Computer-aided Diagnosis of Pulmonary Embolism in Opacified CT Images

A dissertation presented by:

Raphaël Sebbe

to the Department of Circuits Theory and Signal Processing
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Applied Sciences

Committee in charge:

Prof. Marcel Rémy, Chair and Professor at Faculté Polytechnique de Mons
Prof. Emmanuel Coche, Head of Thoracic Imaging at Cliniques Universitaires St-Luc, Université Catholique de Louvain
Prof. Joël De Coninck, Professor at Université de Mons Hainaut
Dr. Roger Lédée, Lecturer at Université d’Orléans
Prof. Paul Lybaert, Dean and Professor at Faculté Polytechnique de Mons
Prof. Benoît Macq, Professor at Université Catholique de Louvain
Prof. João Manuel R. S. Tavares, Professor at University of Porto
Prof. Jacques Trécat, Professor Emeritus at Faculté Polytechnique de Mons

Prof. Bernard Gosselin, Advisor at Faculté Polytechnique de Mons
Prof. Joël Hancq, Co-advisor at Faculté Polytechnique de Mons
Prof. Rachid Harba, Co-advisor at Université d’Orléans

November 2006

Faculté Polytechnique de Mons
Université d’Orléans
Abstract

Pulmonary embolism (PE) is an extremely common and highly lethal condition that is a leading cause of death in all age groups. Over the past 10 years, computed tomography (CT) scanners have gained acceptance as a minimally invasive method for diagnosing PE. In this manuscript, a framework for computer-aided diagnosis of PE in contrast-enhanced CT images is presented. It consists of a combination of a method for segmenting the pulmonary arteries (PA), emboli detection methods as well as a scheme for evaluating their performances. The segmentation of the PA serves one of the clot detection methods, and is carried out through a region growing method that makes use of a priori knowledge of vessel topology. Two different approaches for clot detection are proposed: the first one performs clot detection by analyzing the concavities in the segmentation of the pulmonary arterial tree. It works in a semi-automatic way and it enables the detection of thrombi in the larger sections of the PA. The second method does not make use of PA segmentation and is thus fully automatic, enabling detection of clots farther in the vessels. The combination of these methods provides a robust detection technique that can be used as a safeguard by radiologists, or even as preliminary computer-aided diagnosis (CAD) tool. The evaluation of the method is also discussed, and a scheme for measuring its performance is proposed, including a practical approach to making reference detection data, or ground truths, by radiologists.
Acknowledgments

The first question that comes to my mind right now is: how do you start a thesis? This is probably a complex question and I guess there are as many answers as there are PhDs... My beginnings were a bit original I believe, after a chapter in the speech software engineering field, I wanted to move to the image thing, which has always been a very important topic to me. The interview in the research center concluded in these terms, quite literally: «So you start a PhD, that way we have you by the lower parts of the body.», that I took as a compliment and as their way of signifying their esteem for me. Although it was not in my initial plan, I decided to accept the challenge and later realized that this turbulent takeoff was in no way related to the beauty of the scenery I could get from that place.

Working on a health-specific topic like this one is a great opportunity, as not only the technical aspect of image processing is interesting by itself, but it permits to contribute in a social way as well, possibly contributing to saving lives at a later time. Multiple contacts with the medical field revealed that, although the scientific approach is the same as in the engineering field, there are so many differences in language, expressions, and way of thinking that a learning curve is generally needed from both parties. Being aware of it and accepting it makes working together easier and richer towards achieving the same goals.

I would like to seize this opportunity to thank all the persons who contributed to make this possible, near or far. I am very thankful to Professor Bernard Gosselin’s trust in my developments, to his guidance and feedback throughout the project. I am indebted to Professor Emmanuel Coche, his kindness and efforts for providing me with the necessary medical background, data, and priceless discussions around the goals of this project. Professor Joël Hancq and Dominique Derestiat made sure everything was right around me at all the time, and I am very grateful to them as well. I am also obliged to Professor Benoît Macq who was a key person in making this project possible. My thanks further go to Professor Thierry Dutoit
who provided me wise advices at the important moments of this work. This project was funded by the Région wallonne and the European Social Funds, and I here express my gratitude to these organizations.

During the course of this thesis, I had the opportunity to make a 6-month period at the university of Orléans. This was a very rich experience, both scientifically and humanly. I would like to thank Professor Rachid Harba for hosting me at LESI, Roger Lédée and his numerous contributions and discussions around this work, and all the people at the lab for this warm welcome. I would also like to thank Doctor Richard of the CHRO for providing me images and advices in the medical domain. Orléans was also the place where I was awarded the prize of best poster and oral presentation at the Sciences en Sologne colloquium, as well as the Sadron prize of the Rotary Val-de-Loire. Needless to say these events were much appreciated as a testimony of interest and recognition of this work.

FPMs and the Multitel research center are great places to work at with great people to meet. On a daily basis, many persons contributed directly or indirectly and expressed interest in this work, or simply gave me their support. Having Julien in front of me was the source of passionate human and technical exchanges. The same holds naturally for Vincent, Jean-François, Sébastien, Devrim, Matei, and Céline shared the same doubts and joys as PhD students, and it was always a good moment to discuss all this together, not even mentioning their help in rereading this manuscript. Meetings, discussions, and common conferences with our UCL colleagues were also memorable moments. My thanks go also to the Multivision team at Multitel, to Christophe(s), Bertrand, Derek, Arnaud, Isabel, Sylvie and all the ones I just can’t cite here, for the many discussions, seminars and other events during this period.

Eventually, needless to say that my gratitude goes towards all my family and friends who supported me in this adventure and were (and still are) an inexhaustible source of inspiration, in particular but not limited to, Véronique and Antonin (and Victoria!), David, Samuel, Pierre, and Ana. To Christophe and Aurélie, with whom I shared this experience of going through a thesis, to Fanny, Ricardo, Michel, Patrice. To Mieke and Thierry for their unconditional friendship. To my brother and his wife Sarah, and of course to my parents and grandmother for their everlasting encouragements and assistance.
To my parents,
1 Introduction 1

2 Medical Definitions and Machinery 5
   2.1 Chest and Lung Anatomy 5
   2.2 Pulmonary Embolism 6
      2.2.1 Clot Characterization 8
      2.2.2 Load Scores 9
   2.3 Modalities for Diagnosing PE 10
      2.3.1 Pulmonary Angiography 10
      2.3.2 Nuclear Medicine 10
      2.3.3 Contrast-enhanced Computed Tomography 11
      2.3.4 Other Modalities 12
   2.4 Radiologist’s Expertise with CT 13
      2.4.1 Overview of the Examination Process 13
      2.4.2 Use of CTA and Protocol 13
      2.4.3 Locating Emboli 15
      2.4.4 Pitfalls 15
   2.5 Imaging Apparatus 16
      2.5.1 PACS 16
      2.5.2 DICOM Protocol 16
      2.5.3 CT Scanners 16

3 Related Work 21
   3.1 Segmentation of Vessels 21
      3.1.1 Filter-based, Vessel Enhancing Methods 22
      3.1.2 Region Growing Methods 23
      3.1.3 Knowledge-based Methods 29
   3.2 Active Contours: from Snakes to Level Set 30
   3.3 Detection of Clots 31
   3.4 Other Methods 32
<table>
<thead>
<tr>
<th>Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4.1 Mathematical Morphology</td>
</tr>
<tr>
<td>3.4.2 Distance Maps</td>
</tr>
<tr>
<td>3.4.3 Thin Plate Splines, and Associated Space Transform</td>
</tr>
<tr>
<td>3.4.4 Level Set and Fast Marching Methods</td>
</tr>
<tr>
<td>4 Methods</td>
</tr>
<tr>
<td>4.2 Level Sets and Fast Marching Experiments</td>
</tr>
<tr>
<td>4.3 Slice Marching Algorithm for Vessels Segmentation</td>
</tr>
<tr>
<td>4.4 Prior Knowledge Model Extension</td>
</tr>
<tr>
<td>4.5 Detection Algorithms</td>
</tr>
<tr>
<td>5 Experimental Results</td>
</tr>
</tbody>
</table>
# List of Figures

1.1 Examples of Typical, Opacified CT Slices of Lungs   .......... 2
1.2 Examples of Emboli Images in the PA               .......... 3

2.1 Functional Diagram of Cardiovascular System        .......... 6
2.2 Anatomy of the Lungs                              .......... 7
2.3 Diagram of Pulmonary Arteries Hierarchy           .......... 8
2.4 Examples of Ventilation/Perfusion Scans           .......... 12
2.5 Illustration of Chronic and Acute Clots           .......... 15
2.6 Illustration of CT Scanner Principle              .......... 17
2.7 Illustration of Windowing Process in CT Visualization  .......... 18
2.8 Examples of Windowing Configurations              .......... 19

3.1 Illustration of Agam’s Filters                     .......... 24
3.2 Principle of Zahlten’s method                      .......... 25
3.3 Principle of Stefancik and Sonka’s Method          .......... 27
3.4 Illustration of the Use of Ray Propagation to Detect Vessels .......... 27
3.5 Formulation of Interface Propagation               .......... 34
3.6 Huygens’ Principle of Wavefront Construction      .......... 36
3.7 Illustration of Fast Marching Update Scheme        .......... 38
3.8 Illustration of Fast Marching Algorithm            .......... 39
3.9 Definition of Speed Extension in the Level Set Method .......... 40
3.10 Illustration of Narrow Band Level Set             .......... 41
3.11 Illustration of Narrow Band Scheme                .......... 42
3.12 Limitations of Snake Method                       .......... 43

4.1 Initialization of Fast Marching with a Single Seed Point .......... 47
4.2 Initialization with a Polygonal Contour            .......... 48
4.3 Initialization with a Bézier curve                 .......... 49
4.4 Illustration Fast Marching Schemes against Dimensions .......... 51
4.5 Results of Segmentation with the Level Set Method   .......... 52
4.6 Concept of Vessel Slices                          .......... 54
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.7</td>
<td>Results of Segmentation with the Slice Marching Method</td>
<td>58</td>
</tr>
<tr>
<td>4.8</td>
<td>Examples of Image with Touching Vessels</td>
<td>59</td>
</tr>
<tr>
<td>4.9</td>
<td>Illustration of the Parametric Model of the Vessels</td>
<td>60</td>
</tr>
<tr>
<td>4.10</td>
<td>3-D Visualization of the Model</td>
<td>61</td>
</tr>
<tr>
<td>4.11</td>
<td>Illustration of an Automatic Co-registration Scheme of the Model</td>
<td>63</td>
</tr>
<tr>
<td>4.12</td>
<td>Illustration of a Fiducial Points-based Co-registration Scheme</td>
<td>65</td>
</tr>
<tr>
<td>4.13</td>
<td>Definition of Vessel Potential Functions</td>
<td>66</td>
</tr>
<tr>
<td>4.14</td>
<td>Computation of Initial Speed of Propagation</td>
<td>68</td>
</tr>
<tr>
<td>4.15</td>
<td>Diagram of Segmentation State Machine</td>
<td>68</td>
</tr>
<tr>
<td>4.16</td>
<td>Results of Model-based Segmentation</td>
<td>70</td>
</tr>
<tr>
<td>4.17</td>
<td>Diagram of Method $M_1$</td>
<td>72</td>
</tr>
<tr>
<td>4.18</td>
<td>Results of Method $M_1$</td>
<td>73</td>
</tr>
<tr>
<td>4.19</td>
<td>Diagram of Method $M_2$</td>
<td>74</td>
</tr>
<tr>
<td>4.20</td>
<td>Diagram of Lung Segmentation Method</td>
<td>76</td>
</tr>
<tr>
<td>4.21</td>
<td>Results of Method $M_2$</td>
<td>76</td>
</tr>
<tr>
<td>4.22</td>
<td>Comparison of Detection Methods</td>
<td>77</td>
</tr>
<tr>
<td>4.23</td>
<td>Illustration of the Process of Ground Truths Creation</td>
<td>78</td>
</tr>
<tr>
<td>4.24</td>
<td>Illustration of Difficult Ground Truth Cases</td>
<td>80</td>
</tr>
<tr>
<td>4.25</td>
<td>Comparison of the Results with the Ground Truth Reference</td>
<td>81</td>
</tr>
<tr>
<td>5.1</td>
<td>Medical Figures</td>
<td>84</td>
</tr>
<tr>
<td>5.2</td>
<td>Results of the Whole Processing Chain</td>
<td>87</td>
</tr>
<tr>
<td>5.3</td>
<td>Example of PA Segmentation Failure</td>
<td>88</td>
</tr>
<tr>
<td>5.4</td>
<td>Illustration of Massive Embolism</td>
<td>89</td>
</tr>
<tr>
<td>5.5</td>
<td>Diagram of Software Design</td>
<td>91</td>
</tr>
<tr>
<td>5.6</td>
<td>Illustration of iMed Studio Application</td>
<td>93</td>
</tr>
<tr>
<td>6.1</td>
<td>Visualization of Results with Augmented Display</td>
<td>97</td>
</tr>
</tbody>
</table>
List of Tables

