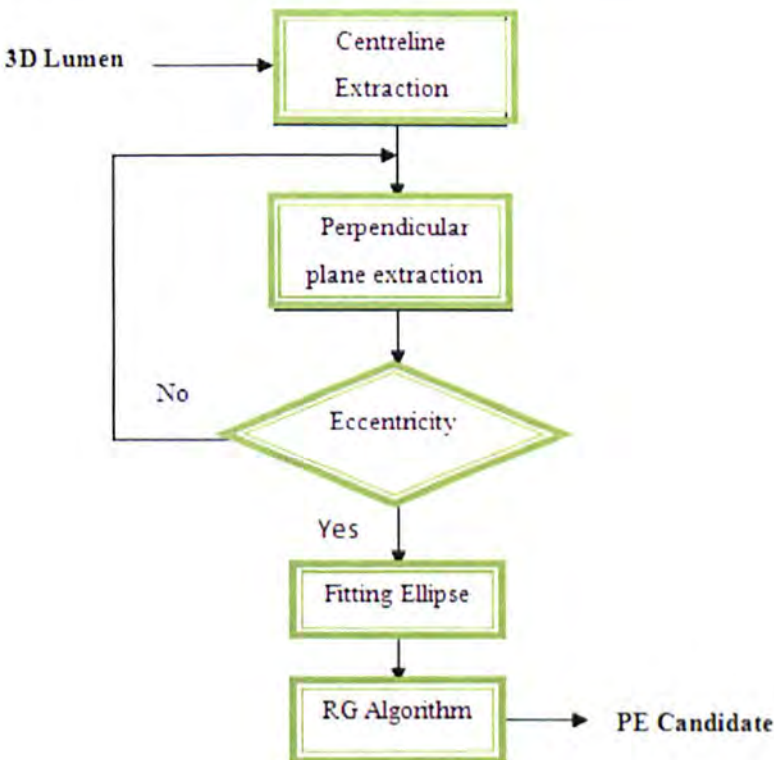
Figure 4-10 shows the flow diagram of the fourth PE candidate detector algorithm.

4.2.4.1. Centreline Extraction

Skeletonization or centreline extraction is a process to produce a central spanning tree to represent the shape of an object in two or three dimensions [94]. In fact, a skeleton of an object is a concise representation of the shape of an object. In medical imaging, the centrelines of tubular structures are applied in order to visualize and analyse tubular structures such as blood vessels, colon and airway structure (trachea and bronchi). For a navigation path planning task, for example to control the movement of a virtual camera inside the human colon, an accurate skeleton of the colon can be very useful [95], [96]. In order to make quantitative measurements such as length, volume and radius of a human organ, the skeleton of the organ can aid measurement [97].

In the proposed CAD system, we extract the centreline of the segmented artery to find the radius of cross-section perpendicular to the centreline of the pulmonary artery to determine the circularity of the cross-section. Several algorithms have been proposed to extract the centreline or skeleton of an object in 2D or 3D. In many skeletonizaton techniques, the centreline of an object is extracted by using a thinning algorithm in which

an object is eroded layer by layer until just a skeleton of the object is left [98]. With thinning algorithms, voxels or pixels are converted to background points while the geometry and the topology of object are not changed. In this method, first the object is segmented and then the skeleton of the mask (resulting from the segmentation process) is extracted. Since the thinning algorithm is grid-based it produces voxel precision results.

To extract the skeleton of a segmented pulmonary artery, we used the method proposed in [99] which is a sub-voxel precision method to extract the skeleton of an object using the fast marching algorithm. Based on this method, in the first step, we segmented the pulmonary artery using a level set algorithm in which the result of the level set algorithm is a sub-voxel precision representation of the original data. The second step, to extract the centreline of the pulmonary artery, we used the fast marching algorithm to calculate a sub-voxel precision of the Euclidean distance field. In fact, in our algorithm the input data is a sub-voxel precision distance field and the result of skeletonization process is a smooth skeleton at sub-voxel precision. A voxel-precision algorithm cannot compute the skeleton of an object that is less than a single voxel thick, but a sub-voxel precision algorithm allows that to be computed. The sub-voxel precision skeleton is more accurate for measurement of an object such as volume or finding vessel cross section and also is a better starting point for virtual navigation [99].

4.2.4.2. Background

In the variation minimum-cost path problem, a curve:

$$C(t): [0, \infty) \to \mathbb{R}^n$$

is found so that it minimizes the cumulative travel cost from a starting point M to another point X in $\mathbb{R}^n$. If the cost H is a function just based on the location X in the given domain, then the minimum cumulative cost at X is defined as follows:

$$T(X) = \min_{C_{MX}} \int_M^X H(C(t)) dt \tag{1}$$

Where, C_{MX} is the set of all curves which links point M to point X and T(X) is the minimum cost for travelling from point M to point X. The curve at which the integral is minimized is called the minimum path. In geometric optics it is shown that the solution of the equation (11) satisfies the Eikonal equation:

$$|\nabla T(X)| F(X) = 1 \tag{2}$$

In the above equation, T(X) can be interpreted as the arrival time of a wave when it crosses the point X with speed $F(X)$ which can be identified as:

$$F(X) = \frac{1}{H(X)} \qquad (3)$$

If the cost function H gives a higher weight for medial points of an object than non-medial points then it can be shown that the minimum path between two medial points (voxels) is a curve skeleton.

4.2.4.3. Skeleton Extraction Algorithm

Figure 4-11 shows the diagram of the main steps of the sub-voxel precision skeleton extraction algorithm. In this algorithm, a sub-voxel precision distance field is as the input data to compute the point which has the largest distance from the boundary of the object (the point with global maximum distance) and to find the speed function as an input for the fast marching algorithm with a step size of unity. The fast marching algorithm is applied to compute the geodesic distance inside the object starting from the global maximum point of the distance field. We used the fast marching algorithm proposed in [100] which is Multi-stencil Fast Marching Method (MSFM) to calculate the shortest distance from a list of points to all other pixels in an image volume. This method gives more accurate distances by using second order derivatives and cross neighbours. The furthest point from the global maximum point of distance field is selected as a start point of the branch. The remaining points of the branch are then calculated by applying the following gradient descent back-tracking procedure on the fast marching time-crossing map:

$$\frac{dC}{dt} = -\frac{\nabla T}{|\nabla T|} \qquad C(0) = P_f \qquad (4)$$

where, $C(t)$ is the centreline of object and P_f is the furthest geodesic distance point from global maximum distance point to the object boundary. The gradient descent allows the tracking of the branch to be performed only along the local gradient of the time crossing map toward the global maximum point from object boundary [99].

By repeating this process for each branch of the skeleton the centreline of object can be extracted.
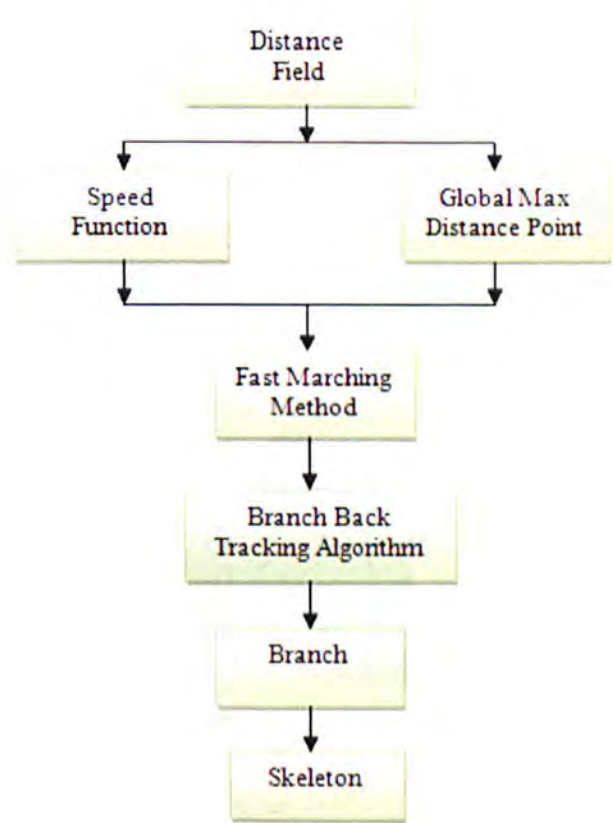
Figure 4-11 shows the diagram of the main steps of sub-voxel precision skeleton extraction algorithm.

Figure4-12 shows the results of the centreline extraction algorithm on original medical images and some synthetic images.

(a)

(b)

(c)
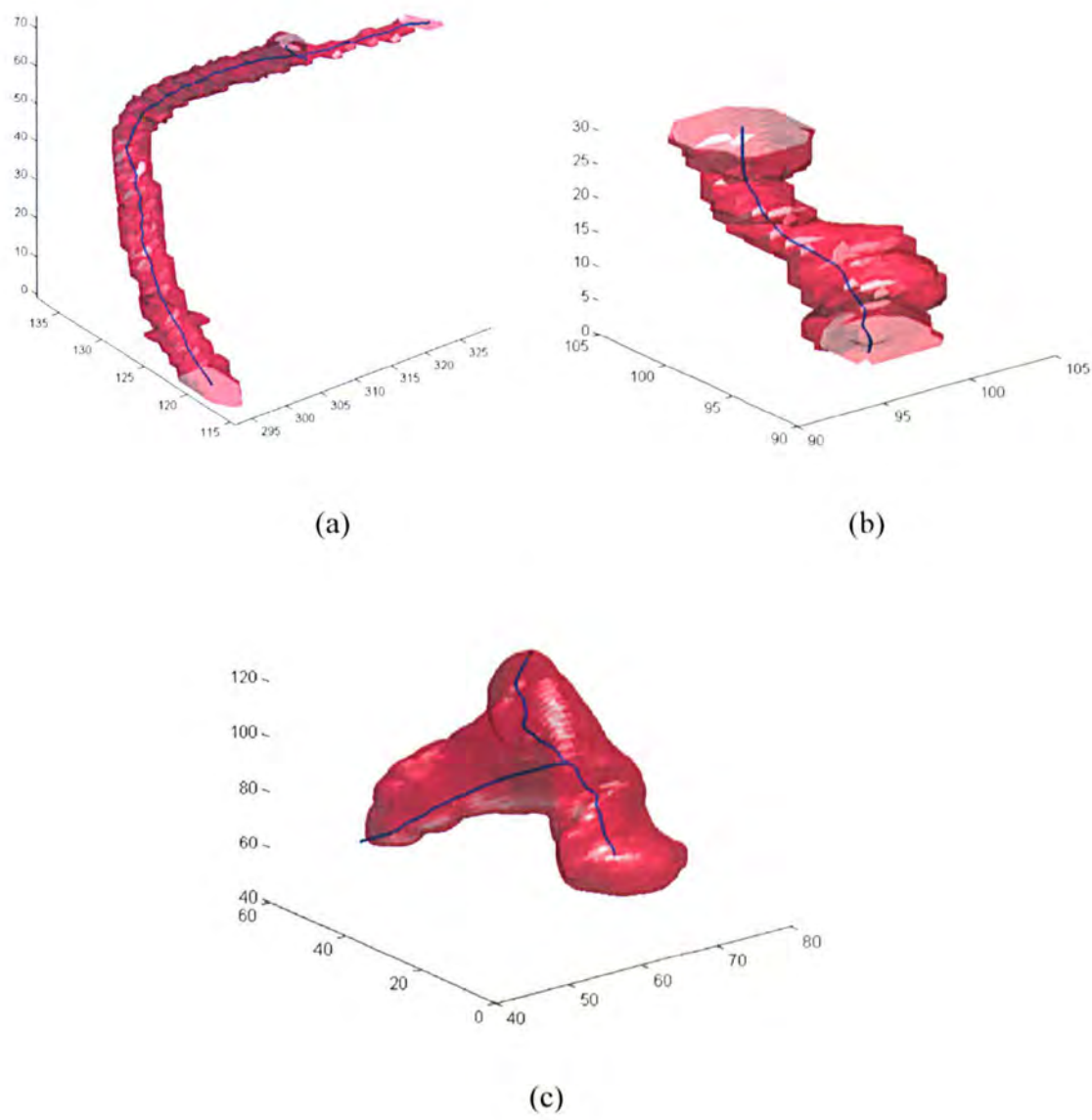
Figure 4-12 shows centreline extraction results by the proposed algorithm. (a) Centreline extracted for small section of the femoral artery. (b) Tibial artery and (c) the centreline extracted from the major pulmonary artery in the heart region.

4.2.4.4. Orthogonal Plane Extraction

The analysis of the oblique planes, especially planes perpendicular to the centreline of a vessel is more likely to improve the performance of the automatic detection of a disease. All CT scanners which produce computed tomographic pulmonary angiography (CTPA)

images provide axial images. So in order to improve the detection rate of pulmonary embolisms, the digital plane orthogonal to the major pulmonary artery is extracted and then is analyzed.

Definition1: (Digital Straight Line)

A digital straight line is a subset of $\mathbb{Z}^2$ described by the following equation:

$$L(a,b,\gamma,\rho) = \{(x,y) \in \mathbb{Z}^2 | \; \gamma \leq ax + by < \gamma + \rho\} \tag{5}$$

where, all parameters are integers , γ is the affine offset of the line, ρ is the arithmetic thickness and (a, b) shows the direction of the digital line.

By letting $\rho = max\{|a|, |b|\}$, an interesting subset of the digital line which is called a naive digital line is extracted.

$$L(a,b,\gamma) = \{(x,y) \in \mathbb{Z}^2 | \; \gamma \leq ax + by < \gamma + max\{|a|, |b|\}\} \tag{6}$$

One of the most important properties of a naive digital line is strict 8-connectivity [119].

Definition2: (Digital Plane)

A digital plane is a subset of $\mathbb{Z}^3$ described by the following equation:

$$P(a,b,c,\gamma,\rho) = \{(x,y,z) \in \mathbb{Z}^3 | \; \gamma \leq ax + by + cz < \gamma + \rho\} \tag{7}$$

where, all parameters are integers , γ is the affine offset of the plane, (a, b, c) is the normal direction of the plane and ρ is the arithmetic thickness which controls the thickness of the plane [136].

- $\rho < max\{|a|, |b|, |c|\}$: The plane has 6-holes.
- $\rho = max\{|a|, |b|, |c|\}$: The plane is strictly 18-connected.
- $\rho = |a| + |b| + |c|$: The plane is strictly 6-connected.
- $\rho > |a| + |b| + |c|$: The plane is thick.

The naive digital planes which have a digital 18-connected counterpart and have no holes for 6-connectivity defined by an equation of the form:

$$P(a, b, c, \gamma) = \{(x, y, z) \in \mathbb{Z}^3 \mid \gamma \leq ax + by + cz < \gamma + max\{|a|, |b|, |c|\}\} \quad (8)$$

The perpendicular plane to be extracted is identified by a point in the input data (centreline points of the pulmonary artery constitute the input data) and a normal vector. The point $M(ox, oy, oz)$ is considered as a point from the major pulmonary artery centreline and $N(a, b, c)$ is the normal vector which is the first derivative of the centreline points at the point $M(ox, oy, oz)$. The first derivative of the centreline points represents the local orientation of the pulmonary artery. The parameters a, b and c determine on which face to project the perpendicular plane. For example, if $a = max\{|a|, |b|, |c|\}$ then the perpendicular plane will be projected onto face (O, y, z), on face (O, x, z) if $b = max\{|a|, |b|, |c|\}$ and onto face (O, x, y) if $c = max\{|a|, |b|, |c|\}$. Figure 4-13 shows the result of perpendicular plane on centreline extraction.
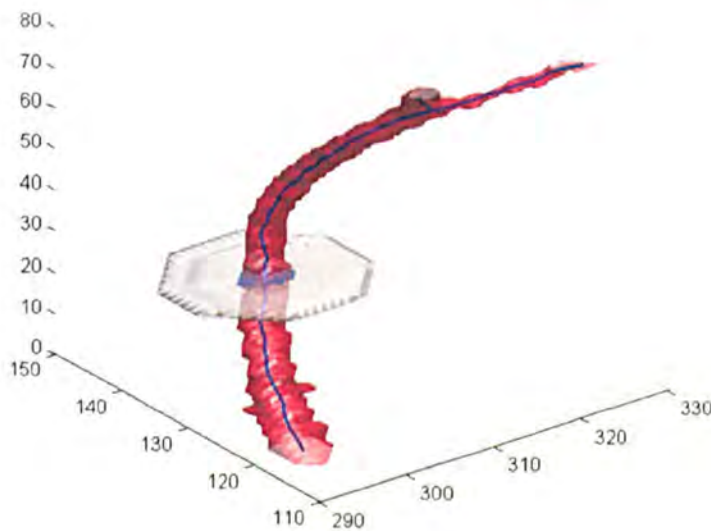
Figure4-13: Perpendicular plane extraction result on centreline of a small section of femoral artery.

The eccentricities of projected planes are computed and those whose eccentricity is less than 0.5 are selected as positions of potential PEs. The convex hull points of the non-

circular object are extracted (Figure4-14a) to estimate a best-fitting ellipse (Figure4-14b). Subtracting the non-circular artery from the ellipse after thresholding will give seed points for the region growing algorithm to detect PE candidate voxels in 3D CT images (Figure 4-14c).

(a)   (b)   (c)
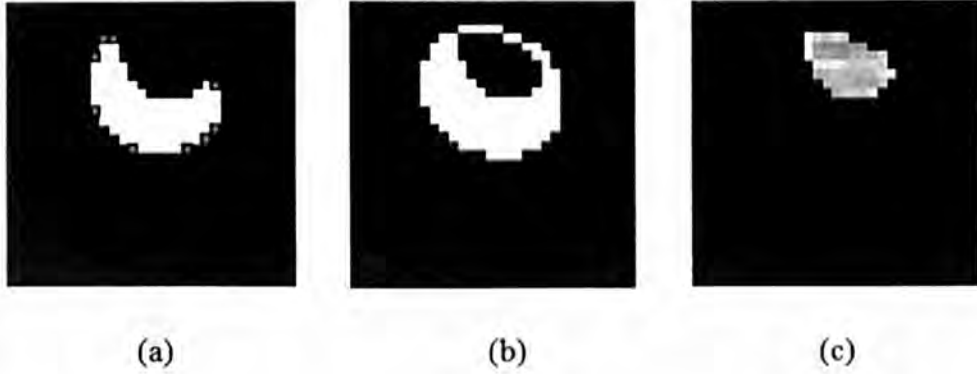
Figure 4-14 Shape analysis of the segmented artery, (a) convex hull points identified on the vessel boundary; (b) the fitted ellipse; (c) the subtraction of non-circular shape from the fitted ellipse shape.

The second feature, based on the 3D shape of the lumen, utilizes the isophote curvature components (maximum and minimum curvatures) to express the local shape of a surface through points of equal intensity. In order to encode the curvature information of the surface two continuous measures or parameters will be used: a shape index, SI, and a curvedness C. These two parameters are applied to quantify and qualify local 3D shape. By using these two measures the shape of the lumen is analyzed to detect the potential PE.

Shape Index: The shape index is an angular measure for describing the local surface topology in terms of maximum and minimum curvatures (principal curvatures). The shape index developed by Koenderink and Doorn [60] at a point of a surface is defined as:

$$SI = \frac{2}{\pi} \arctan\left(\frac{\kappa_2 + \kappa_1}{\kappa_2 - \kappa_1}\right) \quad \kappa_1 \geq \kappa_2 \tag{9}$$

where κ_1 and κ_2 are the two principal curvatures which can be calculated as follows [61]:

$$\kappa_1 = H + \sqrt{H^2 - K} \tag{10}$$

$$\kappa_2 = H - \sqrt{H^2 - K}$$

where H and K are the Mean and Gaussian curvatures. Every distinct shape has a unique shape index value, except a planar shape ($\kappa_1 = \kappa_2 = 0$ on the planar shape).

The range of the shape index is between -1 and 1. A shape index of 1 belongs to those convex surface points which have equal principal curvatures, while the shape index of concave surface points with equal principal curvatures are identified by -1. The shape index measure of 0 is for those surface points (saddle surface points) whose principal curvatures are equal magnitude but with opposite sign. Valley-like surface points have a shape index value of -0.5 and ridge-like surface points have a shape index of 0.5. Table 4.1 shows the different bounds for different classes of surface points.

Table 4-1 Shape index of different surface points

Surface Points	Type of Region	Shape Index Bounds
Dome	Elliptic	[0.625, 1)
Ridge	Parabolic	[0.375 , 0.625)
Saddle Ridge	Hyperbolic	[0.125 , 0.375)
Plane	Hyperbolic	Undefined
Saddle Point	Hyperbolic	[-0.125 , 0.125)
Saddle Rut	Hyperbolic	[-0.375 , -0.125)
Rut	Parabolic	[-0.625 , -0.375)
Cup	Elliptic	[-0.625, -1)

The shape of the lumen surface in a healthy artery (a tubular section) is ridge-shaped (convex in one direction and straight in the orthogonal direction) or saddle-shaped on a bifurcation (convex in one direction and concave in the other direction). The shape of the

surface of an artery with a PE is like a valley (concave in both directions) (Figure4-15). The isophote curvature components are used to distinguish the valley-shaped surface of the lumen which could be candidates for PE location [34].

(a) (b) (c) (d) (c)
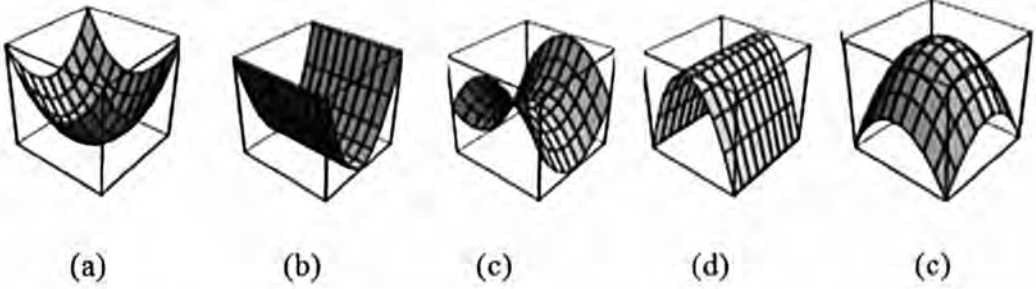
Figure 4-15 Shape index of lumen of segmented artery can be used to describe the local shape of the lumen. (a) the cup shape (-1 for shape index). (b) Ridge shape for healthy vessels (-0.5 for shape index). (c) Saddle shape, which is convex in one direction and concave in the other direction for bifurcation in vessels (0 for shape index). (d) ridge shape for healthy vessels (0.5 for shape index). (e). Cap shape (1 for shape index).