2.1 Table of Hounsfield Units of Various Materials . . . . . . . . 18
4.1 List of Anatomical Definitions of Fiducial Points . . . . . . . 64
5.1 Table of Detection Results Figures . . . . . . . . . . . . . . . . 85
5.2 List of Per-method Processing Times . . . . . . . . . . . . . . 90
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAD</td>
<td>Computer-aided diagnosis.</td>
</tr>
<tr>
<td>CHRO</td>
<td>Centre hospitalier régional d’Orléans.</td>
</tr>
<tr>
<td>CT</td>
<td>Computed tomography.</td>
</tr>
<tr>
<td>CTA</td>
<td>Computed tomography angiography.</td>
</tr>
<tr>
<td>DICOM</td>
<td>Digital imaging and communications in medicine.</td>
</tr>
<tr>
<td>DT</td>
<td>Distance transform.</td>
</tr>
<tr>
<td>EGC</td>
<td>Extruded generalized cylinder.</td>
</tr>
<tr>
<td>FN</td>
<td>False negative.</td>
</tr>
<tr>
<td>FP</td>
<td>False positive.</td>
</tr>
<tr>
<td>GT</td>
<td>Ground truth.</td>
</tr>
<tr>
<td>GUI</td>
<td>Graphical user interface.</td>
</tr>
<tr>
<td>GVF</td>
<td>Gradient vector flow.</td>
</tr>
<tr>
<td>HCT</td>
<td>Helical computed tomography.</td>
</tr>
<tr>
<td>HU</td>
<td>Hounsfield units.</td>
</tr>
<tr>
<td>LESI</td>
<td>Laboratoire d’électronique, signaux et images de l’université d’Orléans.</td>
</tr>
<tr>
<td>MDCT</td>
<td>Multidetector computed tomography.</td>
</tr>
<tr>
<td>MRA</td>
<td>Magnetic resonance angiography.</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic resonance imaging.</td>
</tr>
<tr>
<td>MVC</td>
<td>Model-view-controller.</td>
</tr>
<tr>
<td>PA</td>
<td>Pulmonary arteries.</td>
</tr>
<tr>
<td>PACS</td>
<td>Picture archiving and communication systems.</td>
</tr>
<tr>
<td>PE</td>
<td>Pulmonary embolism.</td>
</tr>
<tr>
<td>PIOPED</td>
<td>Prospective investigation of pulmonary embolism diagnosis.</td>
</tr>
<tr>
<td>PV</td>
<td>Pulmonary vein(s).</td>
</tr>
<tr>
<td>RV</td>
<td>Right ventricle.</td>
</tr>
<tr>
<td>SE</td>
<td>Structuring element.</td>
</tr>
<tr>
<td>TCTS</td>
<td>Laboratoire de théorie des circuits et de traitement du signal de la Faculté Polytechnique de Mons.</td>
</tr>
<tr>
<td>TN</td>
<td>True negative.</td>
</tr>
<tr>
<td>TP</td>
<td>True positive.</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>TPS</td>
<td>Thin plate splines.</td>
</tr>
<tr>
<td>UCL</td>
<td>Université catholique de Louvain.</td>
</tr>
<tr>
<td>V/Q</td>
<td>Ventilation/Perfusion scan.</td>
</tr>
<tr>
<td>VTK</td>
<td>Visualization toolkit.</td>
</tr>
<tr>
<td>WL</td>
<td>Window level.</td>
</tr>
<tr>
<td>WW</td>
<td>Window width.</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Pulmonary embolism (PE) is an extremely common and highly lethal condition that is a leading cause of death in all age groups. PE is the third most common cause of death in hospitalized patients, with an estimated 0.24% annual death rate of the population. Its symptoms are often vague and its diagnosis is a major medical challenge, with 70% of missed diagnosis in people dying from PE in hospitals. If left untreated, approximately one third of patients who survive an initial PE subsequently die from a future embolic episode. Most patients succumb to PE within the first few hours of the event. However, when properly identified, an effective treatment consisting of anticoagulants or thrombolytics is administered that dramatically reduces the mortality rate of the disease.

The diagnosis of this pathology requires medical imaging and has historically been done by using either angiography or, more recently, nuclear medicine. The latest progress achieved in multislice computed tomography (CT) scanners, specifically in terms of spatial resolution and speed of image acquisition, has paved the way to using it as the preferred modality for PE diagnosis. In parallel, the continuous increase in computing power of processors and in memory capacity makes it possible nowadays to analyse very large datasets as the ones provided by CT scanners.

Over the past 10 years, CT has gained acceptance as a minimally invasive method for evaluating patients with suspicion of PE [1]. A CT scanner provides a number of images that are successive slices of an observed volume. Typical slices of a dataset are shown in Figure [1.1] where arteries appear in white color because of the contrast product that is injected to the patient. In that configuration, emboli are dark spots inside the white lumen of the arteries.

Typically, the diagnosis of PE is manually performed by radiologists on
Figure 1.1: Typical slices of a CT volume, that are taken at regular interval across the patient. Vessels appear white because they are opacified with a contrast product.

CT images. It is a time-consuming and error-prone process, in particular because of the huge amount of data and more specifically in the case of sub-segmental and peripheral clots, which are less visible, as shown in Figure 1.2. Indeed, a typical CT dataset that is used for PE diagnosis can have more than 600 slices, the size of the smallest visible volume being in the order of magnitude of one millimeter. The duration of that review process by radiologists, excluding acquisition time, is in the order of 5 to 10 minutes, depending on evidence of the clots. The duration of an exam may become a concern considering that there is currently a lack of radiologists, that is partly due to the availability of new modalities, the increasing complexity of existing ones, and the growing number of new applications of medical imaging (functional imaging, heart CT, etc.).

In that context, a computer aid can be provided whose main goal would be to decrease the time required to perform an exam, by acting either as a safeguard for radiologists or even better, if sufficiently robust and conservative, as a preliminary detection step in a computer-aided diagnosis (CAD) system.

The objective of this contribution is not to replace radiologist’s expertise by an automated tool, but instead to complement it. Identifying clots is a difficult process, and often different radiologists will not identify the same clots from a given dataset. The algorithm developed here should not be used as a way to decide when ambiguity exists among experts because the
knowledge that the method relies on is in all cases a lot more limited than the accumulated experience of radiologists. Instead, the goal of this work is to provide a pre- or post-detection step to help the radiologist improve his job. In essence, the method should be able to tell where potential regions of risk exist in the image, where clots are highly probable, and that the physician can in turn carefully inspect to validate the presence or absence of emboli.

This study takes the form of a framework, where images are collected, segmentation of pulmonary arteries is computed, clots are detected and then the detection results are evaluated against the same detection process made by the radiologist himself. Two methods tackling the problem of emboli detection are presented and later combined. The first one uses a segmentation of pulmonary arteries (PA) and is a semi-automatic process, while the second one works directly on image data and is a fully automatic process. An evaluation scheme is introduced and applied to these methods.

Segmentation of PA is a challenging task by itself, because of the particular shapes of thorax vessels, which are elongated and multi-scale structures. The main difficulty of PA segmentation is that the arteries are surrounded by other vessels, like the vena cava and the aorta, and, on top of that, the PA should not be confused with the pulmonary veins (PV), which share a lot of characteristics. Considering this peculiar application, where images are contrast-enhanced (opacified vessels), a novel, model-based active contour method is specifically developed and presented in this manuscript. This method works by growing a wave front from an initial seed point that is set inside the PA by a radiologist, and the originality is in the way this wave is guided by a higher level information, that is, the a priori knowledge of the topology of vessels.
The first method for detecting clots searches for holes in the segmentation of the PA relying on the expectation that an embolus creates a dark concavity in the lumen of the vessel. Because of its dependence on the segmentation, its performance decreases as we go farther in the vessels, or when vessels are completely obstructed, because the segmentation tends to be less accurate or even fails in those places.

For these reasons, a second detection method is developed which makes assumptions about clot densities and shapes to identify them. Due to the simple criteria that are used, this method leads to many false positives, most of which being outside of the lung volume. Lung volume segmentation is added to tackle that situation. Both clots detection methods are combined and provide a 3-D mask of emboli presence.

Eventually, a scheme for assessing the performance of PE detection methods is introduced. It focuses on the creation of ground truth data by radiologists to allow the comparison of the detection methods to a gold standard, in terms of false negative (non-detection) and false positive (false detection) rates; and these rates are then expressed as sensitivity and specificity. That scheme is used with our method and the results are discussed. This was possible thanks to our collaboration with the radiology departments at the UCL and CHRO hospitals, which provided both the images and the ground truth data.

As for the results, the segmentation method that we proposed could handle the majority of the collected datasets, excepted a few ones because of their poor image quality. Both clot detection methods found the majority of clots with a good sensibility, yet false alarms and non-detections were still present and their causes are discussed later in this text. The proposed validation scheme, combined with proper visualization, permits not only a quantitative but also a visual assessment of the methods, showing where the method works well and where it can still be improved.

The organization of this manuscript is as follows. Chapter 2 gives some background information in the form of medical definitions in the context of pulmonary embolism, its diagnosis, and the modality that is used, CT scanners. Chapter 3 covers the state of the art in image processing in the field of vessel segmentation and clot detection. Chapter 4 is where the proposed methods are exposed, first covering active contour experiments, a presentation of the slice marching method, the use of a model of prior knowledge, the detection methods, and finally the validation scheme. Chapter 5 presents and discusses the results of the methods applied to datasets and the conclusion is drawn in the last chapter.
Chapter 2

Medical Definitions and Machinery

This chapter provides medical background about pulmonary embolism, in terms of definition, symptoms, diagnosis, radiologist’s expertise, and treatment. The imaging devices and protocols that are used to establish the diagnosis are also discussed here.

2.1 Chest and Lung Anatomy

The chest is the part of the body that resides between the neck and the abdomen. It is supported by the ribcage and the spine, and it contains the heart and the lungs.

The lung is an essential organ of respiration in humans. Its function is to provide oxygen to the bloodstream, as well as to excrete carbon dioxide out of the bloodstream. The airways consist of the nose and/or the mouth, the pharynx, the larynx, the trachea, the bronchi, and the bronchioles. Deoxygenated blood travels from the heart, right ventricle, to the lungs through the PA for oxidation. The pulmonary veins (PV) carry oxygenated blood back from the lungs to the heart and then to the rest of the body. A simplified cardiovascular system is illustrated in Figure 2.1.

Airways and vessels in the lungs are dual structures in the sense that they share the same elongated and highly hierarchical shapes to maximize the contacts between air and blood flows. This interesting property of intertwinement is the basis of PE diagnosis with the nuclear medicine modality that is described in section 2.3.2. Airways and vessels meet at the alveoli, where the gas exchange occurs and blood becomes oxygenated. This is illustrated in Figure 2.2.
2.2 Pulmonary Embolism

PE is a very common condition that occurs when a blood clot, also known as a thrombus or embolus, causes a blockage in a pulmonary artery. It was first described by Dr. Virchow in the 19th Century, as the detachment of venous thrombi and their carrying by the blood flow to remote vessels. Its first radiographic description was made by Wharton and Pierson in 1922.

The clot, which is the result of blood coagulation, generally appears in the large veins of the legs and gets dislodged and migrates to the pulmonary arterial tree. The clot travels in the arteries of the lung until it reaches a vessel that is too small and that does not allow it to continue farther. Pulmonary infarct occurs when a part of an artery is not reached anymore by blood. Consequences of PE include a rise in the artery pressure due to the increased resistance to the bloodstream, and a decrease in cardiac output, which in turn can cause heart failure or even death. Symptoms of PE comprise sudden apparition of shortness of breath, rapid breathing, pain in the chest, and cough.

Untreated PE can be fatal and has a mortality rate of 30% that can be decreased to below 10% under rapid detection\[2\]. Once detected, a medical treatment exists, which generally consists of an anticoagulant and thrombolytics therapy\[3\] which is not without complications. Other treatments are also possible, like surgery, but they are uncommon. Death frequently occurs within the first hours after admission, and it arises when embolic obstruction significantly increases the vascular resistance, leading to right ventricle infarction and finally to circulatory collapse. Approximately 10%
2.2 Pulmonary Embolism

Figure 2.2: Lungs anatomy, left and right lungs are shown. The airways (trachea, bronchi, bronchioles) are in black and white, while the PA are in blue.
of patients who develop PE die within the first hour, and 30% die subsequently from recurrent embolism. The administration of an anticoagulant treatment decreases the mortality rate to less than 5%.

PE is the third most common cause of death in hospitals, with more than 650,000 cases per year in the US. Autopsy has shown that the diagnosis was missed in up to 70% of the occurrences, and approximately 100,000 deaths could be prevented with proper diagnosis and treatment. The incidence of PE varies internationally, from country to country, but that variation is likely due to the accuracy of diagnosis rather than the actual incidence of the disease itself.

The diagnosis of PE remains a major medical challenge, and imaging techniques have since played an important role because external symptoms are not sufficient to establish it. Ventilation/perfusion scintigraphy (nuclear medicine) has historically been the preferred imaging modality for evaluating patients with suspected PE, but this is currently shifting towards CT scanner.

### 2.2.1 Clot Characterization

Once clots have been localized in the PA, they can be characterized as acute or chronic. The distinction between both classes may be fuzzy in some cases, depending on image quality and interpretation. A clot is generally said to be acute if it is situated centrally within the vascular lumen. In that situation,
it may be the cause of vessel distension (enlargement). This type of clot
generates sudden pain, and is the one generally found in patients admitted
for suspicion of PE. On the contrary, a clot is qualified as chronic when it
is eccentric and contiguous to the vessel wall, when it reduces the arterial
diameter by more than 50%, or when re-canalization within the thrombus is
present. Chronic clots may decrease the vascular lumen of affected vessels
and may also calcify. Chronic PE typically causes pulmonary hypertension,
although this can also be the case for acute PE.

Clots are further characterized as central or peripheral, depending on
the location of the involved arterial branch. Central vascular zones include
the main pulmonary artery, the left and right main pulmonary arteries,
the anterior trunk, the right and left inter-lobar arteries, the left upper lobe
trunk, the right middle lobe artery, and the right and left lower lobe arteries.
Peripheral vascular zones include the segmental and subsegmental arteries
of the right upper lobe, the right middle lobe, the right lower lobe, the left
upper lobe including the culmen and the lingula, and the left lower lobe.

2.2.2 Load Scores

Various scoring systems have been introduced that permit measurement of
the presence, location, and degree of obstruction of arterial clots. Summaries
of the Walsh, Miller, Qanadli, and Mastora scores are given in [3], and the
Walsh score is expressed in the next section for illustration purpose. Their
use as a predictor of the failure of the right ventricle (RV) seems not possible,
but they are well suited at measuring the severity of PE and treatment
effectiveness. However, these scores are not much used in practice due to
the time it takes to manually evaluate them, which could be overcome with
the availability of a specific, automated tool.

The Walsh Score

The maximum score is 18 for both lungs. Different scores are given for filling
defects and obstructions and for anatomic locations. The following guide-
lines govern quantification of embolic abnormalities: (a) Abnormalities in a
single segmental PA receive a total score that does not exceed 1, regardless
of the type or number of abnormalities. (b) The total maximum score is 3
for abnormalities in a single upper lobar region, 2 for abnormalities in the
middle lobe or lingula, and 4 for abnormalities in the lower lobes. (c) Ob-
structions in central anatomic regions receive scores according to the vessel
involved. (d) If the total score for one lung is greater than 4 without consid-
ering filling defects in central regions, the central filling defects are ignored.
All filling defects in a single central region, whether single or multiple, receive a score of 3. (e) If a single vessel contains both a filling defect and an obstruction, only the obstruction is scored. (f) The sum of scores for all abnormalities in one lung may not exceed a value of 9. The maximum CT obstruction score is 18.

2.3 Modalities for Diagnosing PE

Medical imaging is widely used as a diagnostic means in patients admitted for suspicion of PE, because symptoms are nonspecific. Different techniques exist, with their particular advantages and weaknesses and a shift towards the CT modality is currently observed.

2.3.1 Pulmonary Angiography

Direct X-ray angiography of lungs is considered as the gold standard test for diagnosing PE. Its principle is to inject a radiocontrast agent directly into the target vessels, and to visualize them through X-ray projection. In the case of the PA, catheterisation of the right atrium of the heart is required. This procedure is thus invasive, requiring more expertise and support staff, and these are its main drawbacks against other modalities. Pulmonary angiography is thus the most reliable test for diagnosing PE but it is not used often because it is expensive, invasive, and not readily available in most hospitals.

2.3.2 Nuclear Medicine

Nuclear medicine, also known as scintigraphy, is a branch of medical imaging which permits the visualisation of radioactive substances through a gamma camera. In the context of PE diagnosis, a comparison between airways and arteries activity is made, hence the name, ventilation/perfusion (V/Q) scan. Emboli is possibly present when a defect is visible in the perfusion image that has no corresponding defect in the ventilation, meaning that air activity is not balanced with blood activity. Example images of the lungs with ventilation/perfusion scans are given in Figure 2.4. Technetium (Tc$^{99}$) is used for perfusion (blood imaging), while Krypton is used for ventilation, and images are acquired at various angles. In 1990, a study named PIOPED 1 showed that V/Q scans with normal findings generally excluded PE, but abnormal findings were not sufficient to diagnose PE.

Recently, a team from the LESI lab of the university of Orléans demonstrated that a quantification of the repartition of a radioactive tracer in each
pulmonary lobe was possible [5]. This is done by co-registering an anatomical atlas to V/Q scans and shows that, for a healthy group of patients and for a chosen isotope, a functional atlas could be evaluated. An abnormal decrease in ventilation or perfusion can then be evaluated and used to ease the diagnosis, still based on the comparison between V/Q scans. Figure 2.4 shows the co-registration of the atlas to the image, realized through an affine space transformation that is computed from Fourier descriptors.

The advantages of this technique are that it is not invasive and less irradiant\(^3\) and may be more suitable for patients that are allergic to iodinated agents (CT). Its disadvantage is that the obtained image is not accurate, and can only be used to determine regions of the lungs that are not correctly vascularized, non-obstructing clots remaining invisible. Moreover, the duration of the exam is in the order of 20 minutes, which is slower than other modalities.

### 2.3.3 Contrast-enhancedComputed Tomography

Computed tomography\(^4\) (CT) is a medical imaging modality that is used to generate a three-dimensional representation of the internals of the body (or other objects). It generally displays that information as a series of two-dimensional images as depicted in Figure 1.1. This modality and associated protocol are explained in more details in section 2.4.2.

Diagnosing PE with CT is achieved by first injecting a contrast agent into the patient, through an antecubital vein, that will eventually opacify the PA. This procedure is also known as CT angiography, or CTA. The arteries appear white where the bloodstream is present, and clots can thus be observed as darker spots inside the white mass. The timing and rate of injection are important parameters to obtain a correct opacification.

CT has recently gained acceptance as a minimally invasive method to diagnose PE [1], and has even become the preferred modality in that field. While less accurate than X-ray angiography, its non-invasive characteristic makes it much more practical and less expensive than the angiography. Clots can be seen directly with CT, unlike V/Q scans that provide a rough estimate of region activity, and the time to obtain the image of the patient is much reduced when compared to V/Q, especially considering the latest generation of scanners.

---

\(^3\)imaging devices based on X-rays and nuclear medicine expose the patient to radiations, which should be avoided when possible.

\(^4\)from the Greek *tomos* (slices) and *graphia* (describe)
2.3.4 Other Modalities

Electrocardiogram can generally provide clues for cases of large clots (more than 30% obstruction) because of the right heart strain, which in turn is to be confirmed with other modalities.

Magnetic Resonance Imaging, or MRI, was experimented in the context of PE detection but it provided a lesser sensitivity than CT. Moreover, MRI machines are expensive and cumbersome, but one advantage is that they are not irradiant, unlike X-ray based modalities. Like CT, MRI progresses rapidly in terms of hardware and software developments, and availability of appropriate contrast agents, and we may expect a shift from one to the other in the future.

Vascular Doppler ultrasound is another imaging device that can be used to detect abnormal bloodstream, through the principle of measuring its
speed by Doppler effect. It is not much used for PE because the results it provides generally carry less information than other modalities.

2.4 Radiologist’s Expertise with CT

This section gives an overview of the radiologist procedure for diagnosing PE by looking at how he handles that task from image acquisition to decision.

2.4.1 Overview of the Examination Process

The exam is performed as follows: a patient with suspicion of PE is admitted to the radiology department. He receives an intravenous injection of iodinated contrast agent, with specific rate and timing according to the PE detection protocol. CT scanning is performed, and the radiologist receives the images through a visualization workstation. After selecting the appropriate analysis window (as described in 2.5.3), the radiologist traverses the image volume, looking for clots.

The radiologist relies on his knowledge of the normal vascular and lung anatomy, its common variations and pitfalls acquired through education, as well as the implicit knowledge gained through experience to correctly identify clots. The vague definition of a dark spot inside a white volume may suffice for locating a majority of thrombi, but not always. In practice, a clot that is visible in only one slice of the volume is not enough to mark it as positive, and a minimum of two or three slices is required (spatial coherence), which tells why a back and forth review of adjacent images is often performed. Ambiguous situations where thorax tissues form a concavity in a vessel are generally well handled by the radiologist, although the variability of such cases has not been measured.

The distinction between pulmonary veins and arteries is often an important feature that a radiologist can use to help making the diagnosis. Often, the radiologist traces a vessel back to the point where he knows for sure whether it is a vein or artery, and assesses his conclusion. Interest for an automated tool that would permit the separation arteries and veins was expressed in the course of this research.

2.4.2 Use of CTA and Protocol

The procedure for image acquisition aiming PE diagnosis is expressed here. Scanner and opacification parameters determine what is to be observed in the image and its possible use for PE detection.
Spatial Resolution and Volume

Spatial resolution together with the field of view of the scanner specify the size of a voxel. That size is directly linked to the capacity of detecting small thrombi; the better the resolution, the smaller the clot can be. At the time of writing, the size of each slice expressed in number of voxels, and also known as the matrix size, can be up to 764x764 voxels in modern scanners and for a given field of view, and the spatial resolution between slices is a parameter. The ideal voxel size should be smaller than two millimeters, and isotropic voxel sizes are preferred as they permit easier interpretation. The total volume that is scanned should be sufficient by covering the regions where a risk exists, that is at least from the aortic arch to the lower pulmonary veins, totalling a minimum length of 12cm.

Acquisition Time

The patient must hold his breath for the duration of image acquisition. Indeed, motion that occurs during the time needed to acquire a single slice induces phantom shapes also known as motion artifacts, where structures appear fuzzy, and may confuse the physician. The maximum speed of the table depends on spatial resolution and technology, and is generally more than 5mm/sec corresponding to a duration of less than 30 seconds for an entire chest scan. Modern scanners are always faster than their predecessors.

Contrast Medium

A contrast medium is used to opacify the blood flow, making vessels appear white in the acquired image and making the thrombi appear as darker spots in the arteries. The volume of contrast medium depends on patient weight, and may vary in [70mL, 140mL] with iodine concentration of about 300mg/100mL. The injection rate is generally more than 5mL/s through a catheter in an ante-cubital vein. The rate of injection may increase in the last generations of multidetector computed tomography (MDCT), when acquisition time decreases in order to maintain sufficient iodine concentration within the vessels. The contrast peak arises after 10 to 25 seconds depending on the patient.

voxel stands for volume element, and it is the smallest representable volume in a 3-D regularly spaced grid that has an associated value. This is analogous to a pixel in a 2-D image, or to a sample in a 1-D signal.

smaller quantities are use with modern scanners
2.4.3 Locating Emboli

The iodinated bloodstream is either completely blocked by an embolus, or passes around it. Acute thrombi appear as a contrast defect in the image, and its shape generally defines a hole, or concavity, in the vessel. On the contrary, chronic clots are generally found on the edge of the vessel, with no concavity. This is represented in Figure 2.5.

![Illustration of chronic and acute clots](image.png)

Figure 2.5: Illustration of chronic and acute clots. The bloodstream is in white, and the clot is darker. The emboli on the left are chronic ones, remaining on the boundary of the vessel, with the latter one demonstrating re-canalization. The last two are acute clots with and without contacts with the boundary of the vessel.

The radiologist navigates on successive slices on the visualisation to locate such density defects, looking at one part of the image (left lung, then right) at a time because of the amount of information. When a defect is found in a slice, the immediate neighbor slices are also visualized to confirm or discard the clot.

2.4.4 Pitfalls

Unsatisfying exams represent 1 to 4% of the total number of exams and are generally due to improper opacification and motion artifacts for dyspneic patients. One frequent pitfall is simply to miss a clot because it is small and peripheral, and thus not obvious. Another one is to misinterpret the lymphatic tissue as an embolus that is located between a bronchus and an artery, against the vessel wall. The continuity of such structures in successive slices may help to determine its actual nature. Oblique or transversal arteries might generate hypo-densities in the image, as well as vessel bifurcations. This can generally be solved in most cases by longitudinal reconstruction, relatively to the vessel, or by looking at surrounding slices. Artifacts due to the high concentration of contrast product in the vena cava may also lead to apparent hypo-densities in the neighboring PA, which must not be confused that presents difficult or labored breathing.
with clots.

2.5 Imaging Apparatus

Imaging is performed with scanners that are connected to a network of computers known as PACS which communicate together through the DICOM protocol.

2.5.1 PACS

PACS is the acronym for picture archiving and communication systems, which actually are a network of computers serving the storage and visualisation of various digital medical imaging machines. The various computers in the PACS and the scanners communicate together through the DICOM protocol.

2.5.2 DICOM Protocol

DICOM stands for digital imaging and communications in medicine, is published on NEMA website. It is a comprehensive set of standards for handling (printing, transmitting, storing) medical images. It defines both a communication protocol based on TCP/IP and an image file format.

The acquisition of the images used in this work was made by exporting images from a PACS in the hospital to a CD-ROM in the DICOM file format. These images were then imported into our computer softwares.

2.5.3 CT Scanners

CT Scanners were commercially introduced for the first time in 1972 by Hounsfield, who was awarded a Nobel Prize in medicine in 1979. The principles behind a CT scanner is that an X-ray source placed in front of sensors (as illustrated in Figure 2.6), with both rotating around the volume to image. Attenuation of the X-ray beam is measured for each orientation of the source/sensors, and depends on the traversed materials. These projections are then combined together with a mathematical procedure of reconstruction, known as the inverse Radon transform, to retrieve the original radiographic densities at every location in the volume.

\[\text{at } \text{http://medical.nema.org. NEMA is the acronym for National Electrical Manufacturers Association, which is a U.S.-based association setting many standards for electrical products.}\]
Practically, the X-ray tube and detectors (sensors) of modern CT scanners spin continuously while the patient moves in the direction of the axis of rotation, hence these scanners are known as Helical CT (HCT) or Spiral CT scanners. Multiple detectors are used (MDCT), with up to 64 detector rows at the time of writing, leading to an accelerated acquisition time of just a few seconds for the entire chest.

These estimated radio-densities are computed on discrete volume intervals, or voxels, forming a regular 3-D grid, and are expressed in Hounsfield units (HU). Hounsfield units for various materials are expressed in Table 2.1. Interesting values for diagnosis purpose are in the $[-1000, +3000]$ range, and can be coded as 12-bit per voxel representation. In practice though, a 16-bit representation is chosen in the DICOM format for easier manipulation (8-bit multiple).

Computer displays have a limited 8-bit display capability, roughly corresponding to the human ability to discern nearby shades, and that 8-bit range cannot represent the entire gamut of the Hounsfield scale. For that reason, a process known as windowing is used, which consists of linearly mapping a subset of the Hounsfield scale to the 8-bit interval, $[0, 255]$, using two parameters, the window width (WW) and window level (WL), as illustrated in Figure 2.7. That pair of parameters is also known as a protocol to the radiologist, and must be chosen accordingly to the diagnosis task. For instance, protocols for studying the bones, the abdomen, and the brain respectively are (WL: 1500 HU, WW: 300 HU), (WW: 350 HU, WL: 40 HU), and (WW: 100 HU, WL: 50 HU). Example 8-bit images are given as they appear on a console to illustrate the meaning of that process (Figure 2.8).

Figure 2.6: CT Scanner principle. An X-ray tube rotates around a patient, describing a helicoidal path.
Table 2.1: Radio-densities in Hounsfield units of various materials.