Curvedness: The curvedness of a point of a surface shows how gently or highly curved a surface can be and is estimated as:

$$C = \sqrt{\frac{\kappa_1^2 + \kappa_2^2}{2}} \tag{11}$$

The curvedness is applied to encode the scale differences between objects with the same shapes (e.g. distinguishing between a cricket ball and a soccer ball). A shape index of less than -0.625 is related to points which belong to concave surfaces and a curvedness of more than 0.1 eliminates points associated with flat surfaces.

4.2.5. Fifth PE Detector (Path Tracking)

In the peripheral arteries of patients with PEs, the blockage can be so large that the artery appears un-connected from the rest of the vessel network. So in the peripheral artery segmentation step, before applying the region growing algorithm, a tracking algorithm was used to re-connect the peripheral artery to the unconnected artery. This tracking algorithm, which was inspired by the algorithm in [120], [121] is based on three steps:

first segmentation of the peripheral artery (collecting voxels with CT number for the pulmonary artery), second the segmentation of the unconnected artery caused by PE (collecting voxels with CT number for an embolus) and applying an OR operation to connect the segmentation artery and the dilated unconnected artery. These paths which connect the unconnected vessels can also be considered as a potential PEs.

Figure4-16 shows the steps of the tracking algorithm. Figure4-16a shows a branch of the peripheral artery with a blockage. This blockage created a gap between the rest of the branch and the major branch (Figure4-16b). The second step of the algorithm collects the voxels (the voxels with CT value more than -150 HU and less than 100 HU) that creates the gap between the vessels (Figure4-16c). These voxels (paths) are selected as pulmonary embolism candidates. Figure4-16d shows the final step of the tracking algorithm in which the vessel is completely extracted using an OR operator.

(a)   (b)

(c)   (d)
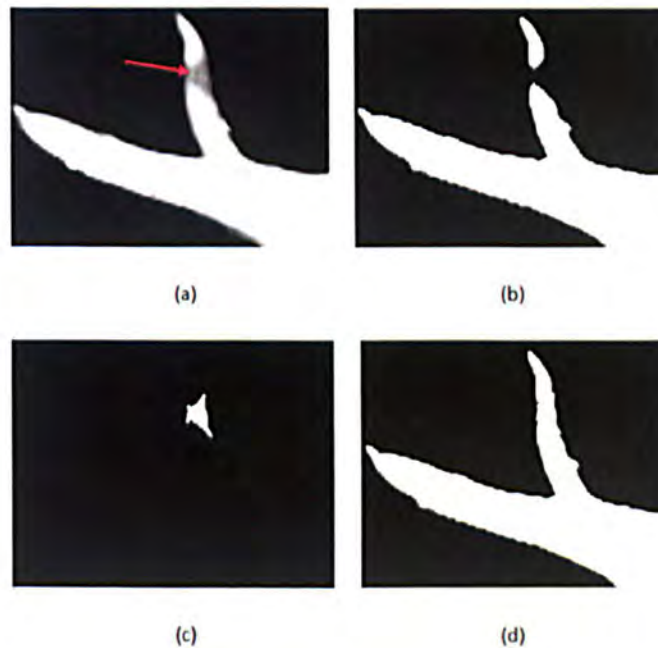
Figure 4-16 shows different steps of tracking algorithm to find potential PE. (a) A branch of pulmonary artery with a blockage. (b) The initial segmentation result showing the gap between the rest of the branch and the major artery. (c) The path between two unconnected vessels which is considered as a potential PE. (d) The completed artery extracted by OR operator.

4.3. False Positive Reduction

4.3.1. Introduction

The output of the five detectors may contain many false positives due to a variety of effects: the partial volume effect on the vessel boundary, soft tissue such as lymphoid tissue surrounding vessels, noise and motion artifacts. By analyzing the connected components of candidate voxels (26-neighbor connectivity or $3 \times 3 \times 3 - 1$ connectivity is chosen aiming to achieve a small number of false detection), the sensitivity of the detection will be improved. To reduce the partial volume effect, voxels on the segmented artery wall are removed according to the radius of the artery, noting that the thickness of the major artery is greater than in the peripheral artery. The other feature classification decision for each PE candidate is based on the size of grouped voxels [25]. Size is used to reject small dark regions associated with noise.

Soft tissue between the bronchial wall and the pulmonary artery is one of the most significant causes of a high false detection rate in most CAD systems (Figure 4-17). By reviewing the proposed system, it showed that the most common type of false positive detection (one third of all false positive detections) is the soft tissue type (see Table 5-1). In order to reduce the number of false positive detections caused by soft tissue, we propose a new feature based on location. The location of soft tissue is anatomically between the pulmonary artery and the bronchial wall. So by tracking the airway system (trachea and bronchus branch) the false positives caused by soft tissue can be reduced. We first segment automatically the airway system and then using a morphological dilation operator a mask is created to filter the image containing connected components (PE candidates) that might be attributed to soft tissue.
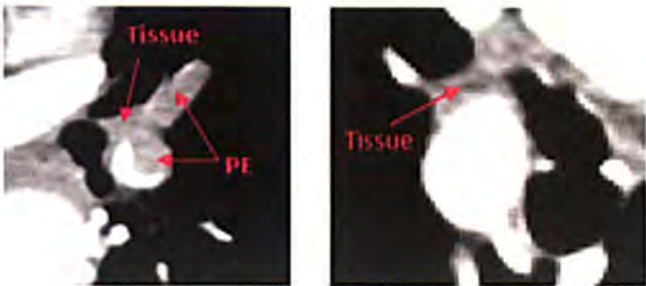
Figure 4-17 shows the soft tissue between the pulmonary artery and the bronchial wall which has the same intensity as the PE.

4.3.2. False Positive Reduction Caused by Partial Volume

In medical imagery, if a single object is present in a voxel then that object will be identified by an intensity value but if more than one object type is present in the voxel then the intensity will be the combination of the contributions of the different objects. This is called the partial volume effect (PVE) in medical imaging. In fact, partial volume effect arises when a single voxel represents more than one object types.

In pulmonary CT angiography images, the pulmonary arteries are bright, surrounded by a dark object in the lung region (lung parenchyma). Therefore, because of the point spread function (PSF) and the partial volume effect, the boundary of the pulmonary artery consists of voxels whose gray levels are equal to the gray level of an embolus. So these voxels are collected by the PE detectors as a PE candidate, which raises the false positive detection rate. Figure 4-18 shows the results of applying the first PE detector to a CT image containing a PE. It can be seen that the voxels on the boundary of the pulmonary artery are detected as PE candidates due to the partial volume effect (Figure4-18b).

(a)   (b)
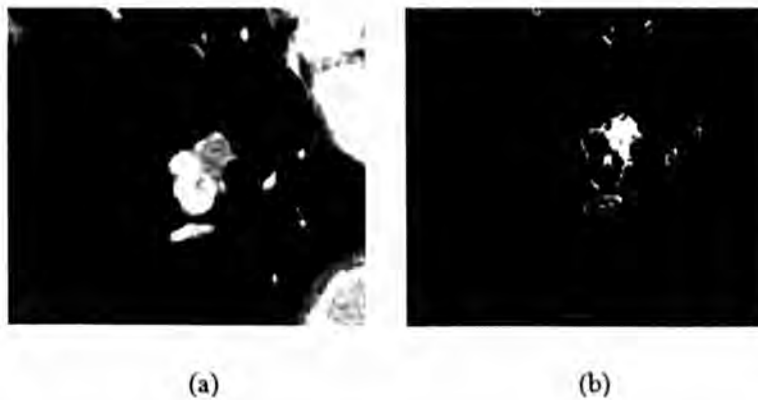
Figure 4-18 shows false positive detection caused by the partial volume effect. (a) A 2D CT image with PE. (b) The result of the first PE detector which collected the boundary voxels as a PE candidate.

In order to remove false positive detections caused by the partial volume effect, voxels on the vessel wall are removed by a morphological operator. The thickness of removal wall depends on the radius of the pulmonary artery, noting that the thickness of the major artery is greater than in the peripheral artery.

4.3.3. False Positive Reduction Based on the Volume Size of Candidate

A PE candidate is a collection or a group of connected voxels in terms of 26-connectivity. So the volume size of the candidate can be considered as a feature to reduce the false positive detection rate. The small flow void and noise create the small dark region in the search area which can be detected as a PE. Therefore, if the volume size of the PE candidate is small enough then that candidate is presumed created by noise or small flow void. Consequently, the volume size of the PE candidate can be used for reducing the false positive detection rate. However, this feature cannot distinguish between a true PE and a large false positive detection.

Another feature that uses the size of connected voxels is the effective length, which is the summation of width, height and depth of the smallest box bounding the voxel group. This feature is used to adjust the sensitivity of detection and eliminates further false positives.

4.3.4. False Positive Reduction Caused by Soft Tissue

The highest false positive rate reported in most developed CAD systems was attributed to soft tissue, such as lymphoid tissue, which surrounds the pulmonary vessel in the lung region. In the pulmonary artery segmentation process, removing the soft tissue is a challenging task due to these regions having a CT value very close to that of a pulmonary embolism. Moreover, in some cases because of other disease, abnormality develops in the lymphoid tissues and hence more false positives can be detected. So it is desirable and reasonable to consider a new feature to discriminate the false positives caused by soft tissues outside the pulmonary artery and the real PE inside the artery, a consideration that has been ignored in some proposed CAD systems.

Since soft tissue is anatomically located between the pulmonary artery and bronchial wall in the lung region (Figure4-19), by tracking the airway system (trachea and bronchus branch) the false positives caused by soft tissue can be reduced. The airway system is first segmented and then with a mathematical morphology dilation operator a mask is created to filter the image containing connected components (PE candidates).

(a)   (b)
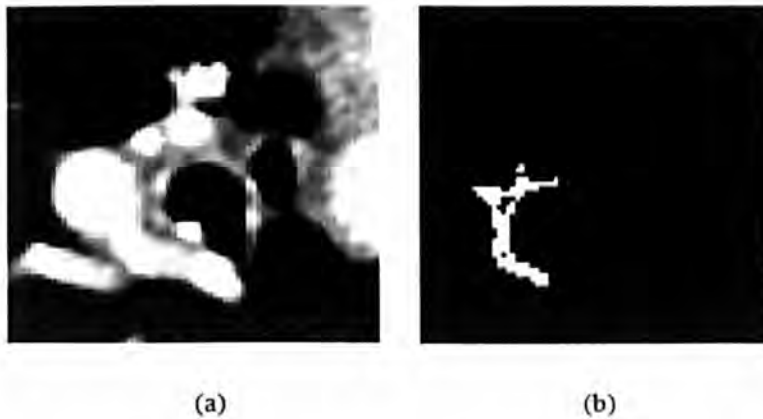
Figure 4-19 : shows false positive detection caused by soft tissue between the pulmonary artery and the major branch of airway system. (a) A 2D CT image. (b) False positive detection.

Segmentation of airway system

The segmentation of the human airway, which includes the hierarchical tubular structure that leads the air into the lungs, namely the trachea and the bronchi tree, plays an important role for many clinical applications and for physiological studies in medical imaging. Some diseases that can affect lung function are emphysema, bronchiectasis, pulmonary embolism and pneumonia. Some of these diseases can be diagnosed by analysing and measuring the thickness of the airway wall, analysing the diameters of airways and determining the rate of trapping of the bronchus. The airways structure is segmented and traced in order to reduce the false positive detection rate caused by soft tissue.

Segmentation of airways can be a challenging task. It can be segmented manually by an image analyst or by a qualified and experienced radiologist, but manual segmentation is time consuming and tedious and can take several hours to analyse [107]. A CT data set for analysing the chest disease with slice thickness less than 1 mm, contains more than 300 slices, so manual segmentation is an impractical task. Several methods have been proposed to segment the airways. Many algorithms use gray level morphological operators [101], [102] or apply the front wave propagation algorithm [103], [104].

But most of proposed algorithms use a region growing algorithm to separate the air and tissue voxels based on a HU (Hounsfield Unit) threshold [105], [106]. Since the region growing algorithm is fast and it does not need prior knowledge of the size and shape of the airway so we use an automatic region growing algorithm to extract airways in CTA

images. In this algorithm the airway is segmented using the region growing algorithm, starting from a seed point which is automatically selected from the top slices by searching for a circle connected component close to the middles of the slices.

To find seed points a 2D connected component labelling is used at the top slices after thresholding, based on the CT number for pure air (-1000 HU). Shape and size properties of extracted components (objects) are measured such as eccentricity, area and solidity. The criterion of the growing in the proposed algorithm is the similarity of voxels values in which the voxels with CT number between -900 HU and -1100 HU are collected as the airway region. Figure 4-20 shows segmentation results of several different 3D binary images of the airway system.

Figure 4-20: shows the results of airway system segmentation algorithms for 3 different datasets.

The segmented airway is then selected to create a filter for reducing the false positive detection rate. Figure 4-21 shows the result of applying this filter and removing the false positive detection caused by soft tissue between the pulmonary artery and airway system.

(a)   (b)   (c)
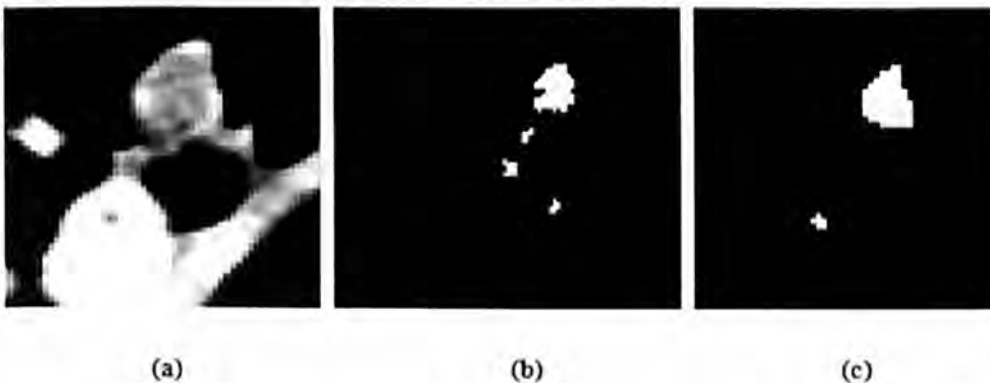
Figure 4-21: shows removal of false positive detection caused by soft tissue. (a) An original 2D CT image with real PE. (b) The result of PE detector in which the both real PE and soft tissue were detected as PE candidates. (c) The result of applying the proposed filter which removed false positive detections caused by soft tissue.

4.4. Summary

In this chapter two major components of the proposed CAD system (PE segmentation and false positive removal) have been demonstrated. Five different detectors have been developed to detect PE candidates inside the segmented pulmonary artery. These detectors respond to different properties of PEs and are based on intensity, geometrical analysis, top-hat transform, detecting the disconnected part of the artery, and based on perpendicular plane of the centreline extraction of segmented artery and ellipse fitting. The first three detectors were implemented inside the segmented artery while the other two detectors have been used to search for PEs outside the segmented artery.

Due to a variety of effects such as the partial volume effect on the vessel boundary, soft tissue such as lymphoid tissue surrounding vessels, noise and motion artefacts, the proposed system produced many false detections. To improve the sensitivity of the system, many features have been applied to reduce the false detection rate.

False detection, caused by the partial volume effect, has been reduced by removing the voxels on the vessel wall with a morphological operator. The size of the PE candidate was considered in order to remove false detection caused by noise. A new feature based on location has been developed to remove false detections caused by soft tissue. Soft tissue is located between the pulmonary artery and the airway system, so by tracking and segmenting the airway system this kind of false detection has been removed.

False detection caused by flow voids in veins is the most important cause for raising the false detection rate in most previous CAD systems. So by separating the pulmonary artery from the veins, the rate of false positive detection was considerably reduced. In fact, reducing the search area was the first factor or feature to reduce the false detection rate.

5. EXPRIMENTAL RESULTS AND DISCUSSION

In this chapter, the performance of the CAD system is evaluated and we demonstrate the effectiveness of the proposed CAD system by performing it on CTA images. As discussed in previous chapters the proposed CAD system consists of three major components which are:

- Pulmonary artery segmentation.
- Pulmonary embolism candidate detection.
- Removal of false positives.

In the pulmonary artery segmentation stage, some pre-processing has first been implemented such as lung and heart segmentation: to reduce the search area, feature map filter extraction is used to remove the connectivity between the pulmonary artery and other organs and applying the diffusion filter in order to smooth homogeneous regions and preserve the edges or locations of boundaries. Then the pulmonary artery was extracted in two stages in the lung and heart region. In pulmonary artery segmentation, all processes have been implemented automatically and all thresholds were selected optimally. In this chapter all these processes and selections of parameters are validated.

For pulmonary embolism candidate detection and removal of false positive detection steps, five different PE detectors were applied to detect PE candidates and some features have been used to remove false detections. These features are based on size, intensity, location and analysis of the shape of the lumen. In this chapter, in the second section all these features are analyzed and validated to show how the amount of information will be contained for PE classification. The sensitivities and specificities of major features are measured and the FROC (detailed on page 125) curve is plotted to evaluate the performance of features and compare whole CAD systems.

This chapter is organized as follows. First we provide a detailed description of the CT data used for validating the major component of the pulmonary artery segmentation algorithm, pulmonary embolism candidate detection and removal of false detection processes and then the major component of the CAD system will be evaluated. The validation of these processes has been carried out by a qualified radiologist from Lausanne hospital.

5.1. CT Data for Validation

The PE CAD system has been validated on a clinical dataset of 55 CTA scans with an average of 220 slices per scan (involving a total of 12100 CT slices). 20 CTA scans (containing 97 emboli) data have been used as training datasets for parameter selection and 35 more CTA scans containing 195 emboli have been used for testing. The scan data collected from Lausanne University are sampled from the top of the aortic arc to the diaphragm into on average 220 slices of 512×512 with a spatial resolution of 0.62 mm and a slice thickness of 1.25 mm. Based on the CT image acquisition protocol the voltage of the X-ray tube ranged from 120kV to 140kV and the mean current of the tube varied from 241mA to 350mA.The number and location of PE's were marked by a qualified radiologist. The radiologist marked the centre of each PE resulting in a total of 176 PE, 68% of which were located in the peripheral arterial network (Segmental and Sub-Segmental) and 32% in the main pulmonary artery (Main and Lobar). Figure 5-1 shows more details of the occurrences of pulmonary embolism in all pulmonary artery types.
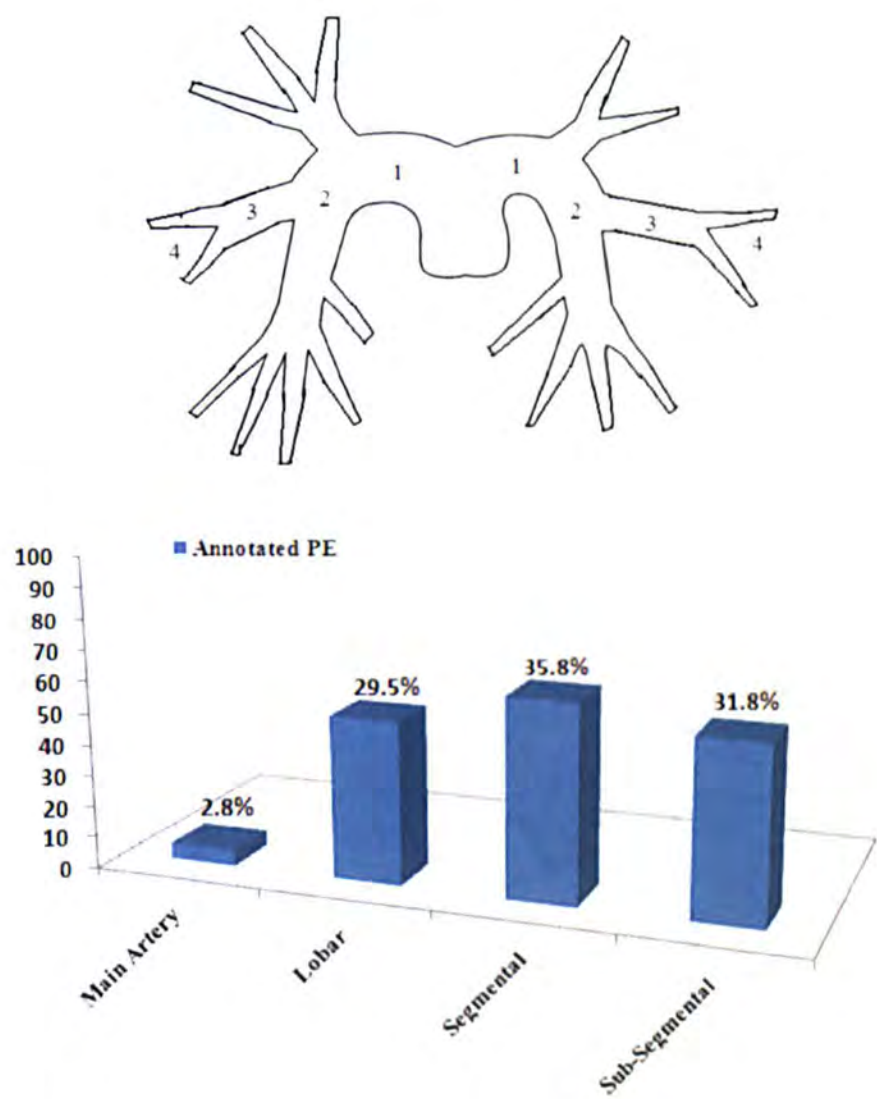
Figure 5-1 shows the occurrences of blockage in different part of the pulmonary artery (1- Main artery, 2-Lobar, 3-Segmental artery, 4-Sub-Segmental artery).