<table>
<thead>
<tr>
<th>Materials</th>
<th>Value (HU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air</td>
<td>−1000</td>
</tr>
<tr>
<td>Water</td>
<td>0</td>
</tr>
<tr>
<td>Bones</td>
<td>+400</td>
</tr>
<tr>
<td>Titanium</td>
<td>+1000</td>
</tr>
<tr>
<td>Iron</td>
<td>∞</td>
</tr>
</tbody>
</table>

Figure 2.7: Windowing procedure consists of mapping the 12-bit gamut of the Hounsfield scale to 8-bit displays.
Figure 2.8: Same volume slice but seen with various window levels and widths. Top row, full gamut on the left (WL: -750 HU, WW: 4500 HU) and lung protocol on the right (WL: -600 HU, WW: 850 HU); bottom row, abdomen (WL: 40 HU, WW: 350 HU) and bones (WL: 300 HU, WW: 1500 HU) protocols.
Chapter 3

Related Work

The intuitive approach that was followed in the initial phase of this project was first to extract, or segment, the PA, and then to inspect them to detect blood clots. This chapter first covers the state of the art about vessels segmentation techniques, then it provides more details about the active contour methods and finally makes a review around the detection of clots. Other techniques, such as the level set algorithm and the thin plate spline transform, which are used by the methods introduced later in the text, are also briefly presented although they are not directly linked to the problem of PE detection.

3.1 Segmentation of Vessels

Segmentation is the process of dividing an image into parts. The segmentation of thorax vessels is a topic by itself, but what can be observed at this point is that it is similar to many other subjects as, for instance, airways segmentation, virtual endoscopy, and also in other medical domains, as retina scans, as well as segmenting routes in satellite images of Earth. A review of all the techniques used in these various domains is beyond the scope of this work, but we nevertheless did not limit ourselves to the study of vessels-only methods and considered relevant techniques in other fields as well.

As a starting point, Kirbas et al. [6] did a classification of techniques for vessels extraction in various modalities with many different classes, as did Felkel et al. [7] about the particular case of CTA datasets. Six main categories are defined by Kirbas that are (1.) pattern recognition techniques, (2.) model-based, (3.) tracking-based, (4.) artificial intelligence-based, (5.) neural network-based methods, and a (6.) miscellaneous category. These are
further refined in subcategories and methods that do not fit one category are cross-listed. Frangi discusses the use of prior knowledge for vessels analysis in his PhD dissertation [8]. The segmentation of vessels and airways is discussed in [9], chapters 14 and 16. We propose a simplified classification of the existing methods.

We define three main categories for vessels segmentation techniques, that are (1.) filter-based methods, which are low-level linear and non-linear filters aimed at enhancing elongated structures, (2.) region growing methods, which propagate information from a chosen seed point, and (3.) knowledge-based methods which make use of prior information. Methods in different categories may share common traits, such as being of multi-scale nature, being fully automatic or not, or their capacity of coping with 2-D or 3-D datasets.

3.1.1 Filter-based, Vessel Enhancing Methods

These methods consist of filtering an image, in the sense of signal processing, to obtain a response that is high for tube-like shapes, and low everywhere else. These filters can be linear or not, and they often share the same characteristic of being sensitive to a particular size and orientation of objects in the image. The majority of these methods can be summed up as defining an appropriate bank of filters, which individually select a corresponding vessel size and orientation and the outputs of these filters are combined or thresholded to highlight vessels. Another approach is to consider the Hessian matrix, whose eigen values, in the case of white vessels against a darker background, correspond to vessel direction (the biggest one) and transversal plane (the smaller ones), but this approach is sensitive to noise.

Poli et al. [10] defined a set of linear filters, whose kernels are a combination of Gaussian distributions. The idea is to enhance both vessels and their contours. This bank of filters can be chosen appropriately in terms of scales and orientations, and gives good results when the scale of the vessel is known in advance. When used to enhance structures of variable scales, the filter responses give blurred contours for vessels that do not match the expected size.

Zana et al. [11] used a bank of morphological operators (opening) with a linear structuring element (SE), combined with a preliminary noise removal filter, in the context of retinal angiographies. An opening filter allows the removal of shapes that are smaller than a chosen SE, as expressed in [12]. With a linear SE and for a chosen orientation, vessel parts that are larger than the SE are preserved, while others are not. Using a bank of such
filters with various orientations permits the retrieval of elongated structures, and an additional condition on curvature continuity enables the removal of elongated background shapes. This method is well suited to 2-D images when the size of the vessels is known beforehand, and when their size has a limited variability within the image.

Koller et al. [13] presented a method for multiscale detection of curvilinear structures. That method extends the developments of Canny [14] to multi-scale and to 3-D data by using Gaussian derivatives on a selected range of scales, and it provides as output the information of centerlines, local direction and scales of the structure.

More recently, Agam et al. [15] demonstrated a filter-based method that made a distinction between vessels, their junctions, and other non-linear structures (as nodules). They aimed at enhancing vessels and their junctions independently of the nodules, to permit the recognition of the latter ones. Their method favorably compares to other Hessian-based methods that generally only partly remove junctions and can only achieve a decrease in nodules intensity (and not completely remove them). This is illustrated in Figure 3.1. Similar techniques based on enhancing vessels through filters were presented in [16, 17].

What makes these approaches not easily applicable (with the exception of Agam’s method) in the specific case of thorax vessels is the high number of bifurcations. Indeed, the linearity hypothesis of these methods is not systematically met. Moreover, the multi-scale nature of the PA is broad; vessels diameters vary from a few centimeters to capillaries. Finally, filtering in 3-D is computationally inefficient for large kernel sizes. We note however that in the happening of an irregular opacification of the PA, which was generally not the case in our images, such filters may be used as a preprocessing step.

Other filtering techniques, not specifically for vessel detection, but for enhancing images are used. This is a preprocessing step for the methods. The idea is to remove unnecessary noise and small artifacts while preserving edges, thus performing a discontinuity preserving smoothing. This is typically achieved by using non-linear filters as the median filter, anisotropic diffusion or bilateral filtering [18].

3.1.2 Region Growing Methods

Region growing methods perform segmentation by starting from a seed point, and by continuously growing to reach the target shape. The first application of such a technique to the case of contrast-enhanced vessels in 3-D datasets is attributed to Zahlten et al. [19]. Their algorithm works by
Figure 3.1: Results of the method presented by Agam et al. (a) shows the original synthetic image, (b) shows the selectively enhanced junctions, (c) shows the enhanced vessels and their junctions, and (d) the enhanced nodules. Courtesy of Professor Agam, Illinois Institute of Technology. Copyright 2005 IEEE.
advancing voxel-wise in the volume from a starting seed through a wavefront mechanism, while reconstructing the vessels bifurcation graph simultaneously. That process is illustrated in 2-D in Figure 3.2. The image is first thresholded, and then 4- or 8-connectivity is used to compute the next wavefront from the current one, and that wavefront is searched for connected components of voxels, that define as many bifurcations. Problems may arise when the vessel opacification is non uniform, and the bifurcation graph is generally larger than the actual one due to over-detection of bifurcations for irregular vessel boundaries.

Figure 3.2: A seed point is placed (a), and a wavefront is constructed and evolves (b, c) by looking at 8-connected neighbors. A bifurcation is detected when the wavefront splits (d).

Brujinns [20] is interested in a more precise detection of branching. He went on and extended that concept introduced by Zahlten to double waves by considering a wavefront which has a depth of two voxels. Skeletonization (the process of finding the centerlines of vessels, or skeleton) is also experienced [21], as well as special handling of bifurcation through sphere probes, which permit tracing vessels. This latter development led to interesting results that seemed more geared towards interactive, user-driven segmentation of vessels.

In the same spirit, Quek and Kirbas presented a vessel extraction technique based on wave propagation for 2-D images in [22, 23], using a concept of cost minimization very similar to the fast marching algorithm which is a speed-based propagation scheme, discussed in section 3.4.4. A cost function is expressed, and a path that minimizes it is searched in the image. The implementation of the algorithm is discussed, and an analogy with a wave propagating in a refractive medium is given. Its main advantages over regular front propagation methods is its limited computational load, and the definition of the cost function is the important point which depends on the application. Also, Masutani et al. proposed a propagation scheme based on bounded space mathematical morphology in [24] where the wave.
front advances inside the vessel by performing successive dilations that are constrained by vessel walls.

Segmentation of 3-D datasets can be done directly by methods that inherently support it, or by using a 2-D method for individual slices and by combining them. That latter approach was used by Wan et al. \[25\] for extraction of vasculature in rats, where an analysis of the voxel neighborhood was made.

We also note that a fuzzy logic approach was experimented by Park et al. \[26\] in the context of thoracic airways segmentation. A seeded, 3-D region growing scheme is used, and a set of fuzzy membership rules are defined for connected regions of various gray-levels, to classify them into airways, airway walls, vessels, and lung tissues.

Image features may vary alongside the segmentation process. Yi et al. \[27\] used a region growing scheme based on statistics that are evaluated in local cubes that contain the active front. Once a cube is defined and its features are evaluated, region growing is performed, constrained to the interior of the cube. The bifurcation graph is computed by connected component labeling on the surface of the cube. Cube definition may be tricky (size, position), and rules are provided to determine the validity of cube.

Stefancik and Sonka presented a method for the segmentation of arterial tree in magnetic resonance angiography (MRA) in \[28\]. Their method combines a dynamic programming approach, where a cost, based on the depth transform, is given to the voxels that belong to vessels, and a conflict resolution scheme is applied to enable the separation of touching vessels. A graph of vessel segments is constructed, and a label is attributed to each segment by traversing the graph from seed points corresponding to different vessels. When a segment receives more than one label, the conflict resolution algorithm is invoked and the segment is split into two. That method is reported to be the only one that permits the separation of contrast enhanced veinal and arterial trees, but at the price of manually initializing the seed points of every vessel.

Another approach based on ray propagation was experimented by Tek et al. \[29\]. It works by casting rays from a point that is manually chosen inside a vessel, looking for a contrast transition with a mean shift algorithm in the traversed voxels. These detected voxels define the vessel contour, that is then post-processed to remove discontinuities (peaks) in its curvature (smooth it). This method has to be manually initialized for each 2-D slice.

1\textsuperscript{a} A depth transform of a contour image computes the distance to the nearest contour at each pixel location.
3.1 Segmentation of Vessels

Figure 3.3: Stefancik and Sonka’s Method. Dynamic programming is used to evaluate vessel centerlines and a bifurcation graph is constructed. Conflicts are detected by traversing the tree from multiple seed points, and they are resolved by a specific scheme. Image courtesy of Professor M. Sonka, University of Iowa, USA.

This is illustrated in Figure 3.4.

Dokládal et al. [30, 31] presented a method for extracting the dense network of liver vessels in 3-D. A two-pass algorithm is described that first establishes vessel skeletons (homotopic erosion) and then filters out the many false detections. Manual input is requested to initiate the segmentation.

Active contour methods can be considered as a subcategory of region growing methods that formalize the propagation scheme. Montagnat et al. [22] proposed a review of such deformable surfaces. The principle of active contour techniques is to evolve a given contour in space through a

Figure 3.4: A series of rays are cast from a seed point; contrast transitions permit the definition of a contour. The graph on the right depicts the luminance of the image along the ray RO, and the first transition to lower values determines the boundary of the vessel.
set of rules, snakes and level set being algorithms based on that concept. They are presented in more details in following sections \[3.2\] and \[3.4.4\]. Malladi demonstrated the use of the level set method for shape modeling in \[33, 34, 35\], and in particular to segment an angiogram image, expressing its superior ability of topology adaptation and insensitivity to initial conditions.

CURVES, which is an implementation of the level set method that uses vessels centerlines to define the initial propagating surface, is presented in \[36\] and it is essentially applied to the MRA modality. It provides vessel radii as by-products that can serve diagnosis purposes. The difficulty of assessing the performance of the method is invoked because of the level of complexity of highly branching structures.

The application of virtual endoscopy (VE), which is the process of defining a path for a virtual camera that evolves into the body to inspect surfaces of organs, presents a similar segmentation interest. Chen et al. \[37\] proposed a region growing method that follows the vessel skeleton, itself defined as a 26-connected set of voxels maximizing the distance to the boundary. Indeed, the VE application enforces paths that are central to the structure because it is meant to be the path of a camera.

Deschamps et al. \[38, 39, 40\] presented a method based on the fast marching algorithm for 3-D vessel segmentation as well as for VE, adding the concept of freezing to contour evolution once the vessel boundaries are reached. This particular feature is a necessary one if the contour has to stop on object boundaries, as the fast marching algorithm uses a strictly positive speed of propagation, meaning that the evolution normally never stops. We used a similar freezing scheme in our slice marching method. Moreover, the fast marching algorithm has the advantage of being fast, which is a requirement in this development, in addition of being flexible through the definition of the speed of propagation.

After experimenting with the level set and fast marching methods, we first proposed \[41, 42, 43, 44\] the slice marching algorithm which is similar to Quek et al. \[22, 23\] and Deschamps et al. \[38, 39\], being based on fast marching with freezing, but advancing as slices inside the vessels. It is presented in section \[4.3\]. Our method can be considered as a unifying wave propagation technique in the same spirit as the methods from Zahlten et al. \[19\] and Bruijns \[21, 20\] but based on contour evolution theory, where slices can be thought of as waves of arbitrary depths. Still, this approach was plagued by the problem of joint vessels, incorrectly segmenting the vena cava and aorta and that problem was later addressed by adding the use of prior knowledge.
3.1.3 Knowledge-based Methods

Methods in this category make use of additional information that can be provided as an *a priori* model of knowledge to perform the segmentation task. The nature of models may vary, some of them are relatively low-level, making hypothesis of local properties of the structure, while others introduce a higher level of abstraction, modeling not only the local properties of the shape, but also the relationship between shapes, their topology, and their intrinsic variability.

In the context of the segmentation of 2-D objects, Tagare \[45\] proposed a template matching method, that demonstrates robustness to noise and to initial conditions. Hall et al. \[46\] presented a method for the 3-D reconstruction of vascular network of the brain, in the difficult context of digital subtraction angiograms (DSA). The reconstruction problem is different from the segmentation one, but the model still presents an interest for segmentation. In that modality, two images are available, a front and a side projections, with no occlusion of contrast-enhanced vessels. 3-D *a priori* models of the vessels are used to disambiguate the limited information and reconstruct the 3-D structures. The distinction of weak and strong *a priori* information is made and vasculature variability among persons is discussed and handled by multiplying models. Each vasculature is modeled as a graph whose nodes represent bifurcation points and whose edges represent vessel segments. The graph thus models the topology and geometric attributes (individual vessel sections) are added as labels to the graph.