5.2. Validation of Pulmonary Artery Segmentation

As discussed in section 3.3 many steps were implemented to segment the pulmonary artery. In this section we only validate and evaluate the major components which were applied to extract the pulmonary artery. The major components are the lung and heart segmentation processes, the process of removing the connectivity by a feature map and finally the level set algorithm implemented to extract the major artery, and the peripheral arteries.

5.2.1. Lung Segmentation

The lung segmentation process was implemented as a crucial pre-processing step prior to segmentation of the pulmonary artery tree in order to reduce the search area. In the proposed CAD system, the lung region was segmented in three steps. In the first step, the initial lung region was extracted using the gray level thresholding method using an optimized threshold. Figure 5.2 shows how the optimized threshold varies in different datasets (20 different training dataset). The figure shows that minimum and maximum thresholds were in the range -748HU and -615 HU respectively (Section 3.3.2).
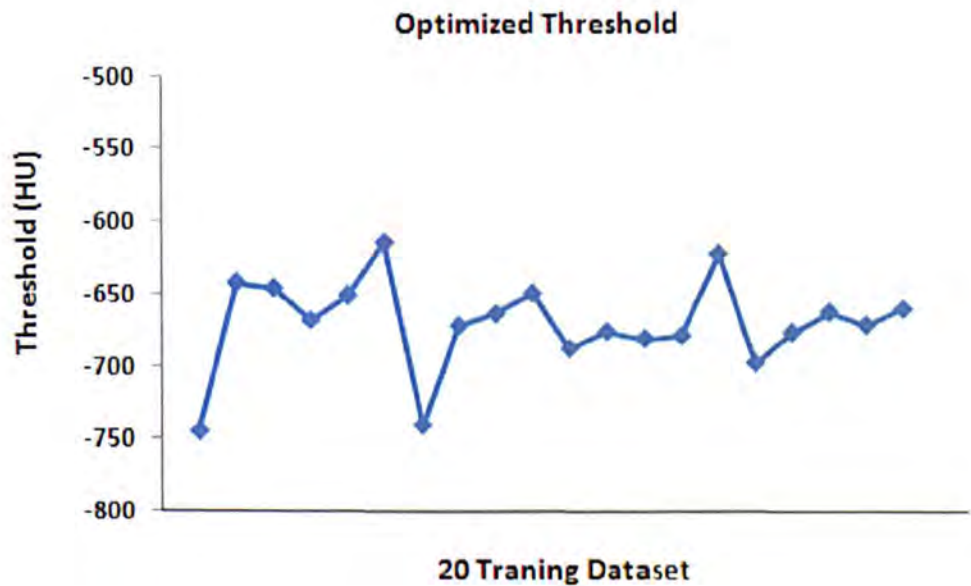
Figure 5-2 shows the graph of optimized threshold for 20 different datasets

In the second step, the left and right lungs were separated by identifying anterior and posterior junctions. Before the separation of the left and right lung, the trachea was extracted for quantitative analysis of lung tissue and facilitates left and right lung separation. The trachea was segmented using an up-to-down 3D region growing algorithm, starting from a seed point which is automatically selected in the top slices by searching for a connected circle shape close to the centre of the image (refer to section 3.3.2.1). The result of seed point selection on 20 training datasets showed that the seed points were correctly selected on all 20 datasets (100% sensitivity).

Figure 5-3 shows the 16 different slices of manually segmented and annotated results and the results of segmentation obtained by the CAD system for scan data.

As the only information available is boundaries outlined by an expert radiologist, the results of lung segmentation algorithm (computer-based) can be evaluated against the results of manually segmentation by radiologist using the method proposed in [130]. To assess the accuracy of segmentation, first the minimum distance between the computer-based border and manually computed border was calculated. For each pixel in the computer-based segmentation result, the minimum distance to the border of manual segmentation results was computed as:

$$d_k = \min_j \left\{ \sqrt{\left(x_k^C - x_j^M\right)^2 - \left(y_k^C - y_j^M\right)^2} \right\} \tag{1}$$

where, $\left(x_k^C, y_k^C\right)$ is the position of a pixel in the computer-based border and $\left(x_j^M, y_j^M\right)$ pixel position of manually detected borders [130]. Then statistics were derived in order to find whether the computer-based segmentation results agree with manually-based segmentation results. For each computer-based border the mean, RMS (Root Mean Square) and the maximum distance to the boundary of the manually-based segmentation result were calculated by the following formulas:

$$d_{Mean} = \frac{\sum_k^L d_k}{L} \tag{2}$$

$$d_{RMS} = \sqrt{\frac{\sum_k^L d_k^2}{L}} \tag{3}$$

$$d_{Max} = \max_k d_k \tag{4}$$

where L is the number of boundary points of the computer-based border. The border pixel accuracy assessed for 200 randomly selected slices and their mean, RMS and maximum distance were calculated. Figure 5-4 shows the comparison between computer-based segmentation results and manually defined results before and after smoothing. The results show that the smoothing step improves the accuracy of segmentation.
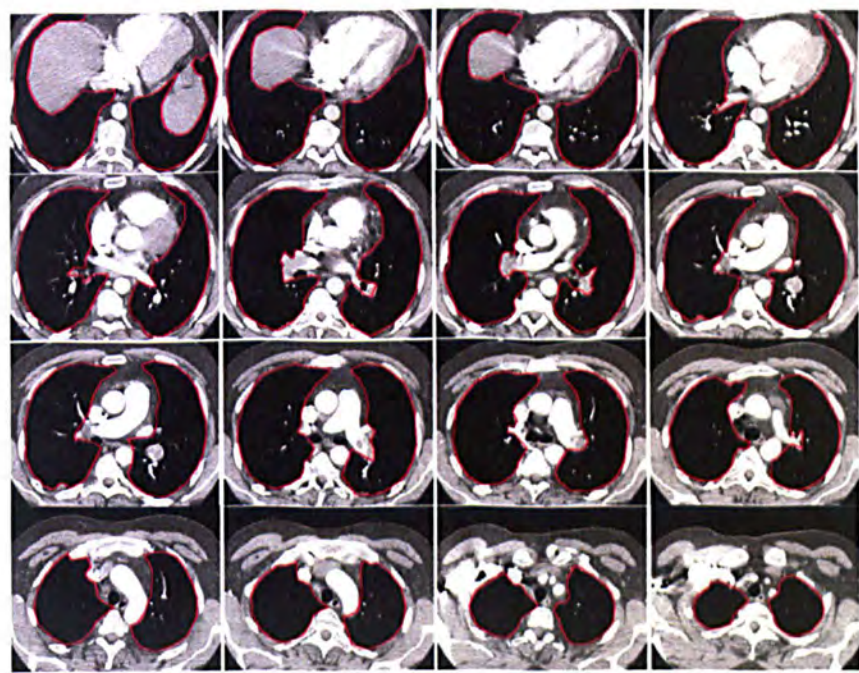
(a)

(b)

Figure 5-3: (a) Shows the lung region annotated by a radiologist and (b) shows the segmentation result by the proposed CAD system
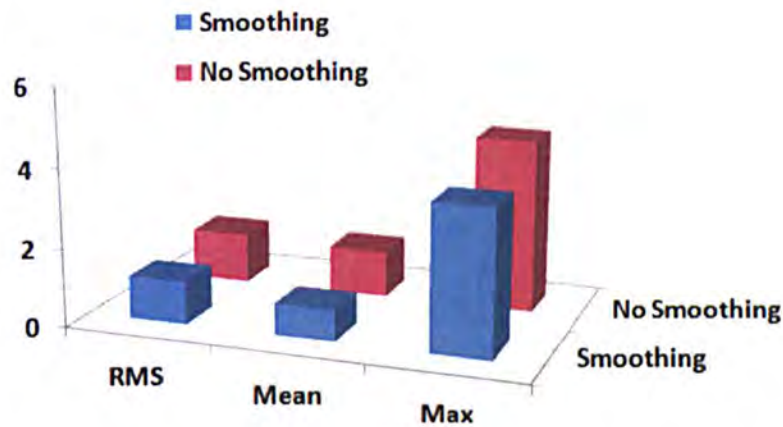
Figure 5-4 shows the accuracy of boundary position when it is assessed by using the distance measures defined from (2)-(4). Figures include mean, RMS, and maximum distance averaged over boundary points of computer-based segmentation results before and after smoothing.

5.2.2. Removal of the Connectivity (Feature Map)

For most previous pulmonary artery segmentation algorithms, the segmentation process fails due to connectivity between the pulmonary artery and the other organs. So a pre-processing was applied to remove the connectivity (feature map). This pre-processing utilized the second derivatives of the Hessian matrix of multi-scaled data. The raw data was scaled with Gaussian filters using $\sigma = [2.5, 3, 3.5]$ (described in more detail in section 3.3.4). These scales were experimentally selected based on the size of the pulmonary artery. Figure 5-5 shows the result of applying the feature map in different slices of a scan data. The right column shows the removal of connectivity (arrowed in middle column) between the pulmonary artery and other organs. The result of applying the feature map on training data set showed good results to remove the connectivity. The filter removed the connectivity in 17 out of the 20 training scan datasets and failed in 3 of the datasets (Figure5-6b). Therefore, a mathematical morphology erosion operator (a disk structure element of radius 2mm) was applied to ensure complete removal of the connectivity (Figure5-6c). Figure 5-6 (This is the same picture described in 3.3.6.1)

shows the removal of connectivity between the right pulmonary artery and superior vena cava in which the feature map filter failed to remove the connectivity but the mathematical morphology erosion operator removed it.

Applying only the morphology erosion operator to remove connectivity would need a much larger structure element (a disk structure element of radius 10 mm). The result of using a large structure element would be to reduce the size of large arteries and remove arteries with small size.
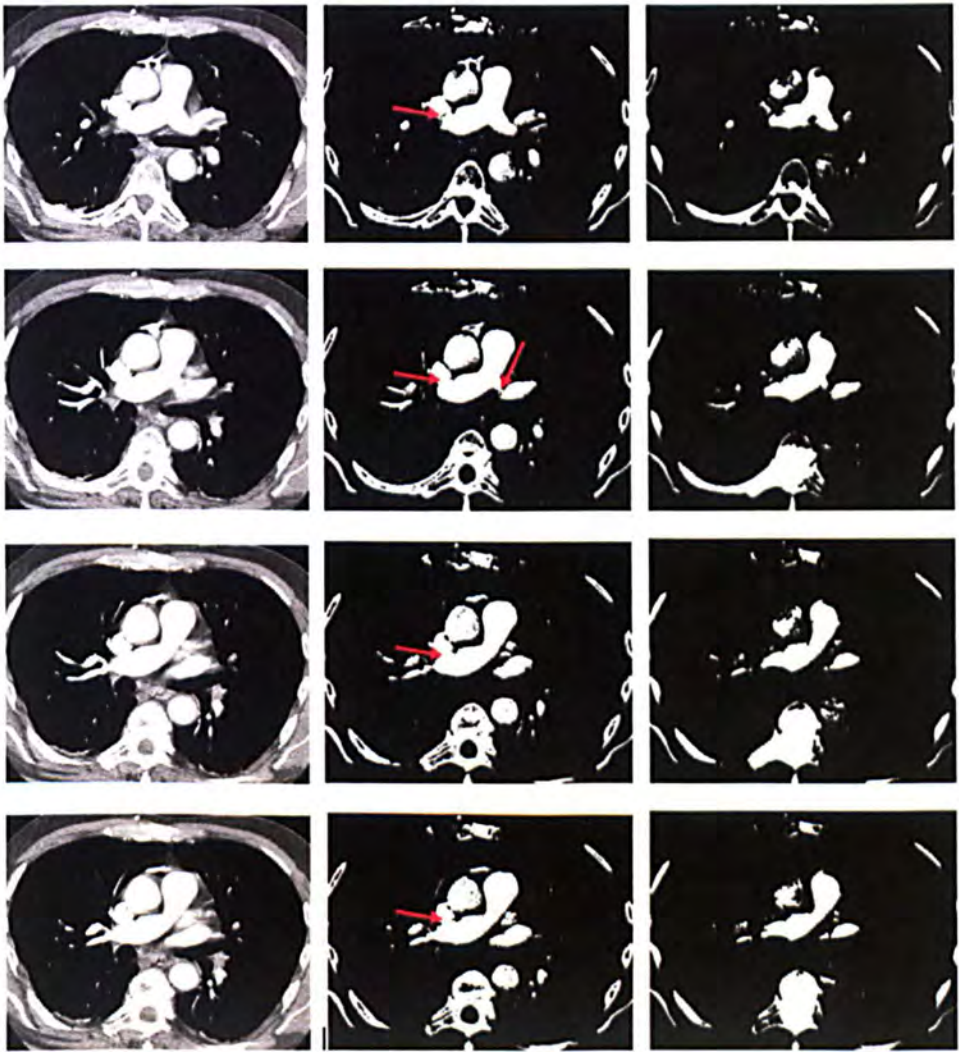
Figure 5-5 shows the results of applying feature maps on CTA images for removing the connectivity (arrowed) between the pulmonary artery and other connected organs.

(a)   (b)   (c)

Figure 5-6 shows the removal of connectivity between the right artery and superior vena cava using the proposed feature map and an erosion operator: (a) connectivity between superior vena cava and right artery (arrowed); (b) shows the failure of applying feature map to remove the connectivity; (c) completed removal of connectivity after erosion.

5.2.3. Validation of Pulmonary Artery Segmentation Algorithm

As discussed in section 3.3.6, the major artery was extracted using a new customized level set algorithm using an initial surface which was automatically selected inside the pulmonary artery. To find the initial surface, a set of seed points was automatically selected inside the pulmonary artery to initiate the region growing algorithm. This seed region was selected by searching the biggest 2D component in the heart region. The result of region growing, which groups voxels with intensity values greater than the optimized threshold, was used as the initial contour for the level set algorithm. In the initial contour selection step, two options were selected automatically: 1) an optimized threshold, 2) a set of seed points.

The optimized threshold was selected automatically by analyzing a local histogram of the heart region. In the heart region there are two main vessel groups: 1) the pulmonary artery, superior vena cava, right atrium and ventricle which have the highest CT value, and 2) the left atrium and ventricle, aorta and pulmonary vein which have a lower average CT value than the pulmonary artery. The histogram of the heart region consists of two maximum peaks which belong to two main groups in the heart region (Figure 5-7). For each dataset the minimum point between two maximum peak points was selected as an optimized threshold. Figure 5-8 shows the graph of optimized threshold selected for 20 different training datasets. The figure shows that optimized threshold for initial PA segmentation varies between two thresholds $T_1 = 92$ HU and $T_2 = 193$ HU
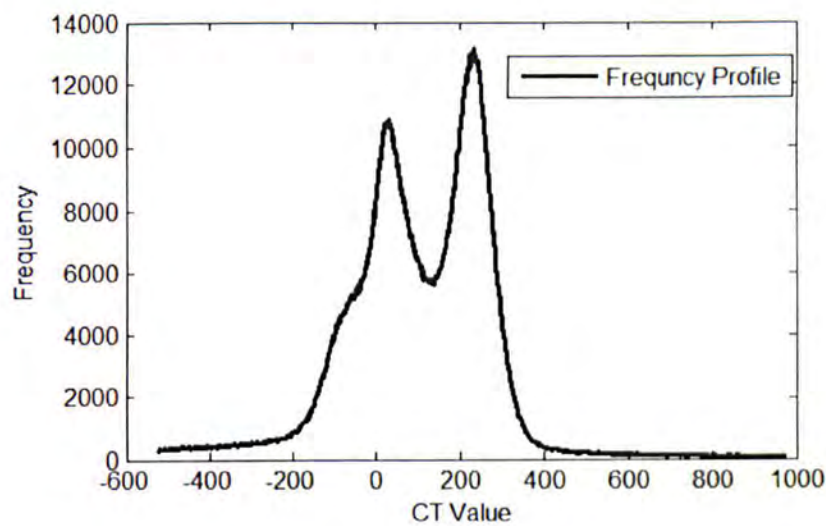
Figure 5-7 shows the intensity distribution for the heart region
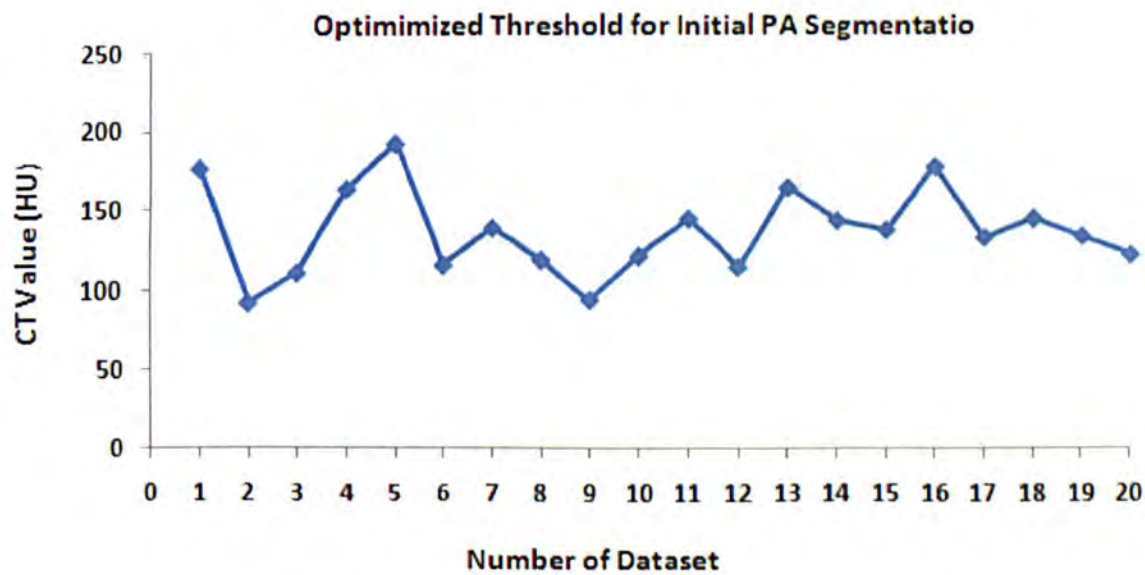
Figure 5-8 shows the graph of optimized thresholds for different datasets (20 training datasets). Figure shows that optimized threshold for initial PA segmentation can be a CT value between two thresholds $T_1 = 92$ HU and $T_2 = 193$ HU

The performance of the seed region selection algorithm was estimated on the training dataset. The seed points were selected by searching for the largest 2D connected

component in the heart region after applying the 2D connected component labelling algorithm. In 3 out of 20 training datasets, the seed point selection algorithm failed and selected the aortic arc as a seed point (Figure 5-9) due to the search area that was selected being the whole slice of the dataset. Since the pulmonary artery is anatomically located between the aortic arc and right ventricle (in the dataset sampled from the top of aortic arc to the diaphragm) so the search area was restricted. The search area was selected as the middle slice of each data scan plus 30 more slices either side of the middle slice. By restricting the search area, the results of seed region selection in more datasets (35 testing datasets) showed that all seed regions were correctly selected in the heart region.
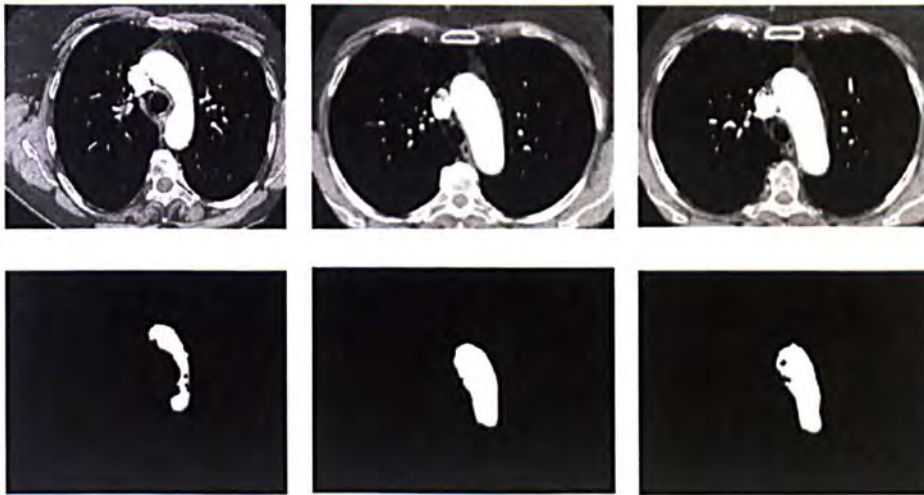
Figure 5-9 shows the failure of seed point selection to segment the initial pulmonary artery in the heart region in which the aortic arc was selected as a seed region. (a) First row shows the original different dataset. (b) Second row shows the binary masks of aortic arc.

After seed point selection, a region growing algorithm starting from the selected seed region was applied to extract the initial segmentation for the major pulmonary artery in the heart region. Figure 5-10 shows that the initial segmentation result cannot be considered as a perfect result for the pulmonary artery segmentation task. So the result of the region growing algorithm was taken as the initial contour for a 3D level set algorithm to push the boundary of the initial contour toward the boundary of the pulmonary artery.
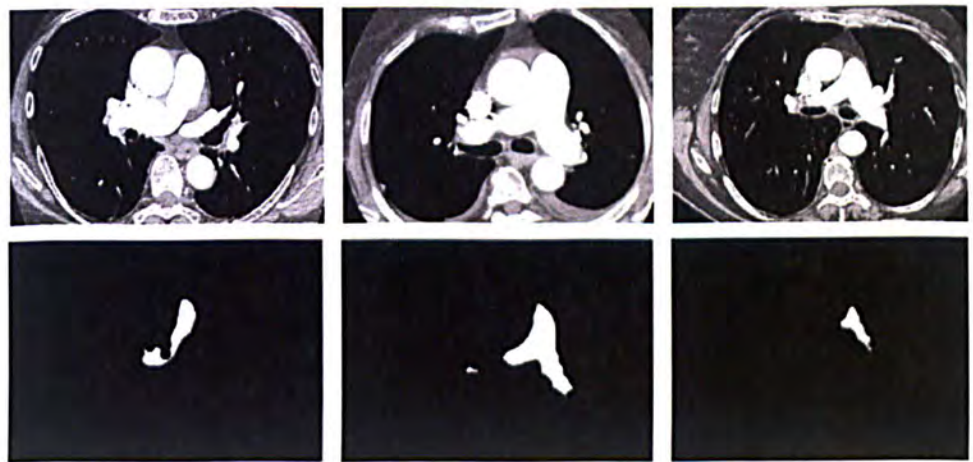
Figure 5-10: shows the initial segmentation results for different original datasets. First row shows 3 different slices of 3 different original datasets while the second row represents the initial segmentation of the pulmonary artery. The results cannot be considered as perfect results so the initial results are enhanced by a 3D level set algorithm.

Figure 5-11 shows the 2D result of using the 3D level set algorithm for segmenting the pulmonary artery. In spite of the fact that the pulmonary artery is connected with other organs such as the superior vena cava, the segmentation results (second row) show the separation of the pulmonary artery from the adjacent connected object (first row) .
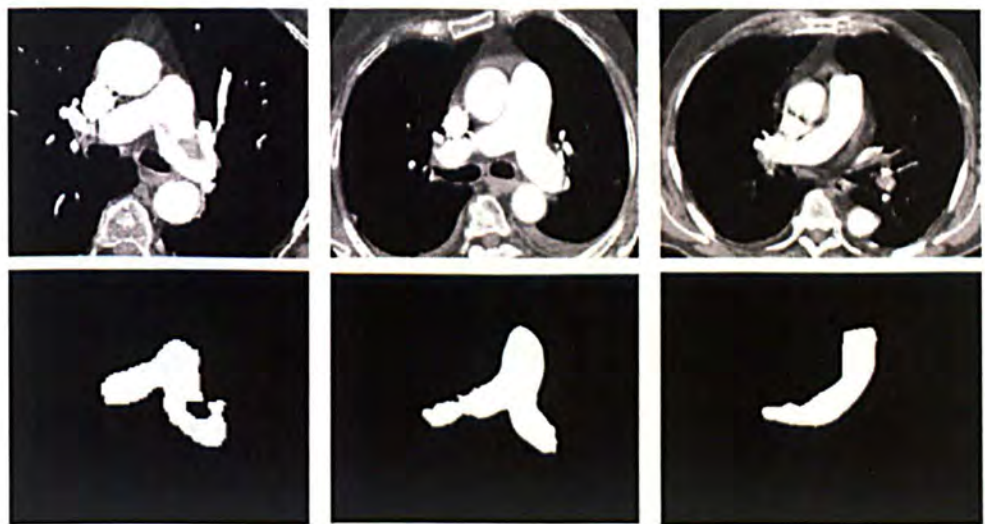
Figure 5-11: shows the segmentation results of the pulmonary artery in heart region. First rows shows the 3 different slices of 3 different scan data in which the pulmonary artery is connected to superior vena cava while the second row shows the results for the major artery using a 3D level set algorithm without leaking into the adjacent connected object.

In segmenting the peripheral artery, the vascular shape was first enhanced using a Frangi filter and then a region growing algorithm collected voxels with intensities greater than - 150 HU in the lung region (Section 3.3.7).

The accuracy of pulmonary artery segmentation results was assessed by comparing the result obtained from the proposed CAD system with the segmentation results that were manually obtained by a qualified radiologist. To assess the accuracy of segmentation, 220 different slices of the 20 training datasets were randomly selected and the boundaries of arteries were annotated by the radiologist. Figure 5-12 shows samples of different slices of the training data which have been annotated by the radiologist (Figure 5-12a) and segmented by the proposed CAD system using the level set algorithm (Figure 5-12b) in the heart region. The minimum distance between the computer-based border and the manually computed (annotated) border was calculated from (1). The border position accuracy was evaluated by calculating the mean, RMS and maximum distance between the computer-based borders and manually defined borders from (2)-(4).
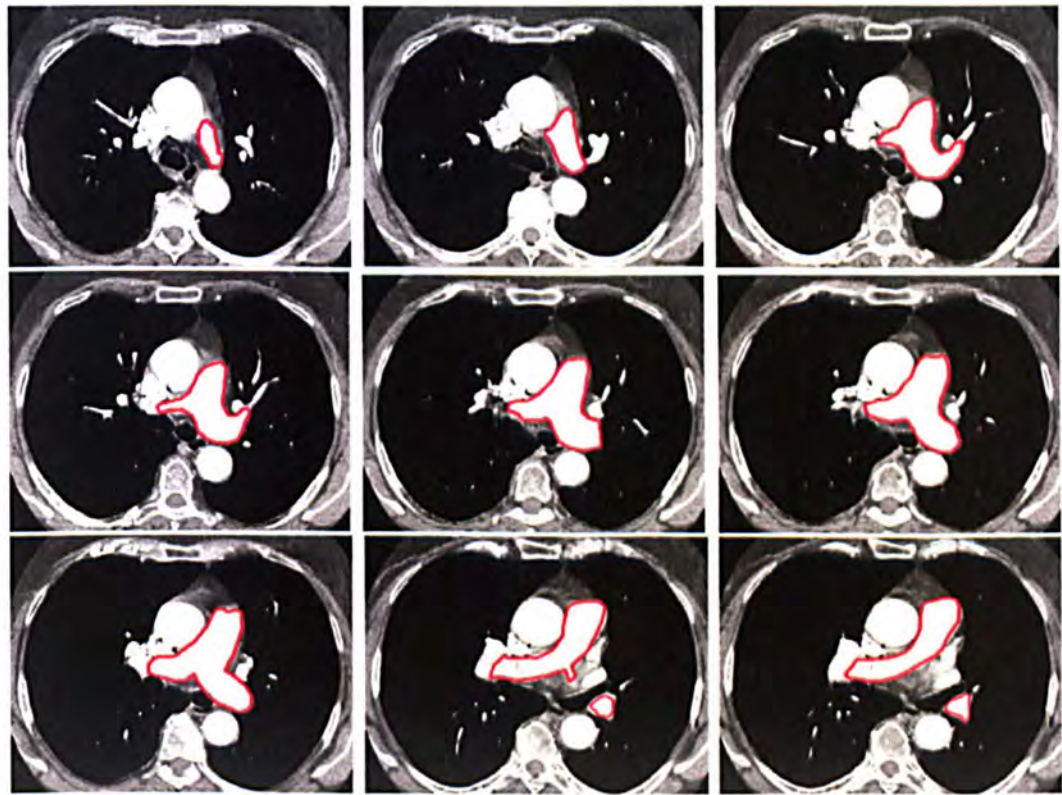
(a)
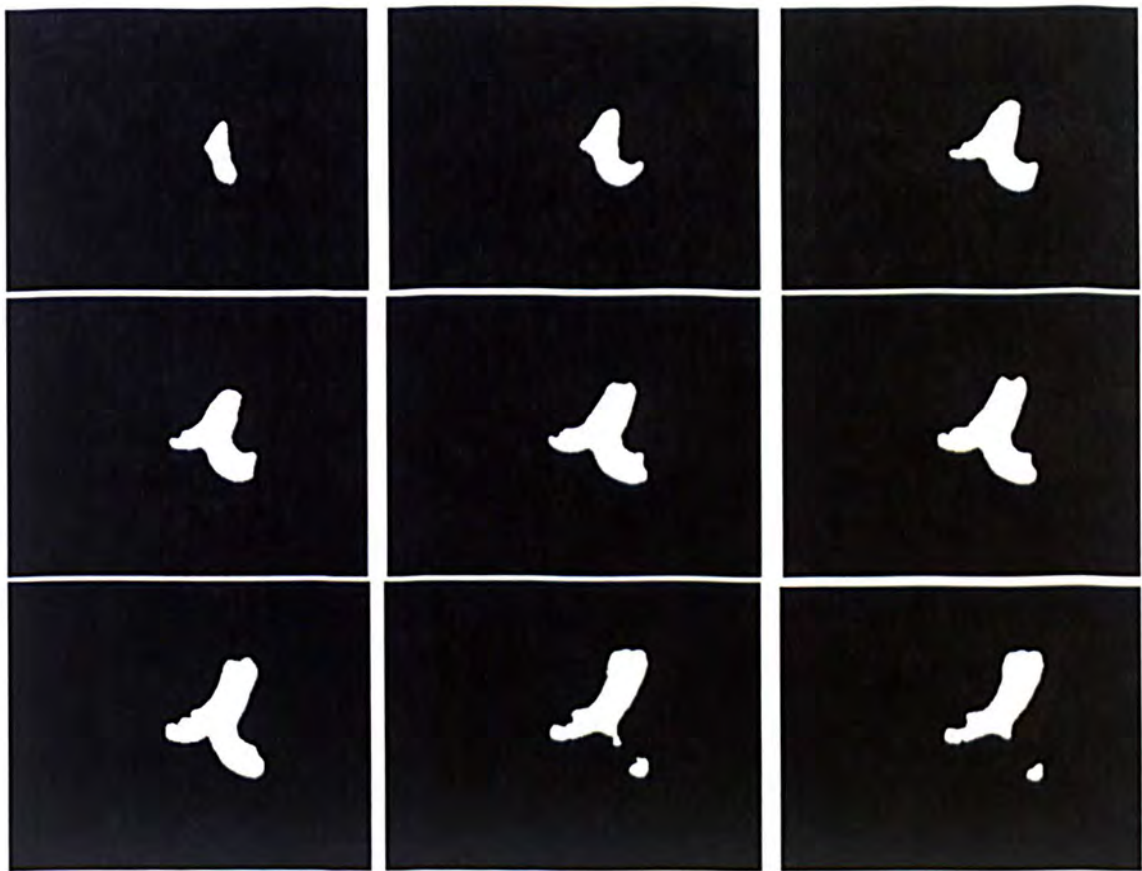
(b)

Figure 5-12: (a) Shows the pulmonary artery region annotated by a radiologist and (b) shows the PA segmentation result by the proposed CAD system

Figure5-13 shows the comparison between computer-based segmentation results and manually defined results. The figure shows the mean of RMS, mean and maximum distance for 220 different slices. The greatest difference between the computer-based border and the manually traced based borders is related to slices in which the CAD system cannot completely extract the pulmonary artery due to the existence of a PE close to the pulmonary artery wall (Figure5-14).
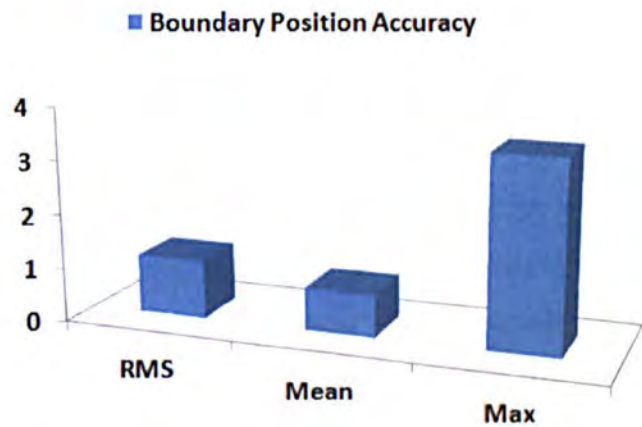
Figure 5-13 shows the boundary position accuracy between computer-based borders and manually-based borders.

(a)

(b)

(c)

(d)
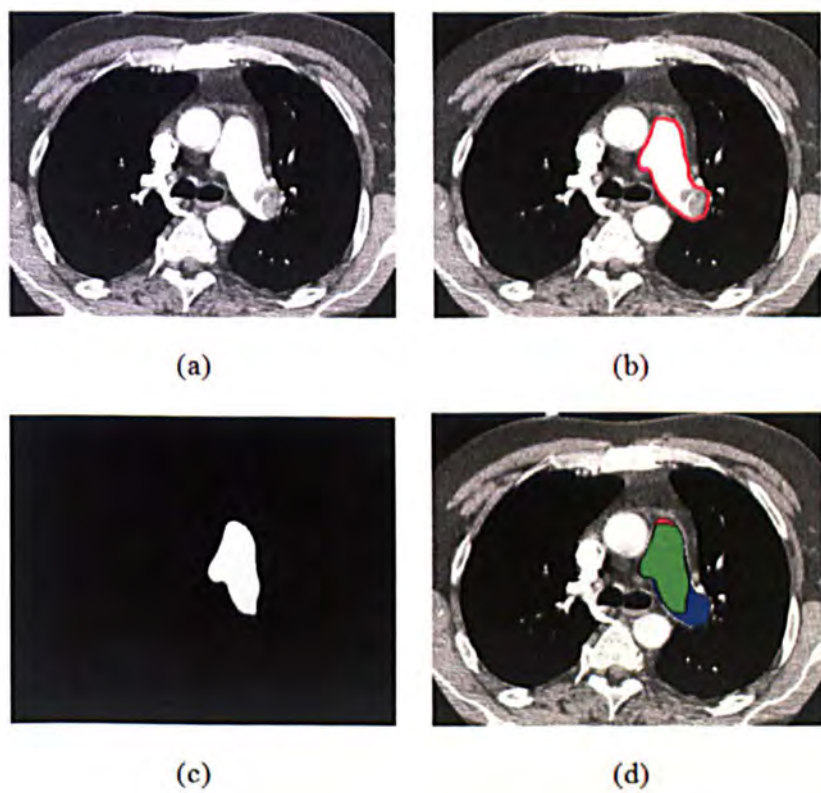
Figure 5-14 shows a case in which maximum distance may happen between the computer-based borders and manually defined borders due to the existence of a PE close to vessel wall. (a) Original CTA image. (b) Annotated or Manual segmentation result by a radiologist. (c) Computer based segmentation result. (d) The green area is intersection between computer based segmentation result and manually based result (True positive), red area false positive segmentation result and the blue area is false negative area (area missed by CAD system).

To validate the artery segmentation result, the computer based segmentation results are also evaluated using quantitative comparison based on two different criteria: 1) Spatial overlap or Dice coefficient (η) and 2) Jaccard similarity index (J) [132], [137].

In order to calculate these two criteria: first the true positive area (TP), which is a group of common pixels in the computer based and manually based segmentation results (Figure 5-15 green area), the false negative (FN) area, which is an area that was missed by the CAD system (Figure 5-15 blue area), and the false positive (FP) area (Figure 5-15 red area) are computed and then the spatial overlap (η) and the Jaccard similarity index(J) are calculated as follows:

$$\eta = \frac{2*[L_M \cap C]}{L_M + L_C} \qquad\qquad 0 \leq \eta \leq 1 \qquad\qquad (5)$$

$$J = \frac{|M \cap C|}{|M \cup C|} \qquad\qquad 0 \leq J \leq 1 \qquad\qquad (6)$$

where, $L_M \neq 0$ and $L_C \neq 0$ are the numbers of pixels for manually based and computer based segmented result, M and C are two sets of manual and computer based segmentation results respectively and $|X|$ denotes the area of X. In the case of $\eta, J = 0$, there is no overlap between manual and computer based segmentation results and there is a complete overlap between two sets when $\eta, J = 1$.
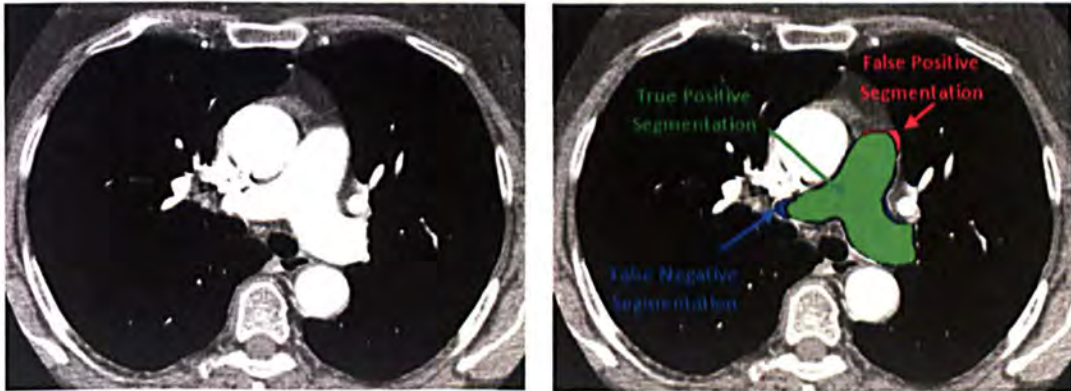
Figure 5-15 shows an original image (left) and the right image shows the comparison between computer based pulmonary artery segmentation and manually based segmentation (expert segmentation). Green colour pixels represent the spatial intersection. Red colour represents false positives (FPs) and blue colour shows false negatives (FNs).

The spatial overlap and the Jaccard similarity index of 220 different slices of 20 training scan data were calculated. Table 5-1 shows the comparison between the spatial overlap and the Jaccard similarity index. The results show that there is no significant difference between the spatial overlap and the Jaccard similarity index for 20 training scan data. The mean and standard deviation of the spatial overlap between the manual based and computer based segmentation were found to be 0.9861 and 0.0041 respectively and the mean and standard deviation of Jaccard similarity index were found as 0.9728 and 0.0082 respectively. As Zijdenbos et al [137] showed a good overlap occurs when $\eta > 0.700$ so the mean of spatial overlap of 0.9861 in the work of the thesis shows an excellent overlap between manual and computer based segmentation result.

Table 5-1 shows the comparison between the spatial overlap and the Jaccard similarity index (**M-B:** Manual based, **C-B:** computer based, **FN:** False negative, **FP:** False positive, **TP:** True positive, η: Spatial overlap, **J:** Jaccard similarity index)

Data	M-B	C-B	FN	FP	TP	η	J
1	45734	44938	859	63	44875	0.9898	0.9798
2	36951	35857	1146	52	35805	0.9835	0.9676
3	39037	38373	696	32	38341	0.9905	0.9813
4	27893	27349	580	36	27313	0.9888	0.9779
5	43028	42042	1033	47	41995	0.9873	0.9749
6	25312	25983	36	707	25276	0.9855	0.9714
7	33972	33571	420	19	33552	0.9935	0.9870
8	40634	39041	1667	74	38967	0.9781	0.9572
9	29453	28916	573	36	28880	0.9895	0.9793
10	41937	40725	1276	64	40661	0.9837	0.9680
11	32663	33498	102	937	32561	0.9842	0.9690
12	43924	42180	1827	83	42097	0.9778	0.9565
13	50295	49022	1387	114	48908	0.9848	0.9702
14	36096	35399	754	57	35342	0.9886	0.9775
15	29936	29241	762	67	29174	0.9859	0.9723
16	43193	41828	1470	105	41723	0.9814	0.9636
17	28391	27544	896	49	27495	0.9831	0.9667
18	41937	41219	831	113	41106	0.9886	0.9775
19	46028	45052	1038	62	44990	0.9879	0.9761
20	28868	29207	84	423	28784	0.9912	0.9826

As discussed in the literature review section (Chapter 3), most previous attempts to segment the pulmonary artery can be categorized into region growing methods, enhancing the vessels with Hessian matrix and fuzzy connectivity methods. In all previous methods, the seed point for the segmentation algorithm was manually selected while the proposed method automatically finds the seed points and allows the algorithm to be included in a fully automatic pulmonary embolism detection system.

Moreover, in most previous proposed methods, the segmentation algorithm sometimes fails as a consequence of merging the pulmonary artery with adjacent tissue caused by acquisition resolution, partial volume effect and noise [133]. But the proposed algorithm applies the feature map filter to remove the connectivity between the pulmonary artery and any other tissues. So, taking into account all the above issues, the proposed algorithm shows encouraging results by tackling all these issues.

5.3. Validation of Pulmonary Embolism Detection Algorithm

In this section we provide a validation and evaluation of major components which have been applied to detect pulmonary embolism candidates and to reduce the false detection rate. In this section by using the concept of sensitivity and specificity and also by plotting the FROC curve, we evaluate the performance of each feature applied to reduce the false detection rate.

5.3.1. Sensitivity and Specificity Measurement

Sensitivity and specificity are two main measures for identifying those people who have or do not have a specific disease in a screening test for a disease. The sensitivity is a measure of the ability for identifying those who have the disease while the specificity is an ability to correctly identify people who do not have a specific disease.

In other words, the sensitivity is defined as the probability that a person has the disease when she/he does actually have it and the specificity is defined as the probability that a person does not have the disease when she/he is disease free. If we display the test results and the true status of the person who is being tested as the following table:

Test Results(T)

True Status (S)		Positive(+)	Negative(-)
	Disease(+)	TP	FN
	No Disease (-)	FP	TN

then, the sensitivity and specificity are defined as follows:

$$Sensitivity = P(T^+| S^+) = \frac{TP}{TP+FN} \qquad (7)$$

$$Specificity = P(T^-| S^-) = \frac{TN}{TN+FP} \qquad (8)$$

where, FP (false positive) occurs when the result of test for a person is positive while the person is free of disease and the rate of false positive is defined as:

$$P(S^-| T^+) = \frac{FP}{FP+TN} \qquad (9)$$

FN (false negative) occurs when the result of a test for a person is negative while the person actually has the disease and the rate of false negative is defined as follows:

$$P(S^+| T^-) = \frac{FN}{FN+TP} \qquad (10)$$

5.3.2. ROC (FROC) Curve and AUC

A ROC (Receiver Operator Characteristic) curve is a graphical plot of true positive rate (sensitivity) versus the false positive rate (1-specificity). A ROC curve starts from the bottom-left to the top-right corner and each point of the graph shows a trade off between false positives and false negatives.

The accuracy of a test is measured by computing the area under the ROC curve. The area under the ROC curve (AUC) is a fraction of the unit square and it's value always is

between 0 and 1. A AUC of 1 shows a perfect result while an area of 0.5 indicates a worthless result.