Interest in modeling vessels was expressed by O’Donnell et al. in \[47\] where they presented a method that is based on their *extruded generalized cylinder* (EGC) model, \[48\]. The most common form of generalized cylinders (GC) is discussed in \[49\]. The key idea is to model tubular structures as surfaces made of a possibly varying cross section that is swept, or extruded, along a path. O’Donnell discussed their representation of the path, which is a 3-D B-Spline curve, and reported successful fitting to a previously segmented aorta.

Extraction of vessels that are not contrast enhanced is a complex problem that was studied by Kitasaka et al. \[50\]. CT values alone do not permit the delineation of vessels because of the mediastinum\(^2\), the proximity of vessels, and a lack of contrast. For these reasons, they used a model-fitting approach, where vessels are represented by their medial axis as B-Splines and are fitted to a likelihood image that is a distance map from visible vessel

\(^2\)Membraneous partition between both lungs, which has similar X-ray densities as other soft tissues.
edges. Automatic fitting of the model is made through a two-step algorithm (a linear transformation is applied to the model followed by a fine-tuning non-linear adaptation).

We have extended our slice marching algorithm [51, 52, 53] by incorporating a model of knowledge where vessels are represented as their centerline and their radius. This model is quite similar to [50] and to a lesser extent to [47], and it is used to disambiguate the case of joint vessels [52], as information in the image alone is not sufficient. Vessels being contrast enhanced in this application, a method that accounts for it is worth developing, as vessels boundaries are most of the time clearly visible (this was not the case for Kitasaka) and one want to give more reliance to the image data. Conversely to O’Donnel [47] and Kitasaka [50], we chose to use Bézier curves instead of B-Splines, as well as a constant section per vessel branch. Also, the co-registration with the image data is performed manually through a small set of fiducial points that must be located for each new patient, and that set then defines the transformation that applies the model to the image. This is discussed in more details in section 4.10.

### 3.2 Active Contours: from Snakes to Level Set

Active contours have their origins in the elastic models introduced by Burr in 1981 [54] and became an important topic in the image processing field and segmentation since the publication of the work of Kass, Witkin, and Terzopoulos in 1988 [55, 56], where they are known as the snakes. In that discussion, contours are defined as a parametric curve, open or closed, that evolves under the action of a force. That force moves and deforms the curve in the image from a starting position, which must generally be chosen close to the expected final position in the image. That model is also known as a physical model because of the analogy with physics (force, elasticity).

An energy functional is defined that is made of three terms and that has to be minimized to find the desired shape. The first term is about the internal energy of the contour, and expresses elongation energy, or tension (first order derivative) and bending energy, related to curvature (second order derivative) and both serve to regulate the shape coherence by avoiding point collapse and sharp corners. The second term is the external energy, and depends on the data in the image. It may be a function of image intensity or the gradient (if image contours are of interest). A special force term named the gradient vector flow (GVF) was defined in [57] to better attract the contour to image edges, even in the case of highly concave shapes.

The last term is about the introduction of global constraints on the curve
and it permits the use of a priori knowledge about the shape for instance, or the evolution itself. A well known example of this is the balloon energy introduced by Cohen in 1991 [58]. Numerous articles were published in that field and a thorough review is out of the scope of this thesis. For further information, the reader is invited to look into review papers such as [59], by McInerney and Terzopoulos, for medical imaging applications of deformable models, [60] deformable surfaces by Montagnat and Delingette or a review of deformable templates by Jain et al. [61].

Snakes cannot be used in this particular application. Indeed, these models do not handle the case of topology adaptation which is required because of the presence of clots in the PA, which are holes in the expected segmented volume. Moreover, the extension to 3-D datasets is cumbersome. Other active contours methods have that property of topology adaptation, such as the level set and fast marching methods presented by Osher and Sethian [61]. Image segmentation is only one of their possible applications. They are presented in more details in section 3.4.4 and more information can be found in the publications of Sethian [62], Osher and Fedkiw [63], and Osher and Paragios [64], or Sethian website [65]. These methods have been studied in this context because of their interesting properties and also because they had already been successfully used recently for vessel segmentation by Deschamps in his thesis [66], of which chapter 2 provides a very good state of the art.

3.3 Detection of Clots

Methods for detection of emboli from CT images have not received much attention, probably because of the recent advent and switch towards this modality. Masutani et al. recently presented a method for emboli detection based on the segmentation of the PA [67, 68, 69]. Once the PA has been localized, features are evaluated to determine clot candidates. Comparison to a reference is given, and sensitivity of up to 100% is reported, but the realization of ground truth and its comparison to the detection results is not explained. Results are very promising and the main limitation seems to be that the detection is tightly linked to the quality of the underlying segmentation of the PA.

Another approach to emboli detection is to not use any prior segmentation of the PA. Such a method based on tobogganing was recently presented by Liang et al. in [70]. Tobogganing [71], without going into too many de-
tails, is a technique for enhancing contrast, that replaces the value of each pixel (or voxel in 3-D) by the value of the one that is the local minimum in contrast. That local minimum is reached by following the path of decreasing contrast (amplitude of the gradient) values. The obtained image has sharper object boundaries that is more suitable to contrast segmentation. Liang’s method gives very good sensitivity of up to 95.7 % and its main reported drawback seems to be the high number of false positives, but this is not much discussed, neither is the way to compute such a statistics.

As we found out, segmentation of lung volume can dramatically reduce the false positive rate of detection methods. Segmentation of the lung volume was studied by van Ginneken et al. for radiographs [72], and we note that the problem is easier in CT scans, because of the access to unprojected raw data. Interest in emboli detection based on machine learning methods and classification was recently expressed [73], but this is not studied as part of this work, although it might be a way to further refine PE detection by decreasing the false alarm rate.

We first tested a detection method that relies on a prior segmentation of the PA that is in many ways similar to the work of Masutani et al., but we saw that this method has the drawback of missing clots that are in smaller vessels (farther in the hierarchy). We then introduced another method that works directly on the image data in the same spirit as Liang et al. but based on mathematical morphology instead of tobogganing. These methods are described in section 4.5.

3.4 Other Methods

The methods presented in this section are used later as building blocks of the proposed algorithm, but are not themselves directly linked to the problem of thrombi detection.

3.4.1 Mathematical Morphology

Mathematical morphology is widely used in this work. It was first proposed by Serra in the sixties, and an overview is given in [12]. It can be used either on binary or grayscale images, through the usual erosion and dilation operators that are combined in many different ways. It can be used either as a means for finding concavities through the closing operator, as well as a way to enforce spatial coherence or filtering out small regions that are below a threshold. This is expressed in the corresponding sections where the algorithms are presented.
In this work, a specific implementation was made in 3-D, using a linear SE (a 3-D cross) for performance reasons. Recent extensions such as SEs of variable sizes were also considered [74], that involves distance map calculation in the process, but computational performance was found to be an issue because of the size of the datasets.

### 3.4.2 Distance Maps

Distance map algorithms (or distance transform, DT) are needed in various phases in this work. The point is to compute a distance from a shape to any pixel or voxel of a domain. The algorithm presented by Cuisenaire in [75] were investigated, but the fast marching algorithm was used instead as it provides a good approximation of the solution to the Eikonal equation 4.1.

### 3.4.3 Thin Plate Splines, and Associated Space Transform

The thin plate splines (TPS) were introduced by Bookstein in 1989 [76], and permit the construction of a continuous function that interpolates a set of known points. More precisely, they permit the evaluation of a continuous function $f(\bar{X}) : \mathbb{R}^n \rightarrow \mathbb{R}$ ($n$ being 2 or 3) given the value of the function for a finite set of points $\{\bar{X}_i\}$. They are a generalization of the cubic spline for 2-D and 3-D cases. Their name has its origin in the physical analogy of bending a thin sheet of metal to make it pass through a set of points whose elevation is known, the solution being the function that passes through the points, minimizing a bending energy functional.

The TPS algorithm can be used to evaluate a continuous space transformation that is expressed as a correspondence between two sets of points. To do this in 3-D, the function that one wants to evaluate is the displacement between both coordinate systems, knowing the displacement in x, y, z for a limited set of points. The algorithm is applied 3 times, one for each dimension ($f_x$, $f_y$, and $f_z$ are evaluated).

This algorithm is used in the co-registration of the model to the image, which is discussed in section 4.10.

### 3.4.4 Level Set and Fast Marching Methods

Level set methods were introduced by Osher and Sethian in 1988, and fast marching methods by Sethian in 1996. [61, 62]. They follow a discussion of solving the problem of propagating fronts using a formulation developed for gas dynamics. That formulation provides a more robust and more stable solution than the discrete representation of contours (snakes), and comes
with intrinsic support for topology adaptation and 3-D (or N-D) representation. A brief overview of these methods is given here; for a more complete theoretical description of these methods, the reader is invited to refer to [62].

**Formulations of Interface Propagation**

Consider an initial contour $\Gamma(0)$, or boundary, or front, at time $0$. It is either a curve in 2-D or a surface in 3-D. That contour $\Gamma(t)$ moves in time $t$ in a direction perpendicular to itself with a speed $F$ that is known (Figure 3.5), and whose positive sign is defined as going outwards. $F$ may depend on local properties of the contour, such as its curvature, on global properties of the front that can be expressed as functionals of the contour, and on independent or external properties such as values derived from an underlying image. Given $F$ and the initial position of the front, the goal is to track its propagation.

Two formulations for this problem are given, the *boundary value formulation* and the *initial value formulation*. In the former, $F$ is assumed to be strictly positive and, because of this hypothesis, the arrival time of the contour can be expressed as a function of space, $T(\bar{X})$, $\bar{X}$ being the space variable (3-D or other). The equation governing the motion of the contour is

$$|\nabla T(\bar{X})| F = 1, \quad \text{with } T(\bar{X}) = 0 \quad \forall \bar{X} \in \Gamma(0),$$  \quad (3.1)
and is known as the *Eikonal* equation when $F$ depends only on position. An intuitive interpretation of this equation is that $T$ grows by $1/F$ when advancing a distance of one space unit. Iso-$T$ lines are thus more dense where $F$ is high, and vice versa.

Conversely, suppose that the speed $F$ is neither strictly positive nor negative. The contour may then pass several times over the same point $\bar{X}$ and another formulation is needed. One possible way of handling this situation is given in the initial value formulation, which embeds the initial position of the front, $\Gamma(0)$, as the zero level set of a higher-dimensional function $\phi$. That function $\phi$, which depends on time $t$, is linked to the contour as the location of the front is given at any time by the zero level set of $\phi$, $\Gamma(t) \triangleq \{ \bar{x}(t) \mid \phi(\bar{x}(t), t) = 0 \}, \quad (3.2)$

$\bar{x}(t)$ being the path over time of a point on the contour.

The chain rule gives this expression,

$$
\phi_t + \nabla \phi(\bar{x}(t), t) \cdot \bar{x}'(t) = 0,
$$

and, considering that $\bar{x}'(t) \cdot \bar{n} = F$, where $\bar{n} = \nabla \phi / |\nabla \phi|$, it provides the level set equation,

$$
\phi_t + F |\nabla \phi| = 0,
$$

with $\phi(\bar{X}, 0)$, or $\phi_0$ that is known.\(^4\) The curvature $\kappa$ of the front is easily derived as the divergence of the unit normal, given here in 2-D,

$$
\kappa = \begin{cases} 
\nabla \cdot \frac{\nabla \phi}{|\nabla \phi|} = \frac{\phi_{xx} \phi_y^2 - 2\phi_x \phi_y \phi_{xy} + \phi_{yy} \phi_x^2}{(\phi_x^2 + \phi_y^2)^{3/2}} \\
\nabla \cdot \frac{\nabla T}{|\nabla T|} = \frac{T_{xx} T_y^2 - 2T_x T_y T_{xy} + T_{yy} T_x^2}{(T_x^2 + T_y^2)^{3/2}}
\end{cases}
$$

\(3.5\)

**Mathematical and Numerical Schemes**

When tracking an interface over time, one can observe the formation of singularities. This is illustrated by an example. Considering a cosine curve, which is continuous and indefinitely differentiable, and making it evolve with a constant speed $F = 1$, then one can observe the apparition of a corner in a finite time (Figure 3.6). The contour is not anymore differentiable at that point, and a weak solution has to be used. A weak solution is one that

\(\text{\footnotesize{\textsuperscript{4}one possible expression of } } \phi_0 \text{ \textsuperscript{is} the signed distance map to the initial contour. This is discussed later in this section.}}\)
Figure 3.6: Huygens construction principle (left) from a sinusoid curve propagating at a constant speed $F = 1$, that sinusoid develops a corner in a finite time (right)

satisfies an integral formulation of the propagation equation. The weak solution that one wants is the one corresponding to the shortest distance, that is, the solution obtained by Huygens construction principle as illustrated in Figure 3.6. This is equivalent to the notion of entropy condition, posed by Sethian, who observed that this solution is the one coinciding with the limit of a viscous solution (one that prevents the apparition of a corner) when the viscous term tends to zero.

Sethian then sets up the basis to build a numerical scheme that will converge to the entropy solution. The path leading to that numerical scheme is beyond the scope of this summary, and we only provide the resulting schemes of the first order as well as a possible interpretation for both formulations of the problem of interface propagation.