Since in all CAD systems, the set of true negatives (TN) are not well defined so the ROC curve and AUC are not useful to compare whole CAD systems [122]. Since there are a lot of points which can be considered as potential true negatives in 3D data sets, the CAD system generates a large number of true negatives. Therefore, in the proposed CAD system, instead of the ROC curve, the FROC (Free-response Receiver Operator Characteristic) curve is used. In the FROC curve, the normalized specificity is replaced by the non-normalized number of false positives per dataset.

In the proposed CAD system in order to avoid having just one threshold for the first three PE detectors which are based on CT value, the FROC curve for each PE detector is analyzed. Moreover, the FROC curve is applied to evaluate the performance of each feature (which is applied to detect true PEs and remove false detections), and to compare the performance of the proposed system with previously developed CAD systems.

5.3.3. Threshold Selection for PE detectors

As described in chapter four, five PE detectors were applied to detect PE candidates within the pulmonary artery. The detectors based on CT values need an optimal threshold to control the trade-off between increased the sensitivity of detection whilst minimizing the number of false detections. The FROC curve of three detectors based on (1) CT attenuation value, (2) eigenvalues of the Hessian matrix and (3) top-hat transform were analyzed to find an optimized threshold. Figure5-16 shows the FROC curve of these three PE detectors. Optimal thresholds were determined by analyzing the FROC curve for a dataset of 20 positive samples of PE (training data). Costs and losses have been used to find the optimal thresholds [34]. The dashed-dot lines represent the lines of equal cost. The slope of dashed-dot lines were selected so that 15 false detections per dataset cost 10% of missed PEs. Based on these costs the best thresholds for the three detectors (CT value, eigenvalues and top-hat transform) were fixed between -25 HU and 75 HU, greater than 5, and between 150 HU and 250 HU, respectively. These thresholds were then used for the remaining tests to evaluate the effectiveness of PE detection.

(a)

(b)

(c)
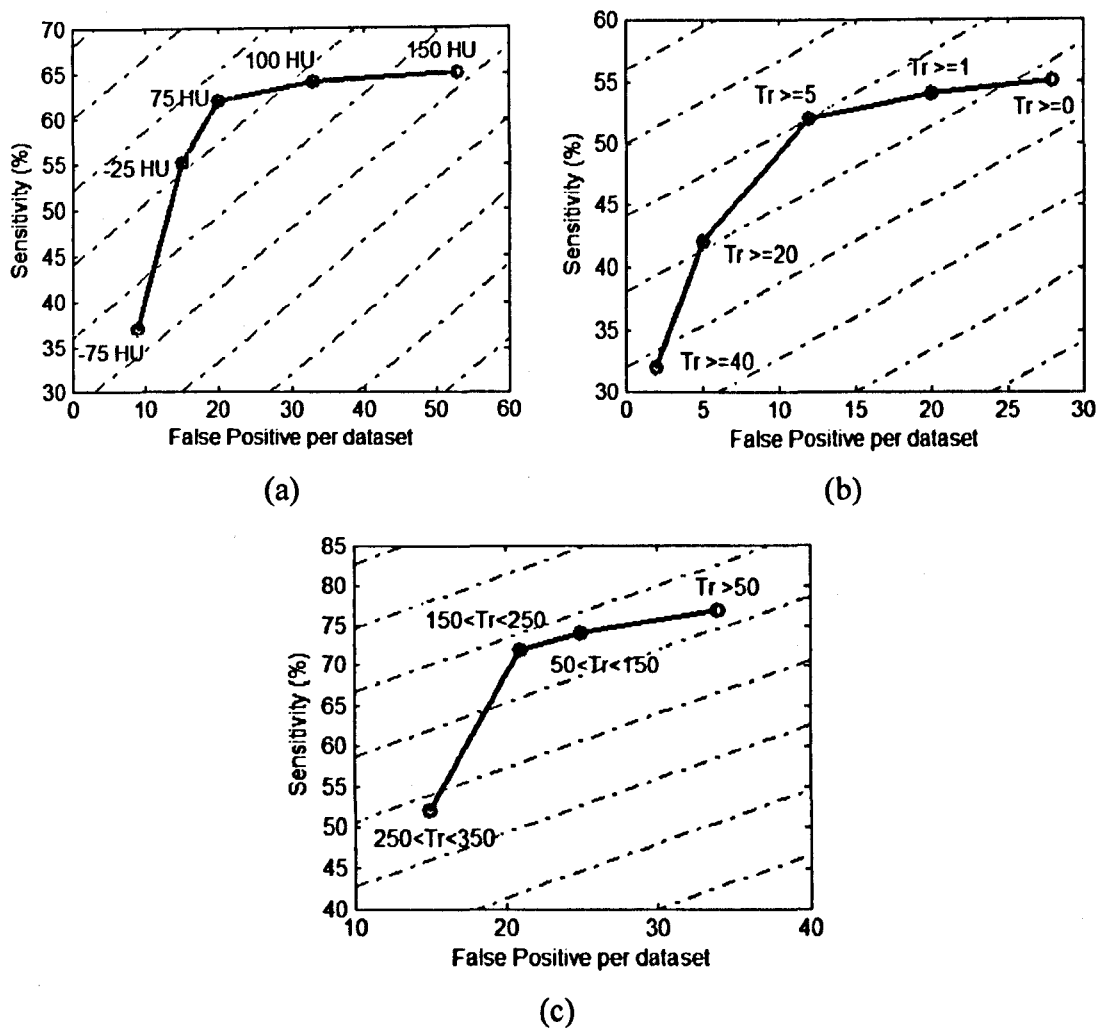
Figure 5-16: (a)-(c) show FROC curves for the thresholds of three intensity-based PE detectors. (a) CT value: The optimized threshold was selected between two thresholds -25 HU and 75 HU. (b) Eigenvalues: The best threshold for PE detector based on eigenvalue was chosen greater than 5. (c) top-hat: The optimized threshold for PE detector based on top-hat transform was selected between two threshold values 150 and 250.

5.3.4. False Positive Reduction

The five PE detectors that respond to different properties of a pulmonary embolism collected different regions. By choosing the union of these regions the sensitivity of detection was improved. These regions were analyzed by connected-component analysis and each group of connected voxels was considered as a PE candidate. However, the numbers of false positives detected caused by the partial volume effect on the pulmonary

artery wall, noise in large arteries, soft tissue and motion artefact were too large. Table 5-2 shows the main causes of false detection in the 20 training data scans.

Table 5-2 shows the percentage of major causes of false detection in 20 training data scans.

Cause of False Detection	FP (%)
Soft Tissue (Bronchial Wall)	37%
Partial Volume Effect	21 %
Noise in Large Pulmonary Artery	19%
Motion Artifact	12%
Lung Disease	7 %
Uncertain	4%

Several features were computed for PE candidates' to make distinctions between the real PE and the look-likes. In most previous CAD systems false detections caused by flow voids in veins are the most important cause for raising the false detection rate [34]. So by separating the pulmonary artery from the veins, the rate of false positive detection was considerably reduced. In fact, reducing the search area was the first factor or feature to reduce the false detection rate.

A PE candidate is a collection of voxels, so the next feature that can be used for rejecting the false detection is the size of PE candidates. The small dark region can be identified as a PE everywhere due to noise. Therefore, the size of voxel groups was used to remove the small voxel groups. Figure 5-17 shows the FROC curve based on the sizes (volumes) of candidates. The figure shows that many false positive detections can be excluded (PE candidates smaller than 42 voxels) without a considerable reduction in the number of true positive detections. The figure shows the larger the threshold for volume size, the lower the sensitivity and false detection rate.
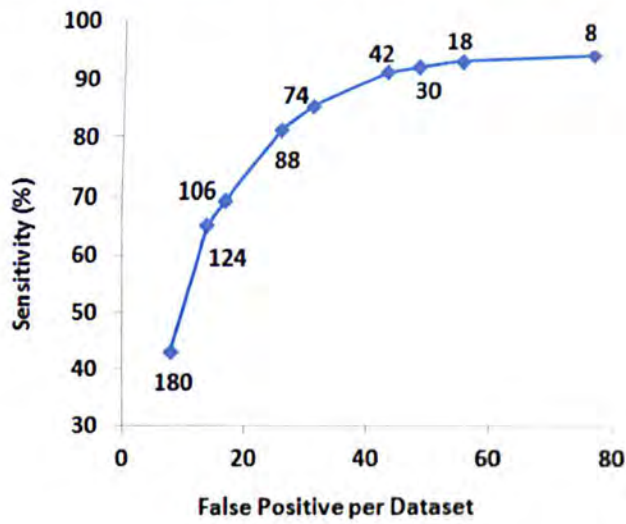
Figure 5-17 shows the FROC curve based on size (volume size) of candidates. PE candidates smaller than the thresholds (required size) are removed and the sensitivity and false positive detections per dataset are shown for different thresholds of size of voxels.

By reviewing the proposed CAD system, we found out the most common cause of false positive detections in the 20 training datasets (one third of all false positive detections) is related to soft tissue between the pulmonary artery and the airway system (Table 5-1). A lymph false positive (soft tissue) connects to the airway system and is anatomically located outside the pulmonary artery and a true pulmonary embolism is inside the pulmonary artery. So a novel approach has been developed with the aim of reducing this false positive type.

In the proposed method, first the airway system was automatically extracted and a filter was created by a mathematical morphology dilation operator on the segmented airway system. Figure 5-18 shows the performance of the proposed method for reducing false detection caused by soft tissue. The sensitivity of detection was adjusted by changing the volume size of candidates and counting the false positives per dataset.
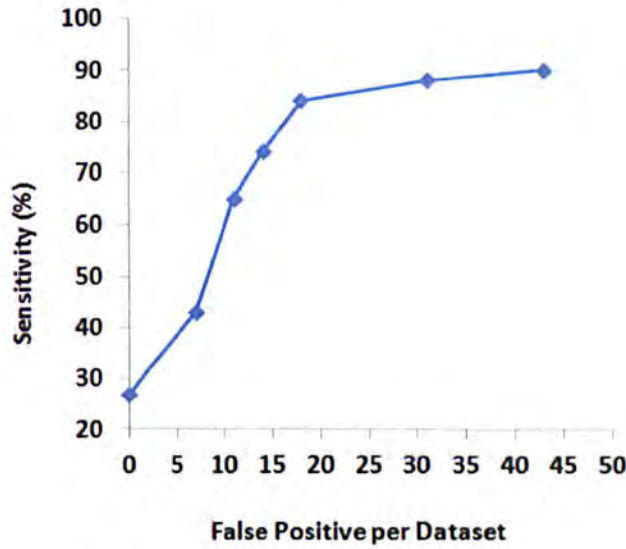
Figure 5-18 shows the FROC curve of detection result after removing false detection by the proposed filter. The sensitivity of detection was adjusted using different thresholds for candidate volume.

5.3.5. Evaluation the Proposed CAD System

The proposed CAD system for automatic detection of pulmonary embolisms in CT images is based on three stages. The first step is pulmonary artery segmentation for reduction of the search area. The second is pulmonary embolism candidate detection and the third step is the removal of false detections. For tuning the detection system and optimizing the whole CAD system, the system was trained on 20 datasets and the estimated performance of the system showed that can be very good. Table 5-3 shows the result of the CAD system on 20 training data scans. The number of detected PEs is 94 and the number of missed PEs is 3. The sensitivity of detection is:

$$\text{Sensitivity} = P(T^+|S^+) = \frac{TP}{TP + FN} \times 100 = \frac{94}{94 + 3} \times 100 \approx 97\%$$

with an average number of false detections of 3.2 per patient.

Table 5-3 shows the result of system on 20 training data scans (for different parts of the pulmonary artery see the diagram on page 109)

No. Dataset	Number of Detected PE				FP before and after Filtering	
	Main	Lobar	Seg	Sub-Seg	Before	After
T1	1	0	2	2	36	4
T2	2	1	1	0	27	2
T3	0	2	0	2	27	2
T4	1	0	4	2	37	2
T5	0	1	1	3	28	3
T6	0	2	3	0	29	2
T7	0	0	2	3	38	3
T8	0	1	0	0	25	1
T9	0	0	4	2	47	8
T10	1	1	2	2	38	4
T11	0	2	0	3	52	6
T12	0	0	2	0	33	2
T13	0	2	3	2	38	3
T14	0	1	2	3	41	4
T15	0	1	2	3	32	3
T16	0	0	2	2	27	0
T17	0	0	1	2	38	3
T18	0	2	0	0	42	3
T19	0	2	3	2	31	6
T20	1	0	2	1	42	3
Total PE and FP Detection	6	18	36	34	708 35 Per Data	64 3.2 Per Data

The CAD system was evaluated on 35 more datasets which were not used in any previous steps. The number of PEs in the evaluation dataset was initially annotated by a radiologist and 176 PEs were annotated in the 35 patient dataset. Table 5-4 shows the result of the system on 35 testing data and Table 5-5 shows the number of PEs annotated by a radiologist, the number of PEs detected by PE detectors in different artery type and the total number of false detections. The discontinuous PE was considered as separated emboli.

Table 5-4 shows the results of system on 35 more testing datasets

No Dataset	Number of PE detected in different PA types				Number of FP before and after filtering	
	Main Artery	Lobar	Seg	Sub-Seg	Before	After
1	0	2	1	0	41	3
2	0	0	2	2	38	2
3	0	1	0	2	27	3
4	1	0	2	3+2	42	4
5	0	2	1	2	43	4
6	0	2	3	0	29	1
7	0	0	2	0+1	35	2
8	0	1	0	0+2	46	5
9	0	3	0	3	49	11
10	1	0	2+1	2	32	4
11	0	2	0	3+2	52	10
12	0	0	3	0	43	2
13	0	4	3	3	48	6
14	0	0	2	1	45	4
15	0	0	2	2	39	3
16	0	2	0	0	24	0
17	0	0	1	2+1	38	3
18	0	2	0+1	0	42	5
19	0	1	5	3	47	3
20	1	0	2	0	44	7
21	0	0	0	1	29	4
22	1	3	3	2	54	4
23	0	4	2	1+2	52	3
24	0	0	3	2	45	4
25	0	3	0	0 +2	46	5
26	0	0	2	0	27	4
27	0	0	2+1	3	43	6
28	0	3	0	2+2	54	7
29	0	3	1	0	41	2
30	0	0	3	5	39	4
31	0	3	4	0	48	3
32	0	2	1	2+1	47	4
33	0	1	3	3	52	4
34	0	0	0	2	39	1
35	0	4	4	2+1	50	6

The proposed CAD system failed to detect 12 PE (6%) while the system could detect 19 (11%) extra PEs which were not annotated by the radiologist. In the second inspection

step, all the extra detected PEs have been confirmed by the radiologist. The detection results show that the system worked for sub-segmental (95%) and segmental (94%) arteries better than lobar (92%) and major arteries (80%). Figure 5-19 shows the distribution of annotated and detected PEs in different artery types.

Table 5-5 shows the number annotated and detected PEs by for the 35 evaluation datasets before false positive reduction

Pulmonary Artery	Annotated PE	Detected PE	
Main Artery	5	4	(80%)
Lobar	52	48	(92%)
Segmental	63	59	(94%)
Sub- Segmental	56	53	(95%)
Total Annotated PE	176	164	(93%)
More detected PE by CAD	19 (+11%)	19	
Total Positive PE	195	183	(94%)
False Positive Detection	-	1470 (42 per Dataset	
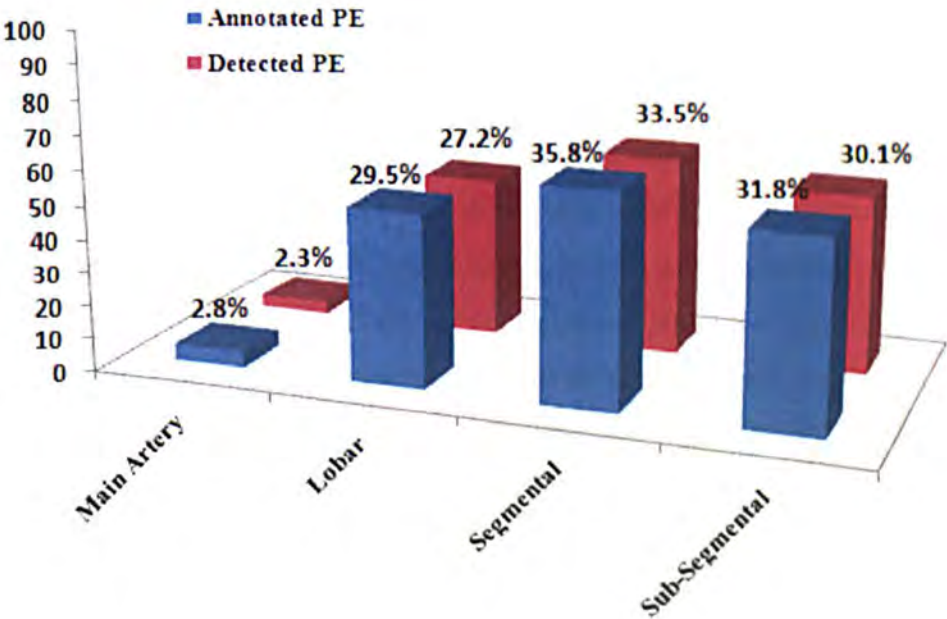
Figure 5-19 Presents the distribution of the number and the percentage of annotated PEs by radiologist and detected PEs using the proposed CAD system in different artery types.

Figure5-20 shows the FROC curve of detection result after removing false positive detection. The sensitivity of detection was adjusted by changing of volume size of candidates and counting the false positive detections per dataset. The FROC curve shows that the proposed CAD system can achieve sensitivities of 76%, 94% and 98% at 2.3, 4.1 and 12.6 false positive detections per dataset respectively.

False Positive per Dataset
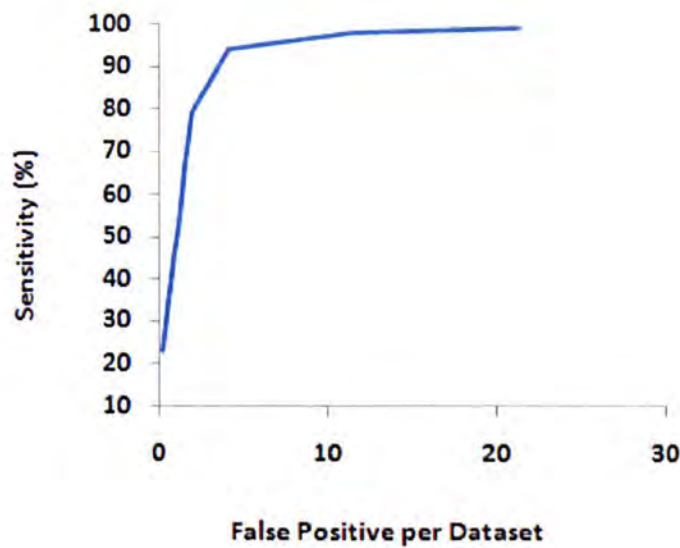
Figure 5-20 FROC curve of the proposed CAD system on the evaluation dataset

The FROC curve for PE detection (Figure5-20) shows the sensitivity of the proposed method is 94% at 4.1 false positive detections per dataset. Table5-6 compares the results of the proposed CAD system with those achieved by other researchers. Das et al evaluated a system to detect PEs which was tested on a large dataset but was applied only in the peripheral vessel region. Zhou et al's approach has a low sensitivity with high false positive detection rate. Although Masutani et al achieved high sensitivity detection, the evaluation was a small dataset with a low number of PEs. Buhmann evaluated a system on a large number of datasets but the results show the low sensitivity performance of his approach. The result of the proposed system by Maizlin et al shows poor sensitivity detection and small dataset selection. Bouma et al achieved sensitivity of 63% (22% extra) with 4.9 false positives per dataset, which corresponds to a lower sensitivity and a higher rate of false positive detection than our system. We believe that the reason why the proposed CAD system achieved high sensitivity (94%) with an average number of false detections of 4.1 per patient is due to the reduction of search area. Table 5-7 shows the

number of false detections (FP) in 20 training datasets by the proposed CAD system (using PA as a search area) and previous proposed CAD system (discussed in the literature review that used all pulmonary vessels as a search area).

Table 5-6 shows the comparison of different systems. The table shows the number of used datasets, the number of true positive PEs in dataset (No.PEs), the achieved sensitivity (Sens) and the number of false positive detections per dataset (FPs).

Ref	Dataset	No.PE		Sens	FP	Comment
Das	33	168	Seg	88	4.0	Only in peripheral artery
		120	Sub_Seg	78		
Zhou	14	225		52	11.4	Low Sensitivity
Masutani	11	21		100	7.7	Low PE
Maizlin	8	45		58	6.4	Low sensitivity and dataset
Buhmann	40	325		47	3.9	Low Sensitivity
Bouma	19	116		63+22	4.9	Low dataset
The Proposed CAD	35	195		94	4.1	High Sensitivity Less false Positive

Table 5-7 shows the number of false detections (FP) in 20 training dataset using the proposed CAD system and the previous proposed system.

Dataset	Previous CAD	Proposed CAD
T1	55	36
T2	38	27
T3	50	27
T4	61	37
T5	32	28
T6	51	29
T7	55	38
T8	48	25
T9	75	47
T10	48	38
T11	67	52
T12	51	33
T13	63	38
T14	73	41
T15	56	32
T16	61	27
T17	55	38
T18	63	42
T19	53	31
T20	68	42
Total FP	1123	708

Figure5-21 shows the display system of detection results. The detection results show that the proposed CAD system is able identify segmental and sub-segmental pulmonary embolisms which are often missed by radiologists. Figure5-21-c (first row) shows that the system detected one false positive.
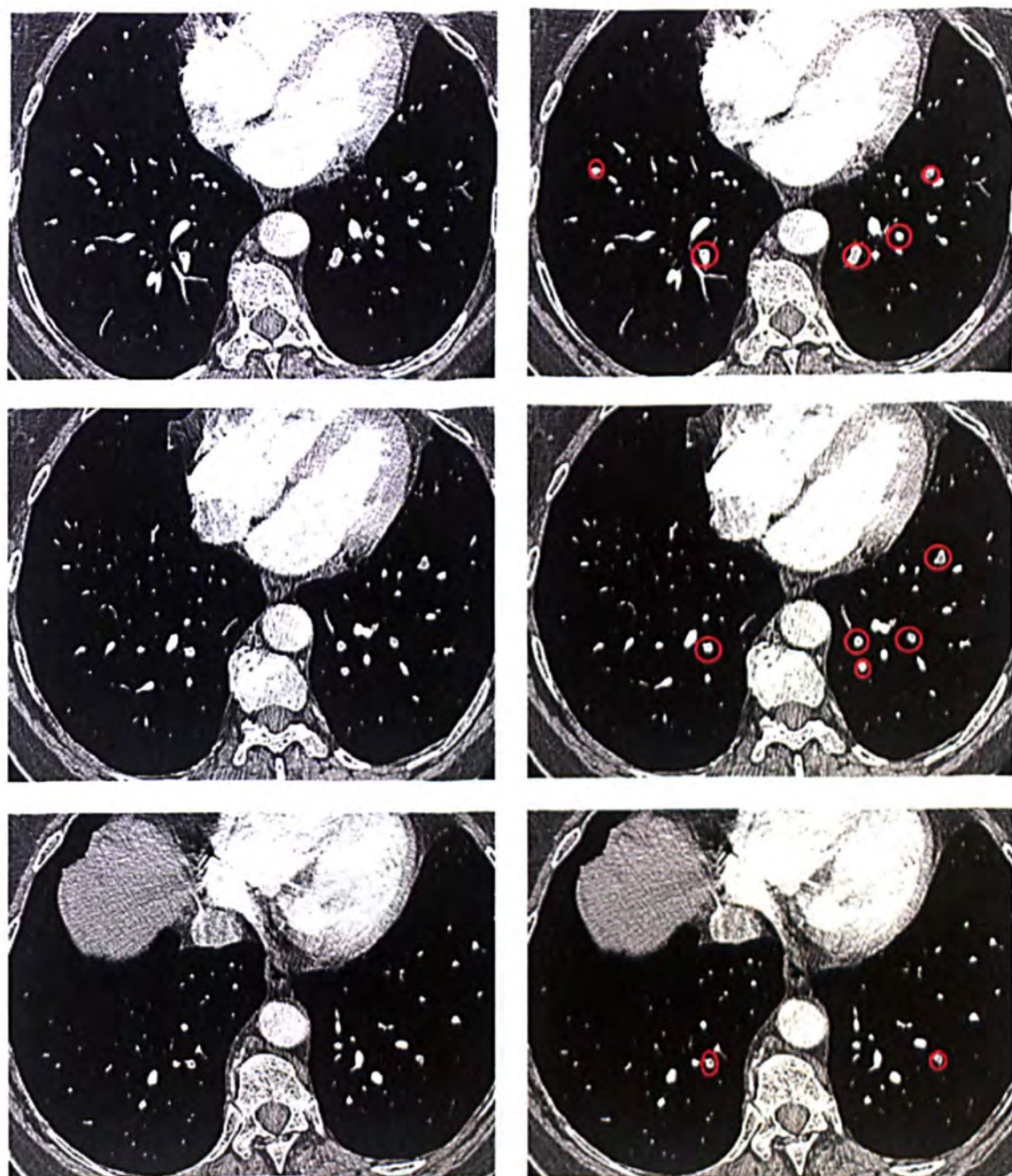
(a)
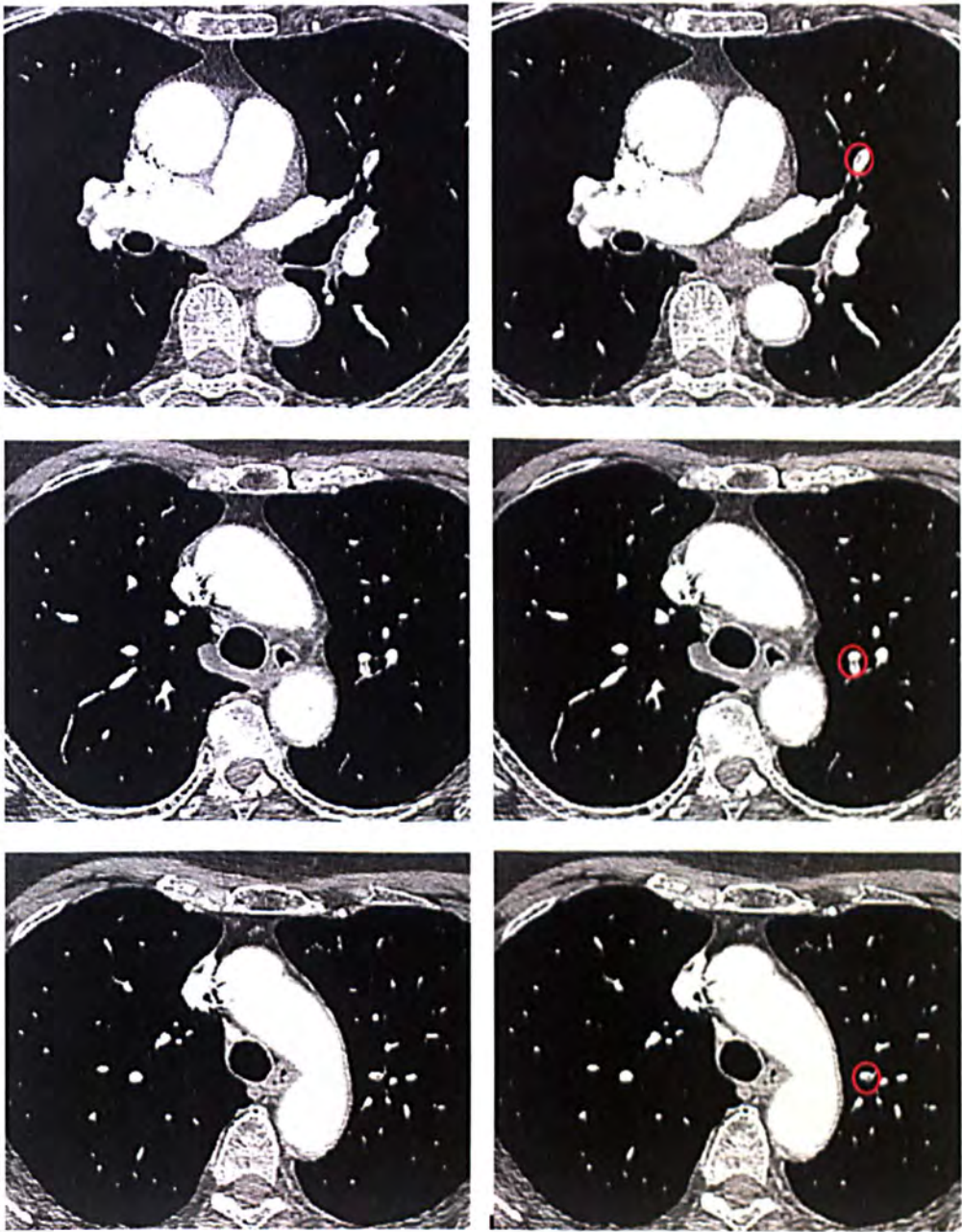
(b)
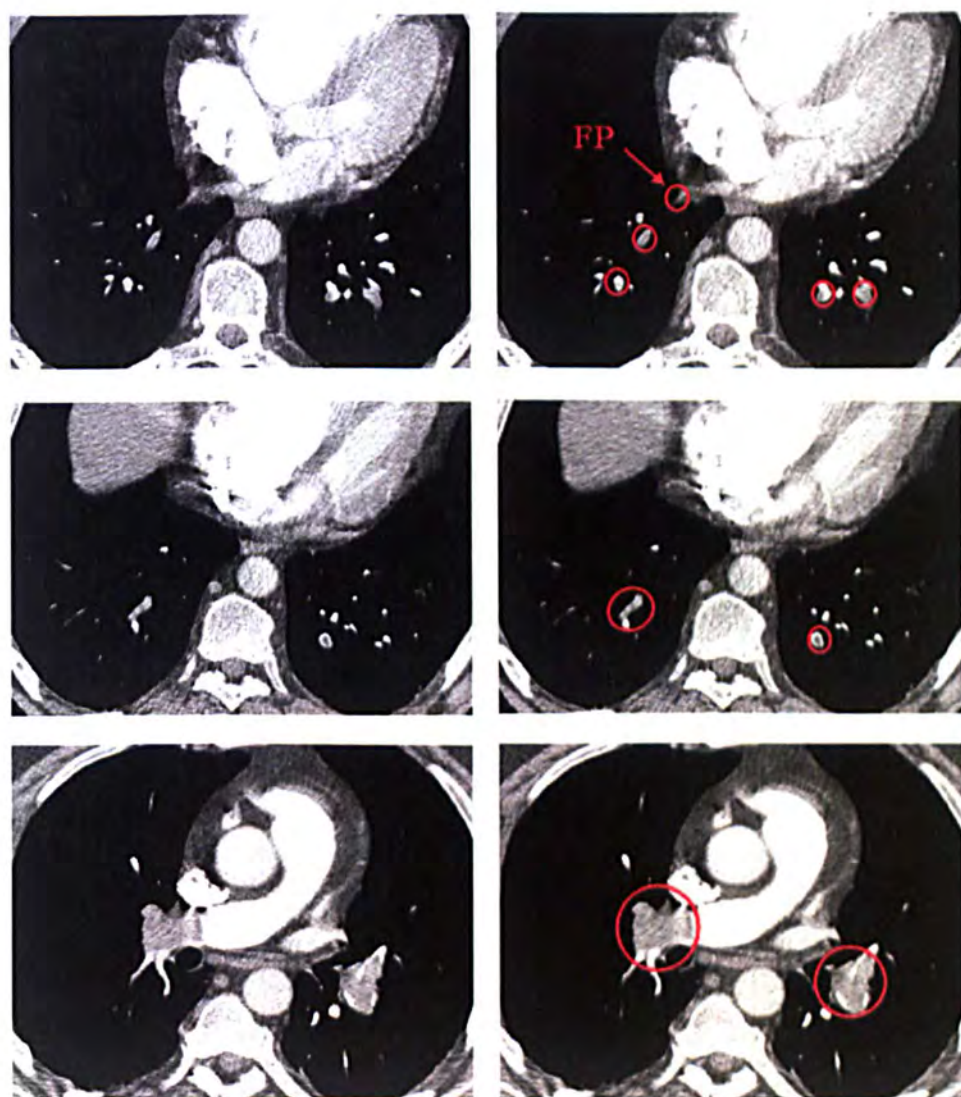
(c)
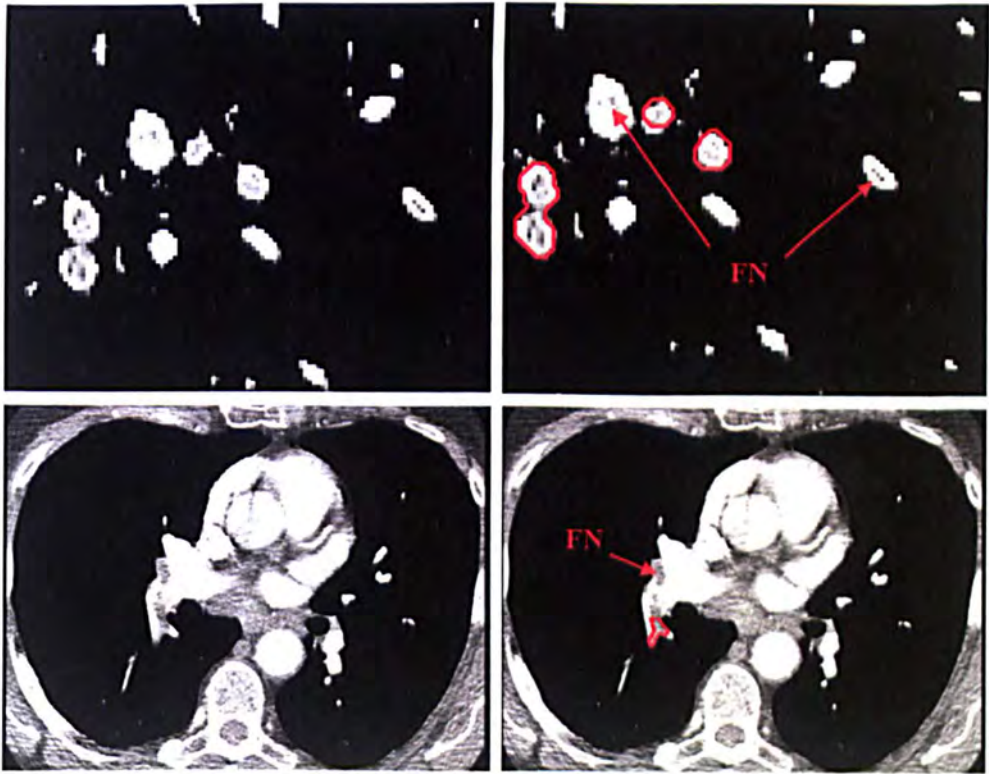
(d)

Figure 5-21 a-d Display the results of the PE CAD system. The left column shows the raw data with PEs and the right column presents the computer detection results.

The CAD system was tested on a PC with 2 GHz Quad core CPU and 8 GB memory. On average it takes 10.23 min per patient. Table 5-8 shows the elapsed time for different stages of the system.

Table 5-8   Average time for the major components of the proposed CAD system

Major CAD Components	Average Elapsed Time Minute
Lung Segmentation Process	1.28
Heart Segmentation Process	1.06
Feature Map Extraction Process	3.17
Pulmonary Artery (PA) Segmentation Process	2.49
Pulmonary Embolism Detection Process	1.34
Total Elapsed Time	**10.23**

The above estimated time was computed from starting stage of the CAD system which is loading the scan data to the final stage which is pulmonary embolism detection and removal of false detection without any optimization.

6. CONCLUSION

In this thesis a novel CAD system for detection of pulmonary embolisms in CTA images has been developed. The proposed CAD system consists of pulmonary artery (PA) segmentation, pulmonary embolism (PE) candidate detection and false positive reduction. The pulmonary artery has been extracted as a search area for pulmonary embolisms. Pulmonary artery segmentation is a challenging task due to the proximity of other organs, its elongated shape and range of vessel diameters. Flow voids in veins are one of the commonest causes for false positive detection of PEs, so separating the pulmonary artery from the vein network resulted in a lower false positive detection rate.

An efficient algorithm for segmenting the Pulmonary Artery (PA) tree has been developed in 3D pulmonary Computed Tomography Angiography (CTA) images. The pulmonary artery tree was extracted in two steps. In the first step, the major artery extracted in the heart region using a region growing method initialized by a seed point which was automatically selected within the pulmonary artery trunk. The segmentation

result of the major artery has been refined using a customized 3D level set algorithm. The major artery was extracted by the following stages:

1) Enhancing the image by a non-linear diffusion filter.
2) Extracting lung and heart region as a search area for pulmonary artery segmentation.
3) Removing the connectivity between the artery and other organs using a feature map based on eigenvalues of the Hessian matrix.
4) Removing the superior vena cava.
5) Segmenting the pulmonary artery using a 3D customized level set method starting from the initialized contour inside the pulmonary trunk.

In level set implementation, a new external energy was proposed which prevents front propagation to the background where the gradient of the boundary of object is weak. A new stopping criterion which uses a measure of the mean curve energy has been used to force the curve toward the object boundaries and to halt the evolution process at an optimum iteration, a consideration that has been ignored in many level set implementations.

In the second step, the peripheral pulmonary arterial network was extracted from the lung regions using a distinct region growing algorithm, starting from seed points which are end points of the left and right arteries and are also automatically selected.

As the proposed method fully automates the extraction of PAs, it allows the CAD system to be included in a fully automatic pulmonary embolism detection system. Unlike the previous proposed methods (discussed in the literature review in chapter2) in which the segmentation algorithm sometimes fails as a consequence of merging the pulmonary artery to adjacent tissue, the proposed algorithm applies the feature map filter to remove the connectivity between the pulmonary artery and other tissues.

Pulmonary embolism (PE) was identified as a dark region within the segmented artery. In the second stage, five different detectors based on intensity, geometric analysis, morphological operator (top-hat transform), analyzing the perpendicular plane to the centerline of the vessel, followed by ellipse fitting and searching disconnected parts of the

vessel, were implemented to find voxels inside the segmented artery as PE candidates. These five detectors respond to different properties of PE.

Soft tissue such as the lymphoid tissue surrounding vessels, partial volume effect, noise and motion artefacts are the main causes of false positive detections. A post-processing step was applied on the 3D features extracted from the detectors to reduce the number of false positive detections.

Small dark regions (caused by noise) can be identified as PEs everywhere. Therefore, the size of voxel groups was used as a feature to remove small voxel groups. The result of using this feature on 20 training datasets showed that much false positive detection can be excluded (PE candidates smaller than 42 voxels) without a significant reduction of the number of true positive detections (91% sensitivity and 46 false positives per dataset).

To deal with the false positive detections caused by the partial volume effect, voxels on the vessel wall were removed by a morphological operator. The thickness of the removal wall depends on the radius of the pulmonary artery noting that the thickness of the major artery is greater than in the peripheral artery.

A new feature based on location was proposed to remove false detection caused by soft tissue. A lymph false positive (soft tissue) connects to the airway system and is anatomically located outside the pulmonary artery and a true pulmonary embolism is inside the pulmonary artery. A filter was created by segmenting the airway to remove false detection caused by soft tissue. The result of applying just this filter on 20 training dataset identified 17 false positives per case at 84% sensitivity of detection.

The system has been tested on 35 CTA scans with average of 176 PEs, 68% of which were located in the peripheral arterial network (Segmental and Sub-Segmental) and 32% in the main pulmonary artery (Main and Lobar). The results showed that the system achieved 94% sensitivity at 4.1 false positives per case. Since the search area was restricted to inside the pulmonary artery, the results of our CAD system performed well in comparison to other CAD systems described in the literature.

The proposed CAD system is most helpful to detect PEs in the peripheral pulmonary arteries (segmental and sub-segmental arteries). This kind of PE is easily missed by a

radiologist and also indicates a risk for recurring of more significant emboli. The system is also able to add 11% more to the PEs manually detected by the radiologist. The main reason for the additional detected PE can be the size and location of the additional PE. Since the proposed CAD system is more capable to detect PE in segmental and sub-segmental artery, this kind of PE can be missed by the radiologist.

Our CAD system can remove false detections caused by partial volume, soft tissue, noise and flow voids (by separating the vein from pulmonary artery). Most of the remaining false positives (4.1 per CT scan) were caused by motion artefacts (Table 5-2). The movement artefacts usually create blurred boundaries in CT images which have the same intensity as a PE. False positives caused by the motion artefacts are an interesting field for the next research to further reduce the number of false positives. Although, this kind of false positive may be less important as a result of using newer CT scanners or improved scanning protocols in future. Nevertheless, future research should be focused on the recognition and rejection of this cause to further reduce the number of false positives or improve the sensitivity. Moreover, we will carry on measuring and labelling the detected PEs. For this purpose, the extracted pulmonary artery should be labelled according to the reference model of pulmonary artery.

REFERENCES

[1]. J. Juni and A. Abass, "Lung scanning in the diagnosis of pulmonary embolism," in Proc. Nucl. Med., vol. 21, pp. 282–296, 1991.

[2]. J. Ravenel, H.P. McAdams, P.C. Goodman, "Computed tomography pulmonary angiography: Diagnostic pitfalls and artifacts," Applied Radiology, vol. 30, no. 10, pp. 35–42, 2001.

[3]. M. Domingo, L. Mart -Bonmat, R. Dosd'a and Y. Pallard'o, "Interobserver agreement in the diagnosis of pulmonary embolism with helical CT," European Journal of Radiology, vol. 34, pp. 136–140, 2000.

[4]. FA. Anderson, HB. Wheeler, RJ. Goldberg, et al. , "A population-based perspective of the hospital incidence and case-fatality rates of deep vein thrombosis and pulmonary embolism", Arch Intern Med; 151:933-938., 1991

[5]. PD. Stein, SE. Fowler, LR. Goodman, et al, for the PIOPED II Investigators, "Multidetector computed tomography for acute pulmonary embolism", N Engl J Med; 354:2317-2327, 2006.

[6]. A. Reißig, J.P. Heyne and C. Kroegel, "Ancillary lung parenchymal findings at spiral CT scanning in pulmonary embolism", European Journal of Radiology, 2003.

REFERENCES

[7]. W.H. Crawford, "Pathology and Medicine", University of California, School of Dentistry, 2004.

[8]. J.A. Heit. The epidemiology of venous thromboembolism in the community: Implications for prevention and management. J. Thrombosis and Thrombolysis, Springer, 21(1):23–29, 2006.

[9]. S.Z. Goldhaber and R.B. Morrison, "Pulmonary embolism and deep vein thrombosis", Circulation, American Heart Association, 106(12):1436–1438, Sept. 2002.

[10]. P.S. Wells, J.S. Ginsberg and others, "Use of a clinical model for safe management of patients with suspected pulmonary embolism", Annals of Internal Medicine, 129(12):997–1005, Dec. 1998.

[11]. J. Wicki, T.V. Perneger, A.F. Junod, H. Bounameaux and A. Perrier," Assessing clinical probability of pulmonary embolism in the emergency ward a simple score", Arch. Intern. Med., 161:92–97, 2001.

[12]. J.H. Ryu, S.J. Swensen, E.J. Olson and P.A. Pellikka, "Diagnosis of pulmonary embolism with use of computed tomography angiography", In Mayo Clinic Proc., volume 76, pages 59–65, Jan. 2001.

[13]. C. Kearon, "Diagnosis of pulmonary embolism", Canadian Medical Association J., 168(2):183–194, Jan. 2003.

[14]. C. Fink, S.O. Schoenberg and M.F. Reiser, "MRI expands options in lung assessment", Diagnostic Imaging Europe, pages 14–19, 47, Nov. 2005.

[15]. M. Remy-Jardin and J. Remy,"Spiral CT angiography of the pulmonary circulation", Radiology (RSNA), 212:615–636, Sep.1999.