The first order numerical scheme for the boundary value formulation on a regular grid is given,

$$
\begin{array}{c}
\max(D^{-x}_{ijk}T, 0)^2 + \min(D^{+x}_{ijk}T, 0)^2 \\
+ \max(D^{-y}_{ijk}T, 0)^2 + \min(D^{+y}_{ijk}T, 0)^2 \\
+ \max(D^{-z}_{ijk}T, 0)^2 + \min(D^{+z}_{ijk}T, 0)^2 \\
\end{array}
= \frac{1}{F_{ijk}}
$$

(3.6)

where $D^{+x}_{ijk}T$ and $D^{-x}_{ijk}T$ respectively denote the forward and backward difference operators, that is

5Sethian clearly states that this choice is made because the front is assumed to be an interface between two regions (think of a flame/air interface) and that one is interested in measuring the progression of one region into the other one.
3.4 Other Methods

\begin{align}
D^{+}_{ijk}T &= T(x_i+h,y_j,z_k) - T(x_i,y_j,z_k), \\
D^{-}_{ijk}T &= T(x_i,y_j,z_k) - T(x_i-h,y_j,z_k),
\end{align}

(3.7)

and \(F_{ijk}\) is the value of the speed of propagation at the discrete position \((x_i, y_j, z_k)\).

The first order numerical scheme for the initial value formulation is here,

\begin{equation}
\phi^{n+1}_{ijk} = \phi^n_{ijk} - \Delta t \left[ \max(F_{ijk}, 0) \nabla^+ + \min(F_{ijk}, 0) \nabla^- \right],
\end{equation}

(3.8)

where

\begin{align}
\nabla^+ &= \left[ \max(D^{-x}_{ijk} \phi, 0)^2 + \min(D^{+x}_{ijk} \phi, 0)^2 + \right. \\
&\left. \max(D^{-y}_{ijk} \phi, 0)^2 + \min(D^{+y}_{ijk} \phi, 0)^2 + \right. \\
&\left. \max(D^{-z}_{ijk} \phi, 0)^2 + \min(D^{+z}_{ijk} \phi, 0)^2 \right]^{1/2},
\end{align}

(3.9)

\begin{align}
\nabla^- &= \left[ \max(D^{+x}_{ijk} \phi, 0)^2 + \min(D^{-x}_{ijk} \phi, 0)^2 + \right. \\
&\left. \max(D^{+y}_{ijk} \phi, 0)^2 + \min(D^{-y}_{ijk} \phi, 0)^2 + \right. \\
&\left. \max(D^{+z}_{ijk} \phi, 0)^2 + \min(D^{-z}_{ijk} \phi, 0)^2 \right]^{1/2},
\end{align}

(3.10)

with \(\Delta t\) chosen according to the CFL condition

\begin{equation}
\max_{\Omega} F \Delta T < \Delta x,
\end{equation}

(3.11)

and \(\Omega\) being the domain of the current front.

An intuitive interpretation of these schemes is that the solution must be computed in a specific, upwind direction, from the known to the unknown. This is done for instance in equation \(3.6\) by discarding negative gradients when going from left to right (and its dual case), thus respecting the domain of dependence of the equation, that is, from smaller time values to bigger ones. This is further illustrated in the following two sections, where these schemes are used to solve their corresponding equations.

The Fast Marching Algorithm

We want to solve the boundary value formulation of a propagating interface on a discrete grid. One approach could be to iteratively apply equation \(3.6\) computing new values for \(T\) based on the values of the previous iteration. In 2-D, one would replace the value of \(T_{i,j}\) in Figure \(3.7\) by a value computed from its neighbors. This could hopefully converge in at least \(N\) iterations for a \(N \times N\) domain.

\(^{6}\)named after their originators, Courant, Friedrichs, and Lewy. It is a condition that ties space and time steps, which must be met for the stability of a numerical scheme solving a hyperbolic partial differential equation.
Figure 3.7: Fast marching update scheme: $T_{i,j}$ value is updated from its neighbors by applying the numerical scheme of Equation 3.6.

The principal idea behind the fast marching algorithm is to use the causality relationship from equation \ref{eq:causality}. This is illustrated in 2-D in Figure 3.8. By considering an initial grid point where $T$ is known and by recalling that $F$ is known at every point, we can update its four neighbors by applying equation \ref{eq:causality}. If we then select the one with the smallest value of $T$ out of these four points and if we state that its value is known, we can move on, updating its neighbors, and successively, the entire domain. That statement is true and can be shown by considering that no path from other points can lead to a value smaller than this one.

Implementation-wise, the algorithm defines three sets of points, the known, the trial, and the far sets, which respectively correspond to the black, gray, and white colors in Figure 3.8. The known set contains the points whose $T$ value is definitely fixed. The trial set contains the points whose $T$ solution has been evaluated, but that may later change when one of its neighbor becomes known. The far set contains all the points that have not yet been visited. The trial set is represented as a priority queue (full binary tree) to get an $O(1)$ complexity for accessing the smallest $T$, and a complexity of $O(\log M)$ for updating it, $M$ being the cardinality of the set.

The initialization of this algorithm is done by providing an initial known set. Various approaches are possible, depending on the nature of the contour. Some of them are expressed in section 4.2.1.

The Level Set Algorithm

The level set algorithm consists of iteratively applying equation \ref{eq:level_set}. To do this, one needs to know an initial value of $\phi$, $\phi_0$, as well as the expression of $\phi$ for unreached points is set to $\infty$, which is equivalent to not using their corresponding terms in Equation 3.6. See section 4.2.3 for more details on how to update the value of $T$.
Figure 3.8: Fast marching propagation scheme. The value of $T$ is propagated from a point where it is known (a) to its neighbors (b). The smallest value of $T$ is marked as known (c) and propagates (d). And the scheme continues, (e), (f).
the speed of propagation $F$ on the entire domain of definition of $\phi$, which generally does not exist$^8$

A straightforward method for providing $\phi_0$ is to compute the signed distance to $\Gamma(0)$. This leads to $\Gamma(0)$ being the zero-level set of $\phi_0$ by construction. The signed distance function can be efficiently evaluated with a two-pass fast marching algorithm.

$F_{\text{ext}}$ is defined as the extension of speed $F$ to the entire domain of $\phi$, from the value of $F$ on the front. Although $F$ may be evaluated in some cases on the entire domain at any time, depending on the application, that value may not be the one to use to evolve $\phi$. In other applications, $F$ may simply be unknown outside of $\Gamma$, but evolving $\phi$ is still needed. For these reasons, just as $\Gamma$ is embedded in a higher dimensional function $\phi$, $F$ is embedded in $F_{\text{ext}}$. This is illustrated in Figure 3.9. Sethian proposes an extension $F_{\text{ext}}$ that has the interesting property that if $\phi(t)$ is the signed distance to the front $\Gamma(t)$, then $\phi(t+\Delta t)$ will still be the signed distance to $\Gamma(t+\Delta t)$. This speed of propagation is obtained by solving this equation$^9$

$$\nabla \phi^{\text{temp}} \cdot \nabla F_{\text{ext}} = 0,$$

where $\phi^{\text{temp}}$ is the signed distance to the zero level set of $\phi$, constructed with the fast marching algorithm. We note in particular that $\phi$ is not directly reinitialized with $\phi^{\text{temp}}$ as this could lead to discontinuities in the motion of the front.

$^8$ $F$ generally depends on the front features, which are defined on the front itself but not on the entire domain.

$^9$ The demonstration of this is given in [62].
Figure 3.10: The narrow band is the set of voxels that are close to the zero level set of the $\phi$ function. A depth of a few voxels is typically chosen. The $\phi$ function is evaluated only inside the narrow band instead of the whole domain.

**Acceleration Technique: Narrow Band Level Set**

The level set method, in its basic form, is computationally expensive. Indeed, each iteration of equation 3.8 must be computed on the entire domain where $\phi$ is defined, and with a total complexity of $O(M \cdot N^3)$ in 3-D ($N$ being the size of the domain, $M$ the number of iterations).

What is done to speed-up the method is to consider the $\phi$ function only in a tight neighborhood around its zero level set, as a narrow band surrounding it. This is illustrated in Figure 3.10. The retrieval of the voxels inside that narrow band is performed by solving the Eikonal equation $|\nabla D| = 1$ starting from the initial contour, $D$ being the distance to the zero level set. This is done twice, once for each side of the front, and the computation is stopped once $D$ has reached the width of the desired narrow-band.

This not only marks the set of voxels that are in the narrow-band, but it provides as well the value of $\phi_0$ for these voxels. Then the classical level set method is iterated until the front reaches a landmine (Figure 3.11b), that is, a voxel on the border of the narrow-band. At that moment, a new narrow-band is built (Figure 3.11b) and the process continues.
Advantages over Discrete Models

Fast marching and level set methods compare favorably to the snake method when it comes to topology adaptation. In the context of an initial contour that is placed inside a structure to be segmented, for instance, and if that structure has a hole that should not be part of the segmentation, the snake method will fail. It will either pass over the hole and segment as if it were part of the shape, or it will stay outside of the hole leaving some part of the shape to segment. Conversely, the level set or fast marching methods will simply go around the hole naturally, as the location of negative $\phi$ function values in the case of level set, or larger values of $T$ in the case of the fast marching method. Topological changes are inherently correctly handled in these methods, while they require workaround and heavy book-keeping if one wants to achieve the same result with the snakes.

Another drawback of the snakes method is that it needs re-parametrization when it becomes elongated (point rarefaction) or compressed (Figure 3.12). Empirical rules can be established that will in turn influence the resulting segmentation. Contrarily, the resolution of the contour is determined by the precision of the domain where $\phi$ or $T$ are evaluated, and level set and fast marching methods do not need that kind of considerations.

Snakes often tend to develop knots when evolving under image-based speeds (Figure 3.12), because of the relative independence of the motion of
its points. These have to be detected and removed appropriately, otherwise they lead to a failure of segmentation. This phenomenon is inexistant in fast marching and level set methods.

Eventually, another advantage of these methods is their easy extension to \(N\)-D, and, in particular, 3-D. This is especially true in our context of CT datasets.

**Applications**

There is a plethora of applications for these front propagation methods, some of them are discussed in [62]. Medical image segmentation with the level set method is just one of these and it was first presented by Malladi et al. [34]. Two other applications are briefly summed up here.

The first one is about robotics path planning. In robotics, one is often interested in finding the shortest path that is going from point A to point B, avoiding obstacles. This can be solved using the fast marching algorithm. To illustrate this, consider the simple case of a ladder that has to be moved in a building for instance. Its configuration is determined by the position \((x, y)\) of its center, and the angle \(\theta\) with the Y-axis. That three-dimensional domain is where one wants to solve the Eikonal equation. We can define a simple speed of propagation \(F(x, y, \theta)\) that has a value of 1 if the ladder does not intersect with any obstacle at \((x, y, \theta)\), and a value of \(\varepsilon\) (close to zero) where there is an intersection. The domain has to be discretized, and the Eikonal equation is solved from point A until point B is reached. The shortest path is given by tracing backwards from point B.
This second application is about simulating etching and deposition in microelectronics. Etching is a process where material is removed from a surface and deposition is a process where material is added. Whenever a physical model is available or can be evaluated, simulation using the fast marching or level set algorithms is possible. A speed of propagation $F$ can be defined correspondingly to the types of etching and deposition that are used, isotropic or not and depending on materials.
Chapter 4

Methods

This chapter details the various algorithms that have been proposed to solve the problem of PE detection, their combination, and the method that is used to evaluate against reference data.

4.1 Overview

Our initial approach for achieving automatic PE detection was first to perform the segmentation of the PA, to extract only the regions of interest, and then to analyse them to find the clots. Looking at the state of the art, we decided to go for a region growing approach similar to the one presented by Zahlten et al. [19], but using modern contour propagation techniques provided by the level sets and fast marching methods. After experimenting a narrow-band implementation of the level set method, we decided to use the fast marching algorithm for its lower computational expense, and we introduced the slice marching algorithm, described in section 4.3.

That segmentation algorithm had a problem for the segmentation of touching thorax vessels, because of its propagation scheme. For this reason, we combined it with a prior model of knowledge of the vessels, to disambiguate nearby vessels. That parametric model had to be co-registered with each patient image, hence a co-registration scheme based on fiducial points was introduced.

The next step was to perform the actual detection of the clots. This was performed from the segmentation of the PA, using a combination of image operators. The clots partially obstructing the PA were generally correctly detected, but we found out this was tightly dependent on the quality of the underlying segmentation. Total obstruction of an artery was still a problem, since the segmentation was literally stopped there and further detection was
not possible. Because of that, another detection scheme that is independent from the PA segmentation was developed.

The idea behind this second detection method was in making assumptions about clots shapes and densities. It proved to be able to detect much more clots than the other method, at the price of increased false alarms.

Finally, a way for assessing the performance of methods for emboli detection was needed and an evaluation scheme was presented and applied to our methods. It is based on the realization of ground truths by radiologists.

This chapter first describes our experiments with the level set and fast marching methods, in terms of definition of the speed of propagation, algorithm initialization, etc. in the next section. Section 4.3 details the initial form of the proposed algorithm, named slice marching and includes the slice definition, considerations about the speed of propagation, and reconstruction of the bifurcation graph. After that, section 4.4 introduces the use of prior information to the slice marching algorithm to cope with the problem of joint vessels. PE detection algorithms are next presented in section 4.5 and a validation is presented in the last section (4.6). The results of these methods are discussed in the next chapter (5).

4.2 Level Sets and Fast Marching Experiments

These active contour methods were implemented for testing, and a number of considerations that are of interest for the following sections are expressed here. A full implementation of the narrow-band level set method on a regular grid is realized, as well as the fast marching algorithm.

4.2.1 Initialization from a Given Contour

The fast marching and level set algorithms require initial data, that depend on the shape of the starting contour. This contour has no constraint of passing through grid points, and that has to be accounted for. The fast marching algorithm requires the known set of points to be available, which is a set where $T$ is known, while the level set algorithm requires an initial value of the $\phi$ function, that is chosen to be the signed distance $D$ from the initial contour. In other words, we want to solve the distance equation,

$$|\nabla D(\bar{X})| = 1, \quad \text{with } D(\bar{X}) = 0 \text{ on } \Gamma, \quad (4.1)$$

using the fast marching algorithm for instance, and it thus comes back to finding the initial known set of that algorithm. A few possible configurations are envisaged here.
4.2 Level Sets and Fast Marching Experiments

**Single Point**

This is the simplest case, where the value of $D$ is set to zero if the point resides exactly on a voxel, or, alternatively, the Euclidian distance is computed otherwise,

$$D(\bar{V}_i) = |\bar{P} - \bar{V}_i|$$

(4.2)

as illustrated in Figure 4.1. The 4 neighbors (8 in 3-D) form the known voxel set.