[16]. M. Remy-Jardin, J. Remy, L. Wattinne, F. Giraud, "Central pulmonary thromboembolism: diagnosis with spiral volumetric CT with single-breath-hold technique-comparison with pulmonary angiography", Radiology (RSNA), 185:381–387, 1992.

[17]. E. Coche, "Evaluation of Multislice CT for the Diagnosis of Pulmonary Embolism", PhD thesis, Universit'e catholique de Louvain, Facult'e de M'edecine, 2005.

REFERENCES

[18]. Remy-Jardin M, Remy J, Deschildre F, et al,"Diagnosis of pulmonary embolism with spiral CT: Comparison with pulmonary angiography and scintigraphy", Radiology.; 200:699-706, 1996.

[19]. Kim K-I, Müller NL, Mayo JR, "Clinically suspected pulmonary embolism: Utility of spiral CT", Radiology;210:693-697, 1999.

[20]. PF Fedullo, VF Tapson, "The evaluation of suspected pulmonary embolism", N Engl J Med;349(13):1247, 2003.

[21]. A. Perrier, N. Howarth, D. Didier; et al, " Performances of helical computed tomography in unselected outpatients with suspected pulmonary embolism", Ann Intern Med;135(2):88, 2001.

[22]. U.J. Schoepf, M.A. Kessler, C.T. Rieger and others, "Multislice CT imaging of pulmonary embolism", European Radiology, 11(11):2278–2286, 2001.

[23]. Winer-Muram HT, Rydberg J, Johnson MS, Tarver RD, Williams MD, Shah H, et al, "Suspected acute pulmonary embolism: evaluation with multi-detector row ct versus digital subtraction pulmonary arteriography", Radiology ;233(3):806-15, 2004.

[24]. P.D. Stein, S.E. Fowler, L.R. Goodman and others, "Multi-detector computed tomography for acute pulmonary embolism", New England J. Medicine, 354(22):2317–2327, 2006.

[25]. Y. Masutani, H. MacMahon, and K. Doi, "Computerized detection of pulmonary embolism in spiral CT angiography based on volumetric image analysis," IEEE Trans. Med. Imag., vol. 21, no. 12, pp. 1517–1523, 2002.

[26]. Zhou C, Chan HP, Patel S, et al, " Preliminary investigation of computer aided detection of pulmonary embolism in 3Dcomputed tomographic pulmonary angiography (CTPA) images", Acad Radiol 12:782-792; 2005.

[27]. Zhou C, Chan HP, Hadjiiski LM, et al, "Automated detection of pulmonary embolism (PE) in computed tomographic pulmonary angiographic (CTPA) images:

multiscale hierachical expectation-maximization segmentation of vessels and PEs", Proc SPIE; 6514:2F1-2F8, 2007.

[28]. Das M, Schneider AC, Schoepf UO, et al. computer aided diagnosis of peripheral pulmonary emboli. RSNA Program Book:351-352, 2003.

[29]. Jeudy J, Flukinger T, White C, "Evaluation of pulmonary embolism using an automated computer-aided detection tool", RSNA Program Book: 255, 2006.

[30]. Schoepf UJ, Schneider AC, Das M, et al, "Pulmonary embolism: computer-aided detection at multidetector row spiral computed tomography", J Thorac Imag; 22:319-323, 2007.

[31]. ZV. Maizlin, Vos , MB. Godoy, et al, "Computer-aided detection of pulmonary embolism on CT angiography", initial experience. J Thorac Image; 22:324-329, 2007.

[32]. Das M, Salganicoff M, Bakai A, Muhlenburn G, Gunther Rw, Wildberger J. "Computer-aided detection of pulmonary embolism: assessment of sensitivity with regard to vessel segments", RSNA, Chicago, IL, 2006; RSNA Program Book:487, 2006.

[33]. Buhmann S, Herzog P, Liang J, et al, " Clinical evaluation of a computer-aided diagnosis (CAD) prototype for the detection of pulmonary embolism", Acad Radiol; 14:651-658, 2007.

[34]. H. Bouma, J. J. Sonnemans, A. Vilanova, F. A. Gerritsen, "Automatic Detection of Pulmonary Embolismin CTA Images," IEEE Trans. Med. Imag., vol. 28, no. 8, pp.1223-1230, 2009.

[35]. J Liang and J Bi. "Computer aided detection of pulmonary embolism with tobogganing and multiple instance classification in CT pulmonary angiography". Springer-Verlag Berlin, Heidelberg ,2007.

[36]. R jain, R Kasturi, B.G, "Schunck, Machine Vision", McGill, 1995.

[37]. M. Donizelli, "Region-Oriented segmentation of vascular structures from DSA images using mathematical morphology and binary region growing", 1998.

REFERENCES

[38]. T. Tozaki,, Y. Kawata, N. Niki, H. Ohmatsu, and N. Moriyama, "3-d visualization of blood vessels and tumor using thin slice ct", in IEEE Nuclear Science Symposium and Medical Imaging Conference, vol.3, pp. 1470-1474, 1995.

[39]. S. Osher, J. A. Sethian, "Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations", J. Comp. Phys., vol. 79, pp. 12-49, 1988.

[40]. Y. Sato, S. Nakajima, N. Shiraga, et al., "Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images," Medical Image Analysis, vol. 2, no. 2, pp. 143–168, 1998.

[41]. K. Krissian, G. Malandain, N. Ayache, R. Vaillant, and Y.Trousset, "Model-based detection of tubular structures in 3Dimages," Computer Vision and Image Understanding, vol. 80,no. 2, pp. 130–171, 2000.

[42]. H. Shikata, E. A. Huffman, and M. Sonka, "Automated segmentation of pulmonary vascular tree from 3D CT images," Physiology, Function, and Structure from Medical Images, vol. 5369 of Proceedings of the SPIE, pp. 107–116, San Diego, Calif, USA, 2004.

[43]. C. Zahlten, H. Juergens, and H.-O. Peitgen, "Reconstruction of branching blood vessels from ct-data", Proceedings of the Eurographics Workshop on Visualization in Scientific Computing, Rostock, pages 41–52, 1994.

[44]. Sebbe, R., B. Gosselin, E. Coche, et al., "Pulmonary arteries segmentation and feature extraction through slice marching,"Proc. ProRISC workshop on Circuits, Systems and Signal Processing, 2003.

[45]. Perona and J. Malik, "Scale-Space and Edge Detection Using Anisotropic Diffusion", IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(7):629-639, 1990.

[46]. R. Uppaluri, T. Mitsa, M. Sonka, E. A. Hoffman, and G. McLennan, "Quantification of pulmonary emphysema from lung CT images using texture analysis," Amer. J. Resp. Crit. Care Med., vol. 156, no. 1, pp. 248–254, 1997.

REFERENCES

[47]. A. M. Boriek, S. Liu, and J. R. Rodarte, "Costal diaphragm curvature in the dog," J. Appl. Physiol., vol. 75, no. 2, pp. 527–533, 1993.

[48]. I. Amirav, S. S. Kramer, M. M. Grunstein, and E. A. Hoffman, "Assessment of methacholine-induced airway constriction by ultrafast high-resolution computed tomography," J. Appl. Physiol., vol. 75, no. 5, pp. 2239–2250, 1993.

[49]. S. Hu, E. A. Hoffman, J. M. Reinhardt, "Automatic Lung Segmentation for Accurate Quantitation of Volumetric X-Ray CT Images", IEEE Transactions on medical imaging, vol. 20, NO. 6, 2001.

[50]. M. S. Brown, M. F. McNitt-Gray, N. J. Mankovich, J. G. Goldin, J. Hiller, L. S. Wilson, and D. R. Aberle, "Method for segmenting chest CT image data using an anatomical model: Preliminary results," IEEE Trans. Med. Imag., vol. 16, pp. 828–839, 1997.

[51]. Lindeberg, "Feature detection with automatic scale selection, "International Journal of Computer Vision. 30(2),79-116, 1998.

[52]. Caselles, R. Kimmel and G. Sapiro, "Geodesic active contours," Int. J. Comput. Vision, 22 (1), pp. 61–79, 1997.

[53]. R. Malladi, J. A. Sethian, and B. C. Vemuri, "Shape modelling with front propagation: a level set approach", IEEE Trans. Patt. Anal. Mach. Intell., vol. 17, pp. 158-175, 1995.

[54]. M. Sussmen, P. Smereka, and S. J. Osher, "A Level Set Method for Computing Solutions to Incompressible Two-Phase Flow", J. Comp. Phys. 114, pp. 146-159, 1994.

[55]. J. A. Sethian, Level set methods and fast marching methods, Cambridge: Cambridge University Press, 1999.

[56]. A. Frangi, W. Niessen, K. Vincken, M. Viergever," Multi-scale vessel enhancement filtering", in: Medical Image Computing Computer-Assisted Intervention, LNCS, vol. 1496, pp. 130–137, 1998.

REFERENCES

[57]. R. Mannieshing, M.A. Viergever, and W. J . Niessen," Vessel enhancing diffusion: A scale space representation of vessel structures", Medical Image Analysis, 2006.

[58]. E. Pichon, C.L. Novak, A.P. Kiraly and D.P. Naidich," A novel method for pulmonary emboli visualization from high-resolution CT images", In Proc. SPIE Medical Imaging, volume 5367, pages 161–170, 2004.

[59]. P. Soille, Morphological Image Analysis. New York: Springer-Verlag, 1999.

[60]. J. Koenderink and A. van Doorn, "Surface Shape and Curvature Scales," Image and Vision Computing, vol. 10, pp. 557–565, 1992.

[61]. L.J. van Vliet and P.W. Verbeek, "Curvature and bending energy in digitized 2D and 3D images", Proc. Scandinavian Conf. Image Analysis, pages 1403–1410, Tromso, Norway, 1993.

[62]. R. Kimmel and J.A. Sethian, "Optimal Algorithm for Shape from Shading and Path Planning," J. Math. Imaging and Vision, vol. 14, no. 3, pp. 237-244, 2001.

[63]. M.S. Hassouna, A.E. Abdel-Hakim, and A. Farag, "Robust Robotic Path Planning Using Level Sets," Proc. IEEE Int'l Conf. Image Processing, pp. 473-476, 2005.

[64]. R. Kimmel, D. Shaked, N. Kiryati, and A. Bruckstein, "Skeletonization via Distance Maps and Level Sets," vol. 62, no. 3, pp. 382- 391, 1995.

[65]. M.S. Hassouna and A.A. Farag, "Robust Centerline Extraction Framework Using Level Sets," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 458-465, June 2005.

[66]. Z. Cao, S. Pan, R. Li, R. Balachandran, M.J. Fitzpatrick, W.C. Chapman, and B.M. Dawant, "Registration of Medical Images Using an Interpolated Closest Point Transform: Method and Validation," Proc. SPIE Int'l Soc. Optical Eng. Medical Imaging, pp. 325-333, 2003.

[67]. H.E.A.E. Munim and A.A. Farag, "A Shape-Based Segmentation Approach: An Improved Technique Using Level Sets," Proc. IEEE Int'l Conf. Computer Vision, pp. 930-935, 2005.

REFERENCES

[68]. D. Adalsteinsson and J. Sethian, "A Fast Level Set Method for Propagating Interfaces," J. Computational Physics, vol. 118, pp. 269- 277, 1995.

[69]. T. Deschamps and L.D. Cohen, "Fast Extraction of Tubular and Tree 3D Surfaces with Front Propagation Methods," Proc. Int'l Conf. Pattern Recognition, pp. 731-734, 2002.

[70]. M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," Int. J. Comput. Vis, vol. 1, pp. 321–331, 1987.

[71]. R. Kimmel, A. Amir, and A. Bruckstein, "Finding shortest paths on surfaces using level set propagation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 17, no. 6, pp. 635–640, 1995.

[72]. C. Li, C. Xu, C. Gui, and M. D. Fox, "Level set evolution without re-initialization: A new variational formulation," in Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2005, vol. 1, pp. 430–436.

[73]. R. Malladi, J. A. Sethian, and B. C.Vemuri, "Shape modeling with front propagation: a level set approach," IEEE Trans. Pattern Anal. Mach. Intell., vol. 17, no. 2, pp. 158–175, 1995.

[74]. T. Chan and L. Vese, "Active contours without edges," IEEE Trans. Image Process., vol. 10, no. 2, pp. 266–277, 2001.

[75]. N. Paragios and R. Deriche, "Geodesic active regions and level set methods for supervised texture segmentation," Int. J. Comput. Vis., vol. 46, pp. 223–247, 2002.

[76]. R. Ronfard, "Region-based strategies for active contour models," Int. J. Comput. Vis., vol. 13, pp. 229–251, 1994.

[77]. A. Tsai, A. Yezzi, and A. S.Willsky, "Curve evolution implementation of the Mumford–Shah functional for image segmentation, denoising, interpolation, and magnification," IEEE Trans. Image Process., vol. 10, no. 8, pp. 1169–1186, 2001.

[78]. L. Vese and T. Chan, "A multiphase level set framework for image segmentation using the Mumford and Shah model," Int. J. Comput. Vis., vol. 50, pp. 271–293, 2002.

[79]. C. Samson, L. Blanc-Feraud, G. Aubert, and J. Zerubia, "A variational model for image classification and restoration," IEEE Trans. Patt. Anal. Mach. Intell., vol. 22, no. 5, pp. 460–472, 2000.

[80]. D. Mumford and J. Shah, "Optimal approximation by piecewise smooth functions and associated variational problems," Commun. Pure Appl. Math, vol. 42, pp. 577–685, 1989.

[81]. D. Mumford, J. Shah, " Optimal approximations by piecewise smooth functions and associated variational problems", Commun. Pure Appl. Math., vol. 42, no. 4 , 1989.

[82]. C. Li, C. Kao, J. C. Gore, and Z. Ding, "Minimization of region-scalable fitting energy for image segmentation," IEEE Trans. Image Process., vol. 17, no. 10, pp. 1940–1949, 2008.

[83]. J.A. Sethian," A fast marching level set method for monotonically advancing fronts", In Proc. Nat. Acad. Sci., volume 93 of 4, pages 1591-1595, 1996.

[84]. J.A. Sethian," Level Set Methods: Evolving interfaces in geometry, fluid mechanics, computer vision, and materials science", Cambridge University Press, 1996.

[85]. H. Zhao, "A Fast Sweeping Method for Eikonal Equations," J. Math. and Computing, vol. 74, no. 250, pp. 603-627, 2005.

[86]. D. Adalsteinsson and J. Sethian, "A Fast Level Set Method for Propagating Interfaces," J. Computational Physics, vol. 118, pp. 269- 277, 1995.

[87]. R. Kimmel and J. Sethian, "Fast Marching Methods on Triangulated Domains," Proc. Nat'l Academy of Sciences, vol. 95, no. 11, pp. 8341-8435, 1998.

[88]. J.A. Sethian and A. Vladimirsky, "Fast Methods for the Eikonal and Related Hamilton-Jacobi Equations on Unstructured Meshes," Proc. Nat'l Academy of Sciences, vol. 97, no. 11, pp. 5699-5703, 2000.

[89]. S. Osher and J. Sethian, "Fronts Propagating with Curvature Speed: Algorithms Based on Hamilton-Jacobi Formulations," J. Computational Physics, vol. 79, pp. 12-49, 1988.

[90]. E. Rouy, A. Tourin,"A viscosity solutions approach to shape-from-shading", SIAM J. Numer. Anal., 29(3):867–884, 1992.

[91]. J. Sethian. Level Sets Methods and Fast Marching Methods. Cambridge University Press, 2nd edition, 1999.

[92]. J. Tsitsiklis, "Efficient algorithms for globally optimal trajectories", IEEE Trans. On Automatic Control, 40(9):1528–1538, 1995.

[93]. J. J. Helmsen, E. G. Puckett, P. Colella, and M. Dorr, "Two new methods for simulating photolithography development in 3D", In Proc. SPIE Vol. 2726, p. 253-261, Optical Microlithography IX, Gene E. Fuller; Ed., pages 253–261, 1996.

[94]. H. Blum, "A transformation for extracting new descriptors of shape," in Proceeding of Models for the perception of Speech and Visual form, edited by Weiant Wathen-Dunn (MIT, Cambridge, pp 362-380, 1967.

[95]. H. Frimmel, J. Nappi, and H. Yoshida, "Centerline-based colon segmentation for CT colonography", Med. Phys., vol. 32, no. 8, pp. 2665–2672, 2005.

[96]. P. Li, S. Napel, B. Acar, D. S. Paik, R. B. Jeffrey Jr., and C. F. Beaulieu, "Registration of central paths and colonic polyps between supine and prone scans in computed tomography colonography: Pilot study," Med. Phys., vol. 31, no. 10, pp. 2912–2923, 2004.

[97]. H. Greenspan, M. Laifenfeld, S. Einav, and O. Barnea, "Evaluation of center-line extraction algorithms in quantitative coronary angiography," IEEE Trans. Med. Imag., vol. 20, no. 9, pp. 928–941, 2001.

[98]. C. Pudney, "Distance-ordered homotopic thinning: a skeletonization algorithm for 3-D digital images," Comput. Vision Image Understand. vol. 72, no. 3, pp. 404–413, 1998.

[99] R. V. Uitert, I. Bitter, "Subvoxel precise skeletons of volumetric data based on fast marching methods", Medical Physics. Vol. 34, pp, 627-638, 2007

REFERENCES

[100]. M. Sabry Hassouna, A. Farag ,"Multistencils Fast Marching Methods: A Highly Accurate Solution to the Eikonal Equation on Cartesian Domains," IEEE Trans on Patern Analysis and Machine Intelligence, Vol. 29, no. 9, 2007.

[101]. C.I. Fetita, F. Prteux, ,C. Beigelman-Aubry, P. Grenier, , "Pulmonary airways: 3-D reconstruction from multislice CT and clinical investigation", IEEE Transactions on Medical Imaging 23(11) ,1353-1364, 2004.

[102]. D. Aykac, E.A. Huffman, G. McLennan, J.M. Reinhardt, "Segmentation and analysis of the human airway tree from three-dimensional X-ray CT images", IEEE Transactions on Medical Imaging 22(8) ,940-950, 2003.

[103]. B.V. Ginneken, W. Baggerman, E.M.V. Rikxoort, " Robust segmentation and anatomical labeling of the airway tree from thoracic CT scans. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 5241 LNCS (PART 1), 219-226, 2008.

[104]. Tschirren, J., Huffman, E.A., McLennan, G., Sonka, M.: Intrathoracic airway trees: Segmentation and airway morphology analysis from low-dose CT scans. IEEE Transactions on Medical Imaging 24(12), 1529-1539, 2005.

[105]. Singh, H., Crawford, M., Curtin, J.P., Zwiggelaar, R.: Automated 3D segmentation of the lung airway tree using gain-based region growing approach. MICCAI 2, 975-982, 2004.

[106]. Mori, K., Hasegawa, J.I., Suenaga, Y., Toriwaki, J.I.: Automated anatomical labelling of the bronchial branch and its application to the virtual bronchoscopy system. IEEE Transactions on Medical Imaging 19(2), 103-114, 2000.

[107]. M. Sonka, W. Park, and E. A. Hoffman, "Rule-based detection of intrathoracic airway trees", Comput. Med. Imag. Graph., vol. 15, pp. 314–326, 1996.

[108]. M.S. Aksoy, O. Torkul, and I. H. Cedimoglu. "An industrial visual inspection system that uses inductive learning." Journal of Intelligent Manufacturing 15.4 : 569(6). Expanded Academic ASAP. Thomson Gale. 2004.

REFERENCES

[109]. Kyriacou, Theocharis, G. Bugmann, and S. Lauria, "Vision-based urban navigation procedures for verbally instructed robots", Robotics and Autonomous Systems 51.1, Expanded Academic ASAP. Thomson Gale, 69-80, 2005.

[110]. R.R. Petrocelli, J. Elion, and K. M. Manbeck, "A new method for structure recognition in unsubtracted digital angiograms", in IEEE Computers in Cardiology, pp. 207–210, 1992.

[111]. R. van derWeide, C.J.G. Bakker, and M.A. Viergever, "Localization of intravascular devices with paramagnetic markers in MRI images", IEEE Trans. on Med. Img., vol. 20, pp. 1061–1071, 2001.

[112]. C. Smets, G. Verbeeck, P. Suetens, and A. Oosterlinck, "A knowledge-based system for the delineation of blood vessels on subtraction angiograms", Pattern Rec. Let., vol. 8, pp. 113–121, 1988.

[113]. U. Rost, H. Munkel, and C.-E. Liedtke, "A knowledge based system for the configuration of image processing algorithms", Fachtagung Informations und Mikrosystem Technik, 1998.

[114]. V. Bombardier, M.C. Jaluent, A. Bubel, and J. Bremont, "Cooperation of two fuzzy segmentation operators for digital subtracted angiograms analysis", in IEEE Conference on Fuzzy Systems, vol. 2, pp. 1057–1062, 1997.

[115]. D.P. Kottke and Y. Sun, "Segmentation of coronary arteriograms by iterative ternary classsification", IEEE Trans. on Biomed. Engr., vol. 37, pp. 778–785, 1990.

[116]. S. Shiffman, G. D. Rubin, and S. Napel, "Semiautomated editing of computed tomography sections for visualization of vasculature", vol. 2707, SPIE, 1996.

[117]. R. Nekovei and Y. Sun, "Back-propagation network and its configuration for blood vessel detection in angiograms", IEEE Trans. on Neural Nets, vol. 6, pp. 64–72, 1995.

[118]. B. Ghaye, D. Szapiro, I. Mastora, et al, "Peripheral pulmonary arteries: how far in the lung does multi-detector row spiral CT allow analysis?", Radiology; 219 :629-636, 2001.

[119]. J.E. Bresenham, "Algorithm for Computer Control of a Digital Plotter", IBM Systems Journal, 4(1), 25-30, 1965.

[120]. T. Deschamps, J.M. L'etang, B. Verdonck and L.D. Cohen, "Automatic contruction of minimal paths in 3D images: An application to virtual endoscopy", In Proc. Computer Assisted Radiology and Surgery (CARS), number 13, Elsevier, 1999.

[121]. T. B"ulow, C. Lorenz and S. Renisch, "A general framework for tree segmentation and reconstruction from medical volume data", In Proc. Medical Image Computing and Computer Assisted Intervention (MICCAI), volume 3216 of Springer LNCS, pages 533–540, 2004.

[122]. E. Konukoglu, B. Acar, D.S. Paik, C.F. Beaulieu and S. Napel, "Heat diffusion based detection of colonic polyps in CT colonography", In Proc. European Signal Processing Conf., 2005.

[123] T. Bulow, R. Wiemker, T. Bla_ert, C. Lorenz, S. Renisch, "Automatic extraction of the pulmonary artery tree from multi-slice CT data", Medical Imaging: Proceedings of SPIE 5746, 730-740, 2005.

[124] E. Z. Susan Wood, J. Hoford, E. Ho_man, W. Mitzner, " Measurement of three-dimension lung tree structures by using computed tomography", Journal of Applied Physiology 79, 1687-1697, 1995.

[125] Y. Masutani, H. MacMahon, K. Doi, "Automated segementation and visualization of the pulmonary vascular tree in spiral CT angiography: an anatomy-oriented approach based on three-dimensional image analysis", Journal of Computer Assisted Tomography 25, 587-597, 2001.

[126] C. Zhou, H.-P. Chan, B. Sahiner, L. Hadjiski, A. Chughtai, "Automatic multiscale enhancement and segmentation of pulmonary vessels in CT pulmonary angiography images for CAD applications", Medical Physics 34, 4567-4577, 2007.

[127] H. Shikata, G. McLennan, E. Ho_man, M. Sonka, "Segmentation of pulmonary vascular trees from thoracic 3D CT images", International Journal of Biomedical Imaging, 2009.

REFERENCES

[128] J. Zhou, S. Chang, D. Metaxas, L. Axel, "Vascular structure segmentation and bifurcation detection", IEEE International Symposium on Biomedical Imaging, 872-875, 2007.

[129]. J. N. Kaftan, A. Bakai, M. Das, T. Aach, "Locally adaptive fuzzy pulmonary vessel segmentation in contrast enhanced CT", Proceedings 5th international symposium on biomedical imaging: from nano to macro, 101-104, 2008.

[130]. V. Chalana and Y. Kim, "A methodology for evaluation of boundary detection algorithms on medical images," IEEE Trans. Med. Imag., vol. 16, pp. 642–652, Oct. 1997.

[131]. K. N. Chaudhury and K. R. Ramakrishnan, "Stability and Convergence of the level set method in computer vision", Pattern Recognition Letters, vol. 28, pp. 884-893, 2007.

[132]. T. Kapur, W. E. Grimson, W. M. Wells, III, and R. Kikinis, "Segmentation of brain tissue from magnetic resonance images" , Med. Image Anal., vol. 1, pp. 109- 127, 1996.

[133]. J. Wala, S. Fotin, J. Lee, A. Jirapatnakul, A. Biancardi and A. Reeves, "Automated segmentation of the pulmonary arteries in low-dose CT by vessel tracking" , Med. Image Anal, CoRR abs/1106.5460, 2011.

[134]. R. Collins, G. Cranny, J. Burch, et al. "A systematic review of duplex ultrasound, magnetic resonance angiography and computed tomography angiography for the diagnosis and assessment of symptomatic, lower limb peripheral arterial disease", *Health Technol Assess.*;11(20):iii-iv, xi-xiii, 1-184, 2007.

[135]. DW. Ramzi, KV Leeper, "DVT and pulmonary embolism: Part I. Diagnosis", Am Fam Physician; 69(12):2829-2836, 2004.

[136]. O. Figueiredo, "Advances in discrete geometry applied to the extraction of planes and surfaces from 3D volumes", PhD Thesis, Lausanne, EPFL, 1999.

REFERENCES

[137]. A. P. Zijdenbos, B. M. Dawant, R. A. Margolin, A. C. Palmer, " Morphometric analysis of white matter lesions in MR images: method and validation", IEEE Trans Med Imaging;13:716–724, 1994.

APPENDIX A

Differential Geometry

In this project, in order to pre-process or post-process volume data sets some features need to be calculated. For example for classifying the voxels based on their shape, locations and texture several features of image must be extracted. The following is a brief description of using differential geometry tools for extracting some features which will be used for doing the project.

Gaussian Smoothing Filter

The Gaussian filter is a linear smoothing filter whose weights are chosen based on the shape of the Gaussian function:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{1}$$
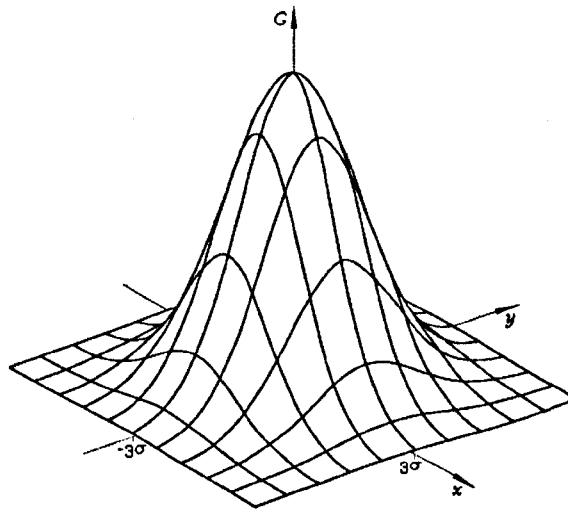
Figure 0-1 2D Gaussian function and distribution

The Gaussian smoothing filter is capable of removing the noise which is drawn from a normal distribution.

First Order Derivatives Operator

An edge in an image is where there is a discontinuity in the intensity function or a very steep intensity gradient. By taking the first order derivatives of intensity values and finding the maximum values for the derivatives, the edges of image are marked. The gradient of an image I which contains of first derivatives operators is defined as:

$$\nabla I = \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, \frac{\partial I}{\partial z}\right) \tag{2}$$

The gradient at each voxel has the following properties:

- The gradient direction at each voxel is the direction of steepest ascent at that voxel and the gradient magnitude:

$$|\nabla I| = \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2 + \left(\frac{\partial I}{\partial z}\right)^2} \tag{3}$$

is the steepness of that ascent at that voxel. This feature can be used for detecting of edges in an image.

- The direction of gradient at each voxel is the normal to the level curve at that voxel

- By using the gradient operator the first derivatives of an image at each voxel in any direction can be calculated:

$$L_u = \overline{u} \cdot \nabla L \tag{4}$$

Fig.ure0-2 shows an ideal steep edge and the profile of its derivatives.
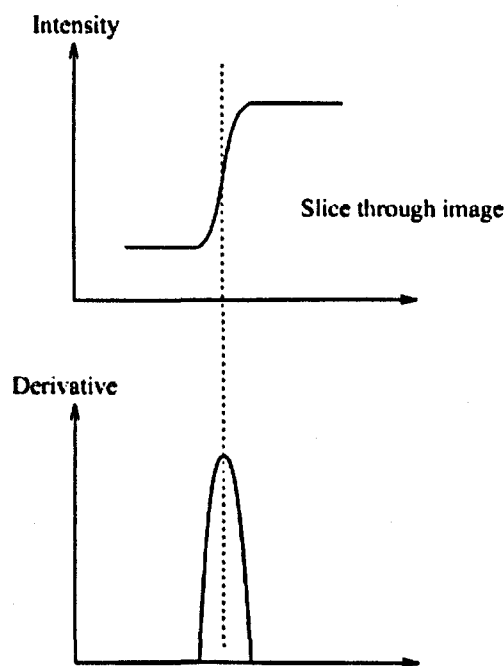
Figure 0-2 shows a steep edge and its derivative's profile

Second Order Derivative Operators

In second order geometry, the matrix of second order derivatives or the Hessian matrix is equivalent to the first order derivatives of the gradient operator. The information of second order derivatives can be used for marking an edge and extracting local shape of a gray scale image. The second order derivative operator of an image I or Hessian matrix is defined as follows:

$$H(I) = \begin{pmatrix} \dfrac{\partial^2 I}{\partial x^2} & \dfrac{\partial^2 I}{\partial x \partial y} & \dfrac{\partial^2 I}{\partial x \partial z} \\[2mm] \dfrac{\partial^2 I}{\partial y \partial x} & \dfrac{\partial^2 I}{\partial y^2} & \dfrac{\partial^2 I}{\partial y \partial z} \\[2mm] \dfrac{\partial^2 I}{\partial z \partial x} & \dfrac{\partial^2 I}{\partial z \partial y} & \dfrac{\partial^2 I}{\partial z^2} \end{pmatrix} \tag{5}$$

The eigenvalue and eigenvector of the Hessian matrix has a few geometric meanings:

- The direction of the greatest curvature (maximum curvature) can be identified by the first eigenvector (the eigenvector whose corresponding eigenvalue is largest absolute value).

- The eigenvector whose corresponding eigenvalue is the smallest one in the direction of the minimum curvature.

- The values of curvatures (maximum and minimum) are the corresponding of eigenvalues.

The Laplacian operator of an image I at each voxel which is defined as follow:

$$\nabla^2 I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} + \frac{\partial^2 I}{\partial z^2} \tag{6}$$

measures the second spatial derivatives of an image and shows the rapid intensity changes.

Mean and Gaussian Curvature

The mean curvature at point M of a surface S is the average of principal curvatures κ_1 and κ_2 and the Gaussian curvature is the product of these two principal curvatures:

$$\kappa_{Mean} = \frac{\kappa_1 + \kappa_2}{2}$$

$$\kappa_{Gaussian} = \kappa_1 \kappa_2 \tag{7}$$

where, κ_1 is the minimum of all curvature κ_i of curve C_i passing through at point M on surface S, and κ_2 is the maximum of them.

If the surface S is represented by the implicit function φ :

$$S = \{(x, y, z) \mid \varphi(x, y, z) = 0\}$$

then, the mean and Gaussian curvature of surface S at each point $M(x, y, z)$ can be calculated as follows:

$$\kappa_{Mean} = \frac{\varphi_{xx}(\varphi_y^2 + \varphi_z^2) + \varphi_{yy}(\varphi_x^2 + \varphi_z^2) + \varphi_{zz}(\varphi_x^2 + \varphi_y^2) - 2(\varphi_x\varphi_y\varphi_{xy} + \varphi_y\varphi_z\varphi_{yz} + \varphi_x\varphi_z\varphi_{xz})}{2(\varphi_x^2 + \varphi_y^2 + \varphi_z^2)^{3/2}}$$

(8)

$$\kappa_{Gaussian} = \frac{\varphi_x^2(\varphi_{yy}\varphi_{zz} - \varphi_{yz}^2) + \varphi_y^2(\varphi_{xx}\varphi_{zz} - \varphi_{xz}^2) + \varphi_z^2(\varphi_{xx}\varphi_{yy} - \varphi_{xy}^2) + 2[\varphi_x\varphi_y(\varphi_{xz}\varphi_{yz} - \varphi_{xy}\varphi_{zz}) + \varphi_y\varphi_z(\varphi_{xy}\varphi_{xz} - \varphi_{yz}\varphi_{xx}) + \varphi_x\varphi_z(\varphi_{xy}\varphi_{yz} - \varphi_{xz}\varphi_{yy})]}{(\varphi_x^2 + \varphi_y^2 + \varphi_z^2)^2}$$

Based on the mean and Gaussian curvature of a surface, the region of surface can be classified as follows:

- If the $\kappa_{Gaussian} > 0$ and $\kappa_{Mean} > 0$ then the curvature in any direction is positive and the surface region is a convex region.

- If the $\kappa_{Gaussian} > 0$ and $\kappa_{Mean} < 0$ then the curvature in any direction is negative and the surface region is a concave region.

- If the $\kappa_{Gaussian} < 0$, then the curvature in some direction is positive and others are negative (Hyperbolic patches).

Shape Classification Based on Eigenvalues

The eigenvalues of the Hessian matrix are used for locally describing of structure of an image (Figure0-3). Frangi at el [56] applied the elements of the Hessian matrix to classify the shape of a vessel by defining a multi-scale vesselness measure. They obtained the vesselness measure on the basis of all eigenvalues of the Hessian matrix aiming to develop a vessel enhancement filter.
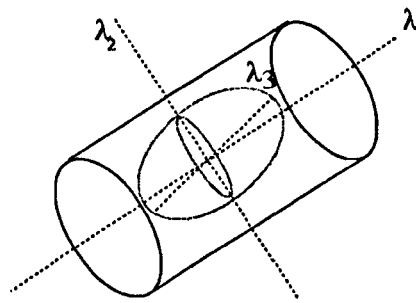
Figure 0-3: the second order ellipsoid which describes the local principal directions of curvatures

All three eigenvalues of the Hessian matrix have an important role to discriminate the local orientation pattern. Table 0-1 is a brief description and properties of eigenvalues of the Hessian matrix to detect the different structures [56].

Table 0-1 Shows the possible pattern in 3D based on eigenvalues of Hessian matrix λ_k (the eigenvalues are ordered as: $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$) in different size (H for high, L for low, N for noisy and +/- indicate the sign of eigenvalues.

λ_1	λ_2	λ_3	The orientation of pattern
N	N	N	Noisy, no preferred direction
L	L	H^-	Plate-Like structure (bright)
L	L	H^+	Plate-Like structure (dark)
L	H^-	H^-	Tubular structure (bright)
L	H^+	H^+	Tubular structure (dark)
H^-	H^-	H^-	Blob-like structure (bright)
H^+	H^+	H^+	Blob-like structure (dark)

Frangi proposed some ratios of eigenvalues to distinguish the shapes of different structures. The following is a briefly description of some ratios:

1). The following ratio is considered for deviating from a blob-like structure but cannot distinguish a plate-like structure from a line-like structure.

$$\left(\frac{|\lambda_1|}{\sqrt{|\lambda_2 * \lambda_3|}}\right) \tag{9}$$

For the blob-like structure, the ratio is maximum value and is zero whenever $\lambda_1 \cong 0$, or λ_1 and λ_2 tend to vanish.

2). The plate-like structures and the line-like structures are distinguished by the following ratio:

$$\left(\frac{|\lambda_2|}{|\lambda_1|}\right) \tag{10}$$

The two proposed ratios are gray-level invariant and remain constant under re-scaling and give just geometric information of the image.

APPENDIX B

3D Level Set Algorithm

$$\Phi_t = |\nabla\Phi|\left(\text{div}\left(\text{Fext}\,\frac{\nabla\Phi}{|\nabla\Phi|}\right) + v\text{Fext}\right) \quad \text{in } [0,\infty[\,\times R^3$$

$\Phi(x,y,z,0)$: The initial surface which is automatically selected

inside the pulmonary artery

κ: Mean curvature of evolving curve

F_{ext}: An external energy function composed of two energy

functions.

$$F_{ext} = F_{edge} \times F_{intensity}$$

v : $v > 0$ is a constant inside the curve to increase the propagation

speed.

Input:

- 3D original data
- 3D initial contour (In this algorithm the initial contour is a 3D surface inside the object which will be segmented).
- The number of iterations, which is an aggressive number but the algorithm terminates where the variance of the mean energy profile falls to zero.

Output:

- The segmentation results (binary image).

Step1: Load the row data (I)

Step2: Load the initial contour (Φ)

Step3: Finding the signed distance transform of initial contour (signed distance respect to boundary of initial surface).

Step4: Smooth the original image using Gaussian kernel.

Step5: Defining the edge detector in level set method

$$(F_{edge} = e^{-|\nabla I(x,y,z)|} \quad , \text{I: Original Image})$$

Step6: (Level Set Iteration)

For   Iteration = 1: number of level set iteration number

   Step6_1:

       Finding the narrow band of signed distance transform

$$idx = \left\{ (i,j,k) \middle| -\beta \le \Phi_{i,j,k}^{Iter} \le \beta \right\}$$

       (2β is width of narrow band)

Step6_2:

Finding the curvature of $\Phi^{\text{Iteration}}$ in idx (mean curvature)

$$\kappa_{mean} = \frac{(\Phi_{yy}+\Phi_{zz})\Phi_x^2 + (\Phi_{xx}+\Phi_{zz})\Phi_y^2 + (\Phi_{xx}+\Phi_{yy})\Phi_z^2 - 2\,\Phi_x\Phi_y\Phi_{xy} - 2\,\Phi_x\Phi_z\Phi_{xz} - 2\,\Phi_y\Phi_z\Phi_{yz}}{\left(\Phi_x^2+\Phi_y^2+\Phi_z^2\right)^{3/2}}$$

$$\Phi_x \approx \frac{\Phi_{i+1,j,k}-\Phi_{i-1,j,k}}{2h_x} \qquad \forall i \neq 1, m \quad , [m,n,k] = size(image)$$

$$\Phi_x \approx \frac{\Phi_{2,j,k}-\Phi_{0,j,k}}{2h_x} \quad , \qquad \Phi_{0,j,k} = \Phi_{1,j,k} \qquad i = 1$$

$$\Phi_x \approx \frac{\Phi_{m+1,j,k}-\Phi_{m-1,j,k}}{2h_x} \qquad \Phi_{m+1,j,k} = \Phi_{m,j,k} \qquad i = m$$

$$\Phi_{xx} \approx \frac{\Phi_{i+1,j,k}-2\Phi_{i,j,k}-\Phi_{i-1,j,k}}{h_x^2} \qquad \forall i \neq 1, m$$

$$\Phi_{xx} \approx \frac{\Phi_{2,j,k}-2\Phi_{i,j,k}-\Phi_{0,j,k}}{h_x^2} \qquad i = 1$$

$$\Phi_{xx} \approx \frac{\Phi_{m+1,j,k}-2\Phi_{1,j,k}-\Phi_{m-1,j,k}}{h_x^2} \qquad i = m$$

$$\Phi_{xy} \approx \frac{\Phi_{i+1,j+1,k}-\Phi_{i-1,j+1,k}-\Phi_{i+1,j-1,k}+\Phi_{i-1,j-1,k}}{4h_xh_y}$$

$$\Phi_{xz} \approx \frac{\Phi_{i+1,j,k+1}-\Phi_{i-1,j,k+1}-\Phi_{i+1,j,k-1}+\Phi_{i-1,j,k-1}}{4h_xh_z}$$

(The first and the second derivative in y and z direction are the same and h_x, h_y and h_z are the step size in x, y and z direction)

Step6_3:

Computing the value of the evolving coefficient $C_{evolving}$

$$C_{evolving} = |\nabla\Phi^{Iter}| \left(div \left(F_{ext}^{Iter} \frac{\nabla\Phi^{Iter}}{|\nabla\Phi^{Iter}|} \right) + v F_{ext}^{Iter} \right)$$

$$\delta t = \frac{min(h_x, h_y, h_z)}{A_{max}}$$

(δt, is computed based on CFL condition for stability of level set algorithm and A_{max} is maximum of elements of $C_{evolving}$)

Step6_5:   Evolving the surface in idx

$$\Phi^{Iter+1} = \Phi^{Iter} + \delta t C_{evolving}$$

Step6_6:

Reinitialize the Φ^{Iter+1}

(Using Susmand method algorithm)

End For

$$\text{Output} = \Phi^{Iter+1} \geq 0$$

Susmand Method Algorithm:

The following equations are used to reinitialize the level set function:

$$\varphi_t + S(\varphi_0)(|\nabla\varphi| - 1) = 0$$
$$\varphi(x,y,z,0) = \varphi_0(x,y,z)$$

φ_0: Initial level set function

$S(\varphi_0)$: Sign function

Discrete Algorithm:

For Iter = 1, L

  For i, j, k = 1, dim   dim=[m,n,p]=size(φ_0)

    Solve:

$$\varphi_{i,j,k}^{Iter+1} = \varphi_{i,j,k}^{Iter} - \Delta t S_\epsilon\left(\varphi_{0,i,j,k}\right)G(\varphi_{i,j,k}^{Iter})$$

    Given:

$$S_\epsilon\left(\varphi_{0,i,j,k}\right) = \frac{\varphi_{0,i,j,k}}{\sqrt{\varphi_{0,i,j,k}^2 + \epsilon^2}}$$

$G(\varphi_{0,i,j,k}^{Iter})$

$$= \begin{cases} \sqrt{\max((a^+)^2,(b^-)^2) + \max((c^+)^2,(d^-)^2) + \max((e^+)^2,(f^-)^2)} - 1 & \varphi_{0,i,j,k} > 0 \\ \sqrt{\max((a^-)^2,(b^+)^2) + \max((c^-)^2,(d^+)^2) + \max((e^-)^2,(f^+)^2)} - 1 & \varphi_{0,i,j,k} < 0 \\ 0 & otherwise \end{cases}$$

Where:

$$z^+ = \max(z,0)$$

$$z^- = \min(z, 0)$$

a, b : backward and forward space derivatives in x direction(two matrix)

c, d : backward and forward space derivatives in y direction(two matrix)

e, f : backward and forward space derivatives in z direction(two matrix)

End For

End For

Fast marching algorithm

-Initialization step

1. Alive points: Choose A as set points of all grid points (i_A, j_A, k_A) which represent the initial contour.
2. Narrow band: Select all grid points neighbours of A as set of narrow band or trial points
3. Far away points: Choose all other points as set of far away points and set:

$$T_{i,j,k} = \infty$$

    For all far away points.

-Marching forward step

1. Start the loop: Find the grid point $(i_{min}, j_{min}, k_{min})$ in narrow band where the arrival time T has the smallest value.
2. Remove the grid point $(i_{min}, j_{min}, k_{min})$ from narrow band points and add it to alive points
3. Tag all points $(i_{min}-1, j_{min}, k_{min})$, $(i_{min}+1, j_{min}, k_{min})$, $(i_{min}, j_{min}-1, k_{min})$, $(i_{min}, j_{min}+1, k_{min})$, $(i_{min}, j_{min}, k_{min}-1)$, $(i_{min}, j_{min}, k_{min}+1)$ which can be either in narrow band or in far away set points. If it is in far away set points, remove it from far away set points and add it to the set of narrow band points.
4. Re-compute the value of T at all neighbours by using the Eikonal equation (15) and find the largest possible solution for the quadratic equation (17).
5. Return and do the loop.