![Figure 4.1: Fast marching initialization. $T$ is computed on neighbors of $P$, $P$ being the point where $T$ is known to be 0.](image)

**2-D Polygonal Contour**

In this case, the intersections of the contour with the grid axes are computed, $s$, $t$, $u$, and $v$ being the distances from a voxel center to these intersections respectively for positive X and Y axes, and their negative counterparts, as expressed in Figure 4.2. A 1-D or 2-D scheme is used depending on the number of axes reached by the contour,

$$\frac{1}{D^2(V)} = \frac{1}{\min^2(s, u)} \quad (1-D),$$

(4.3)

$$\frac{1}{D^2(V)} = \frac{1}{\min^2(s, u)} + \frac{1}{\min^2(t, v)} \quad (2-D).$$

(4.4)

When multiple solutions are possible, the one with the smallest value is used, noticing the 2-D scheme always gives a smaller value than the 1-D one, because the 1-D scheme overestimates the distance while the 2-D scheme is exact. Various configurations are represented in Figure 4.2. The sign is chosen according to the relative position of the point with the polygon, inside (-) or outside (+).
Figure 4.2: Various configurations when computing the distance to a polygon. For (a) and (c), the 1-D scheme is used, and for (b) and (d), the 2-D scheme is used.

3-D Polyhedron Surface

This case is a direct extension to the previous one, this time with a 3-D scheme,

$$\frac{1}{D^2(V)} = \frac{1}{\min^2(s,v)} + \frac{1}{\min^2(t,w)} + \frac{1}{\min^2(u,x)}$$  \hspace{1cm} (4.5)

3-D Bézier Spline

Distance from a Bézier spline is needed in a method presented later in the text (section 4.4.4). The fast marching algorithm can be used in that case, and we provide this initialization scheme. This is not to be confused with tracking the evolution of a Bézier contour itself but instead it means evolving a surface that grows from the curve. To initialize the known set, the Bézier curve is sampled with a sufficiently small parameter step, leading to a space distance between successive samples that is well below the grid spacing. This is illustrated in Figure 4.3. For each evaluated point of the Bézier curve, the Euclidian distance to the four surrounding grid points (eight in 3-D) is computed, and that distance is attributed to those points if smaller than the one currently attributed (smallest up to now). The error of this scheme is proportional to the space-step of the sampling of the curve.

4.2.2 Expression of Speed

The expression of the speed determines the evolution of the contour. What has to be noted is that it can depend on any feature of the image, the contour itself, or a priori information. The fast marching method requires that the speed be strictly positive, but the level set method has no such limitation. A value of $F = 1$ means that the contour moves with a constant speed.

\footnote{Nomenclature is adapted to the 3-D case.}
4.2 Level Sets and Fast Marching Experiments

Figure 4.3: The Bézier curve is evaluated with a fixed parameter step, and the computed distances are attributed to neighboring grid points (in 2-D, same principle in 3-D).

outwards (by convention), and never stops, which is related to the balloon force introduced by Cohen [58].

If we are interested in segmenting contours, we want that the curve slows down or even stops when a higher gradient amplitude is met in the image. Such a speed may have the following expression:

$$F(\bar{X}) = \frac{1}{1 + |\nabla I(\bar{X})|},$$  \hspace{1cm} (4.6)

where $I(\bar{X})$ is the density at voxel $\bar{X}$, or image intensity. This primarily means that the speed of the contour is low when the gradient is high in the image.

If one wants to prevent the apparition of sharp corners, one can define a speed that decreases when the curvature $\kappa(\bar{X})$ increases. Work to prevent the contour from leaking in small gradient holes in the image was presented by Caselles et al. [77], where an additional term for pulling the contour is introduced. Prior information such as the expected area or perimeter of the contour can also be used to modify the speed of propagation.

The gradient of the image is a local characteristic, meaning that if the contour is located where the gradient is zero, but not far from a high amplitude gradient in the image, it will not move. The gradient vector flow (GVF) was introduced by Xu et al. in [57] specifically to solve that issue, with as intent to pull the contour towards these visible image contours through a vector diffusion equation that is applied to the image gradient.

Another important point is the stop criterion of the front, or how to decide when to stop its evolution. This can be done by measuring the amplitude of displacement (or speed) between two iterations in the level set
algorithm and stopping if the value is smaller than a chosen amount. Any other criterion can be defined depending on the specific application.

### 4.2.3 Expression of Numerical Schemes

Starting from the fast marching upwind scheme given in 3.6, we want to solve \( T \) knowing its values at the neighboring positions, that is:

\[
\frac{(T - T_x)^2}{\Delta_x^2} + \frac{(T - T_y)^2}{\Delta_y^2} + \frac{(T - T_z)^2}{\Delta_z^2} = \frac{1}{F^2},
\]

where \( T_x, T_y, \) and \( T_z \) respectively represent the selected minima on each axis, either in the positive or negative direction. Solving this second degree equation brings the following expression:

\[
T = \frac{T_x \Delta_y^2 \Delta_z^2 + T_y \Delta_z^2 \Delta_x^2 + T_z \Delta_x^2 \Delta_y^2 \pm \sqrt{\rho}}{\Delta_y^2 \Delta_z^2 + \Delta_z^2 \Delta_x^2 + \Delta_x^2 \Delta_y^2},
\]

where

\[
\rho = (T_x \Delta_y^2 \Delta_z^2 + T_y \Delta_z^2 \Delta_x^2 + T_z \Delta_x^2 \Delta_y^2)^2 - (\Delta_y^2 \Delta_z^2 + \Delta_z^2 \Delta_x^2 + \Delta_x^2 \Delta_y^2) \cdot \left(T_x^2 \Delta_y^2 \Delta_z^2 + T_y^2 \Delta_z^2 \Delta_x^2 + T_z^2 \Delta_x^2 \Delta_y^2 - \frac{\Delta_x^2 \Delta_y^2 \Delta_z^2}{F^2}\right).
\]

Multiple cases are possible, with two, one, or no roots. For the case with two solutions (\( \rho > 0 \)), we understand that the smallest solution is the one with the contour evolving from the considered point to the three others, and the largest is the one coming from the the others to the considered point. The latter is the one of interest because we want to respect the upwind direction, as discussed in section 3.4.4 the + sign is chosen in Equation 4.8.

Next, we consider the case with exactly one solution when \( \rho \) is zero. It can be shown that this is possible only if \( F = \infty \) and \( T_x = T_y = T_z \). The meaning of this is that the contour evolves with an infinite speed at that point, and that is only possible if the arrival times on the neighbors are the same.

Finally, there can be no solution to this equation (\( \rho < 0 \)). This expresses that the solution \( T \) cannot come from the three directions simultaneously but instead from a dimensional subset. This corresponds to a shock, as expressed in [62] and in that case, we select the \( N - 1 \) smallest \( T \) values and try the \((N - 1)\) dimensional scheme (\( N \) being 3 or 2).

The 2-D and 1-D schemes are given here,

\[
T = \frac{T_x \Delta_y^2 + T_y \Delta_z^2 + \Delta_x \Delta_y \sqrt{\frac{\Delta_x^2 + \Delta_y^2}{F^2} - (T_x - T_y)^2}}{\Delta_x + \Delta_y} \quad (2-D),
\]

\[
T = T_x + \frac{T_y}{F} \quad (1-D).
\]
We remark the easy interpretation of the 1-D scheme which is that the time required to travel a distance equals that distance divided by the speed of propagation.

Eventually, here are the simpler expressions of the isotropic peculiar case when grid sizes are the same along the three axes, $\Delta_x = \Delta_y = \Delta_z = \Delta$,

$$
T = \frac{(T_x + T_y + T_z + \sqrt{\frac{3\Delta^2}{F^2} - (T_x - T_y)^2 - (T_y - T_z)^2 - (T_z - T_x)^2})}{3} \quad (3-D),
$$

$$
T = \frac{(T_x + T_y + \sqrt{\frac{2\Delta^2}{F^2} - (T_x - T_y)^2})}{2} \quad (2-D),
$$

$$
T = T_x + \Delta/F \quad (1-D).
$$

(4.11)

### 4.2.4 Narrow-band Level Set Application

We tested the narrow-band level set algorithm on 2-D images. This is illustrated in Figure 4.5. The level set algorithm is by nature slower than the fast marching algorithm, but has the advantage of being able to handle a broader range of speeds, and in particular, negative ones.

The expression of the speed of propagation $F$ that is used in this particular experiment is

$$
F(\tilde{X}) = \frac{1 - \varepsilon_\kappa(\tilde{X})}{1 + |\nabla G_\sigma(\tilde{X}) \otimes I(\tilde{X})|},
$$

(4.12)

where $\kappa(\tilde{X})$ is the curvature of the front, $G_\sigma$ a Gaussian kernel, and $I(\tilde{X})$ the luminance of the image.

With this speed of propagation, the contour continuously grows at a constant speed over a uniform area of the image, and it slows down over areas of high gradient amplitude because of the denominator term. The choice of
Figure 4.5: Illustration of the evolution of a contour using the level set algorithm. The contour is in red, and the narrow-band is painted in blue (only where $\phi$ is positive). Notice that the narrow-band is reevaluated when the front reaches its border and that the final contour accurately fits the target structure.

The $\sigma$ parameter of the Gaussian kernel is dependent on the size of the objects that are to be segmented. The numerator ensures that the contour remains smooth at all time, allowing negative speeds when the contour tends to develop sharp corners (high curvature).

Three issues are depicted when using this algorithm for segmentation purposes. Firstly, the stop criterion is not necessarily easily determined, and often one can continue to iterate the formula without knowing when to stop. The second point is that the time step that is used to evolve the $\phi$ function is determined by the largest value of $F$ on the front, which in essence means that local properties can influence the speed of convergence globally. This is especially disadvantageous in our application, because of the large extent of the structure that we want to segment and slowing down the convergence when arriving at smaller vessels because of a condition near the seed point.
is not effective. Lastly, this method is computationally expensive, because each iteration involves a modification of the $\phi$ function on the entire domain (or the narrow band) and only provides, in the best case, a displacement of the contour of one unit (a one-voxel distance) per iteration.

The second and third drawbacks do not exist in the fast marching algorithm, which steadily converges independently of the underlying data, with a very low computational load of $O(N \log(N))$. The time constraint requirement of our application suggests that the level set algorithm is not the one to choose here. For these reasons, we opted for a method based on the fast marching algorithm and, the first drawback not being intrinsically solved in this algorithm, we have proposed a freezing extension that is exposed in the next section.

### 4.3 Slice Marching Algorithm for Vessels Segmentation

The principle of this algorithm is to advance inside the vessel, slice by slice, using a propagation scheme based on the fast marching algorithm. The branching tree of the vessel can then be reconstructed by looking at the connectivity of the slices.

#### 4.3.1 Front Propagation

The concept of wave introduced by Zahlten \[19\] is interesting in that it permits the reconstruction of the vessels tree by analysing the connectivity of the wavefront. The wave depth was extended to two voxels by Bruijns \[21\] with the concept of double waves. Our approach enables the use of arbitrarily deep slices of voxels, and in that sense, can be considered as a generalization of the concept of wave.

Starting from a seed point that the radiologist localizes at the beginning of the PA, we let a contour (a surface in 3-D) grow. To accomplish this, we want to solve the Eikonal equation \[4.1\] using the fast marching algorithm as expressed in section 3.4.4.

If we suspend the computation of the solution when the crossing time $T(\bar{X})$ becomes higher than a given threshold $T_{\text{slice}}$:

\[
\text{if } (T(\bar{X}) > T_{\text{slice}}) \Rightarrow \text{suspend} \quad (4.13)
\]

and if we define our speed function $F(\bar{X})$ in a way that it is close to one inside the vessel, and by noticing that time and distance are equivalent when
speed is unity, we can tell that $T_{\text{slice}}$ is the depth of the slice (Figure 4.6) expressed in voxel units, which may be a non integral number.

Then, we iterate using a time threshold $T_{\text{slice}}$ of the form:

$$T_{\text{slice}} = T_k = k \Delta T_{\text{slice}}$$

$\Delta T_{\text{slice}}$ being the desired depth of the slice, each time getting a new slice. We define a slice $\mathcal{S}_k$ as:

$$\mathcal{S}_k \triangleq \{ \bar{X} \mid T_k \leq T(\bar{X}) < T_{k+1} \}$$

where a temporal criterion actually defines spatial zones. This slice definition is conceptually similar to the cluster definition introduced by Masutani et al. in [24]. Unlike Yi et al. [27], the growing of slices is not constrained to custom shapes (cubes in their developments) but solely to image data through the speed of propagation.

### 4.3.2 Speed of Propagation

The speed function should be close to one inside the vessel to travel at full speed there, and close to zero outside to stop (slow down) near vessel boundaries. But the speed of propagation is required to be strictly positive by the fast marching algorithm. Considering an input voxel density $I(\bar{X})$ between 0 and 1, we provide the corresponding speed by

$$F = \varepsilon \quad \text{if} \quad I(\bar{X}) < I_{th}$$

$$= I(\bar{X}) \quad \text{otherwise}$$

This slice concept is not to be confused with slices of CT images.
where \( \varepsilon \) is a positive, near zero value and \( I_{th} \) a threshold on luminance (we used a value of 90\%, using the hypothesis that vessels are correctly opacified).

An important point is that the speed of propagation \( F \) can be used for controlling the target shape and this is a major difference between this and other region growing schemes. In particular, Masutani et al. [24] used bounded space morphology, and Zahlten [19] and Bruinjs [21] used a more simple 8- or 26-connectivity scheme. When compared to these approaches, this method adds the flexibility of permitting the introduction of external information to the segmentation process. This is inherently supported in the algorithm through the modification of the speed of propagation which, in turn, influences the reached shape. We exploit that feature to permit the use of a model of prior knowledge in the segmentation task. This is described in section 4.4. Another difference with the other schemes is that they grow in a discrete manner, adding one row of voxels at a time, contrarily to this one, which supports non integral depths. This probably has a limited influence on big structures with size of tenths of voxels or more, but may have an importance for small vessels whose size is near the limit of CT resolution.

### 4.3.3 Freezing the Front

Deschamps et al. [39] described a mechanism of freezing, that is, preventing \( T \) from continuously growing after it has reached vessel borders. In their freezing mechanism, they compute a second solution \( T_2 \) alongside \( T \) to equation 4.1 this time with \( F \) set to one but following the exact same path as \( T \), and they consider the difference between \( T \) and \( T_2 \) to determine which voxels should be frozen. Frozen voxels are removed from the trial set (see section 3.4.4), preventing further increase. In contrast, we avoid the computation of \( T_2 \) by comparing \( T \) to the time corresponding to the start of the slice.

Next, by choosing a speed function \( F(\bar{X}) \) that is small enough outside the vessel, we can freeze the active front (denoted as the trial set in the fast marching algorithm [62]) by not considering voxels whose arrival time is higher than a second threshold, \( T_{\text{bound}} \):

\[
\text{if } (T(\bar{X}_t) > T_{\text{bound}}) \Rightarrow \text{remove } \bar{X}_t \text{ from trial set } \quad \text{(4.18)}
\]

Frozen voxels are not anymore part of the active front, and this has the side effect of reducing the computational load of the method, as the trial set is smaller and its update quicker.
4.3.4 Detection of Bifurcations

The PA are a highly hierarchical structure, and we are interested in tracking bifurcations to be able to model that tree. As the front advances inside the vessels for recovering the entire arterial tree, the hierarchical structure of the vessels is built, where child slices are defined relatively to their parent according to the order of traversal by the front. Voxels of slices of the same generation (same $k$) are checked for their mutual connectivity with connected component analysis. Each unconnected group of voxels defines a new slice. The false bifurcation phenomenon reported by Zahlten [19] is avoided because the bifurcation decision is taken on the slices themselves, which are deeper structures than one-voxel deep waves are thus less sensitive to the rough surface of vessels.

4.3.5 Estimation of Vessel Features

Interesting vessel features can be evaluated by analysing these slices, such as the center of gravity $\bar{C}_k$ (4.19) of successive slices, their section $S_k$ (4.20), inter-slice direction $D_k$ (4.19), and the vessel curvature $\kappa_k$ (4.20). These features can in turn be re-injected in the process through the speed function $F(\bar{X})$. This can be used for instance to accelerate voxel traversal on the outside of a turning vessel to keep the front perpendicular to the vessel.

$$\bar{C}_k = \frac{1}{N_k} \sum_{X_i \in \mathcal{S}_k} X_i$$
$$D_k = \bar{C}_{k+1} - \bar{C}_k$$

$$S_k = \frac{N_k}{T_{k+1} - T_k}$$

$$\kappa_k = |D_k - D_{k-1}|$$

where $N_k$ is the number of voxels belonging to slice $\mathcal{S}_k$. Yi et al. also used features that were evaluated in local cubes containing vessel parts, although they served mostly to derive contrast thresholds for the growing scheme.

4.3.6 Initialization and State Management

As pointed out in section 4.3.1, the seed point is set at the start of the PA. This is not far from the heart and before the first bifurcation of the PA (left-right).

The segmentation algorithm must perform differently depending on its location in the vessel hierarchy. In effect, the speed of propagation is set to a value near zero for the direction going towards the heart, as we want to prevent the contour from going back to other vessels through the heart. At a
4.4 Prior Knowledge Model Extension

later time in the segmentation, we want to restrict the front from moving to vessels with larger sections than the one being segmented, as this probably means that the front is switching to a vein for example and is travelling in another vessels tree.

This is realized by tracking the section $S_k$ of the slices as the front propagates, along with a low-pass version of it. That is used to control a segmentation state machine which permits the rejection of slices, and also permits to choose different expressions of the speed of propagation. This relies on concepts discussed later in the text and it is presented in more details in section 4.4.6.

4.3.7 Illustration of the Method

The results of the slice marching algorithm can be observed in Figure 4.7 and they are discussed in more details in the next chapter. What can be observed is that slices of vessels are defined as the front advances in the vessels that are concentric spheres at the start (unconstrained growing). Features are computed for these slices, as well as their relationship with each other (tree). An important remark here is that the emboli stay outside of it because of their lower radio-densities. As expected, the fast marching algorithm it relies on easily copes with topology changes (holes) introduced by the clots.

4.4 Prior Knowledge Model Extension

An extension to the initial slice marching algorithm is proposed to tackle the problem of touching vessels.

4.4.1 Problem of Nearby Vessels

Due to the limited resolution of CT scanners, as well as acquisition artifacts and image noise, nearby thorax vessels typically present large contact surfaces, over which there is no visible decrease in terms of radio-density. In other words, you can freely move from one vessel to another one without reaching voxels of smaller densities. This is illustrated in Figure 4.8.

Because regular region growing algorithms only make use of image information, the contour naturally moves to adjacent vessels which is not correct because it segments unwanted vessels. Moreover, we have found that situation to be recurring in our test cases and it had to be addressed. A
Figure 4.7: Slice marching results. From a chosen seed point, the contour grows in 3-D, until the entire PA is reached. The segmented region is in blue, and successive slices are separated by yellow surfaces (curves in the 2-D sections).
Figure 4.8: Nearby thorax vessels often have luminosity contacts, which prevent region growing methods from correctly segmenting them. Different vessels are highlighted on the second row, and a surface of contact of the same size as the vessels themselves is observed.

morphological operation of opening was initially tested, in the hope of removing these contact regions, but it did not succeed because the size of that contact between the structures was generally comparable to the sections of the involved vessels. Another technique, this time based on selective opacification of the PA and not the other ones through specific injection timing and rate is not practically reproducible.

For these reasons, we decided to make use of a higher level prior information, just as it is done by the radiologists themselves. We opted for a parametric model in the same spirit as Kitasaka et al. where the topology and geometry of vessels of the thorax are modeled as parametric curves. This model is then registered to the patient image and then used in the segmentation method.

4.4.2 Definition of the Model

This prior knowledge model is defined as a set of fourth order, 3-D Bézier curves. More specifically, each part of the vessels comprised between two bifurcations are modeled as a number of Bézier segments, as illustrated in Figure 4.9 where 6 such segments compose the represented vessel skeleton. A Bézier segment is made of four control points that can be viewed as two ending points and two tangents relative to these ending points.

B-Spline could have been used instead, but we wanted a tool that would make it easy to define a model. One interesting property of the Bézier splines
Methods

Segment i

Figure 4.9: Vessels are modeled as a set of Bézier curve segments that represent their centerline. The radius of the vessel is attributed to each ending control points. Radiologists create that model by clicking where curve should pass on a chosen patient image.

is that they interpolate their ending points, which basically means that the radiologist clicks on a point where he wants the curve to pass through, unlike B-Splines.

In practice, the model is realized on a chosen dataset, and registered to other patients. Taking into account usability concerns, the tool permits the automatic computation of points $\bar{P}_{i,2}$ and $\bar{P}_{i,3}$ of Bézier segment i using the position of $\bar{P}_{i-1,1}$ and $\bar{P}_{i+1,4}$ in such a way that the physician only has to click on successive points on the medial axis of the vessels. Another goal is to preserve the geometric continuity between successive Bézier segments.

The conditions to make the curve continuous are that

$$\bar{P}_{i-1,4} = \bar{P}_{i,1}$$

and $\bar{P}_{i-1,3}, \bar{P}_{i,1}, \bar{P}_{i,2}$ be co-linear.

(4.21)

Our scheme for automatically computing $\bar{P}_{i,2}$ and $\bar{P}_{i,3}$ from surrounding $\bar{P}_{i,1}$ is given by these equations:

$$\bar{P}_{i,2} = \bar{P}_{i,1} + \alpha_i \frac{|\bar{P}_{i+1,1} - \bar{P}_{i,1}|}{|\bar{P}_{i+1,1} - \bar{P}_{i-1,1}|} (\bar{P}_{i+1,1} - \bar{P}_{i-1,1}),$$

$$\bar{P}_{i,3} = \bar{P}_{i+1,1} - \alpha_i \frac{|\bar{P}_{i+1,1} - \bar{P}_{i,1}|}{|\bar{P}_{i+2,1} - \bar{P}_{i,1}|} (\bar{P}_{i+2,1} - \bar{P}_{i,1})$$

(4.22)

where $\alpha_i$ is a parameter that can be adjusted between 0 and 1 to make the curvature sharper or smoother (we generally use $\alpha_i = 1/3$). The value of the radius is provided as $R_i$, and is chosen to be constant between two bifurcations.

---

Bézier curves are actually a subset of B-Spline and the interpolation of points can be realized in B-Spline as well by augmenting the multiplicity of these control points, but that process is less intuitive.
Figure 4.10: 3-D views of the model. The model is showed as a colored wireframe, that is sliced by a descending axial plane for display. The aorta is selected and appears in orange color, other vessels are in red. The user creates vessels by adding point where it should pass in the image. Branches can be defined as well, and a radius is attributed to each vessel section.

The vessels that have been modeled are the aorta, the PA, the PV, and the vena cava. Figure 4.10 shows 3-D wireframe views of that model as they represented in the actual computer software.

4.4.3 Co-registration of the Model to the Image

Fully Automatic Approach

Our initial intention was to make the co-registration of the model fully automatic and this was meant to be done by registering the model at the same time as the contour advanced inside the PA. This approach was experimented by measuring the reached distances along successive vessel slices by accumulating the $|D_k|$ of equation 4.19 and by identifying the corresponding points on the parametric model. That correspondence could then be used for the co-registration as described later in this section, and in turn to influence the segmentation as expressed in the next section. But two difficulties arose, namely

- the scale of the model (vessel lengths) greatly varies among patients,

- the locality of the registration sometimes prevents forthcoming segmentation.

5 excepted for the seed point for the segmentation
The former point means that a shift appears and grows as we advance in the PA because the correspondence is established on a model that is either larger or smaller than the actual vessels in the dataset, leading to incorrect registration after a finite segmentation time. The latter point is more dependent on the geometry of each vessel as well as their relative positions with each other, which vary among patients. Assuming for instance that the aorta is closer to the PA in a given patient than in the model, this might lead to some parts of the aorta being incorrectly segmented as the PA, which in turn would make the co-registration of the model fail.

A more sophisticated scheme was also tested, by defining a likelihood of correspondence $L(i, t_i)$ between the model and the segmented slices (Figure 4.II). This likelihood takes into account various features of the vessels, weighted accordingly to their relevance. Its expression is

$$L(i, t_i) = \alpha |\bar{C}_i - C(t_i)|^2 + \beta (\bar{C}'(t_i) \cdot \bar{D}_i) + \gamma (\kappa_i - |\bar{C}''(t_i)|)^2 + \delta (\Delta - t_i)^2,$$

where $t_i$ are the parameter values to the parametric curves and $C(t_i)$ the corresponding point on the curve, $\Delta$ is the slice depth and $\alpha, \beta, \gamma, \delta$ are four weighting coefficients. $L(i, t_i)$ has to be maximized over the choice of $t_i$,

$$\{t_i\} = \arg \max_{t_i} L(i, t_i).$$

We came to the conclusion that such a system, where segmentation and co-registration depend on each other, is inherently unstable as any error in both processes leads to a divergence. For this reason, we reverted to a manual, yet as practical as possible, co-registration method.

**Fiducial Points based Approach**

As expressed in the previous section, the model is defined on an arbitrarily chosen dataset. In that chosen image, a number of fiducial points are chosen at precise locations surrounding the model. These are picked in such a way that they are easily located on another dataset. The set of points that have been used in practice have their anatomical definitions expressed in Table 4.II.

Once this set of fiducial points has been located on another patient’s image, a thin plate spline 3-D space transformation is evaluated from the correspondence with the original set. This non-linear transformation can
Figure 4.11: Automatic co-registration of the model to the segmented PA slices so far. The segmentation advances by a few slices and these are matched to corresponding points on the model. The model is in turn deformed and the process is iterated.

then be used to register the model to the new patient, as illustrated in Figure 4.12.

Because this method is not automatic as it requires the intervention of the radiologist to locate the fiducial points, it is not subject to instability and it is a lot more convenient than recreating a model from scratch on each new patient.

The number of fiducial points is an important parameter. Taking into account the fact that only the centerlines of vessels in the immediate surrounding of the PA are of interest in this co-registration method, a rough transformation is acceptable. Our experiments with 4, 7, and 12 fiducial points revealed that 4 is not enough, as an affine transform cannot handle the non-linear variability of vessels, and that 12 is probably too much work for only a slight increase in precision. 7 fiducial points are chosen as a compromise and that configuration is found to give a satisfactory matching of the model to the dataset.

4.4.4 Using the Model as Segmentation Guide

Once the model is correctly registered to the patient image, it is used to influence the segmentation method. This is executed elegantly and consis-
Table 4.1: Fiducial points and their anatomical definitions. These points are chosen to encompass the volume where the model is needed. They are manually located on each new patient.

<table>
<thead>
<tr>
<th>Fiducial Name</th>
<th>Anatomical Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_1$</td>
<td>Center of the contact section between the second carotid and the arch of aorta.</td>
</tr>
<tr>
<td>$F_2$</td>
<td>Inner point on PA border at the separation between left and right PA.</td>
</tr>
<tr>
<td>$F_3$</td>
<td>Vena cava center point closest to the right PA.</td>
</tr>
<tr>
<td>$F_4$</td>
<td>Aorta center point closest to the left PA.</td>
</tr>
<tr>
<td>$F_5$</td>
<td>Aorta center point at the same height as $F_3$.</td>
</tr>
<tr>
<td>$F_6$</td>
<td>Center of the small PV at the left of the left PA at the same height as left PA center.</td>
</tr>
<tr>
<td>$F_7$</td>
<td>PV center point at the left/right bifurcation.</td>
</tr>
</tbody>
</table>
Figure 4.12: Fiducial points-based co-registration. A patient is chosen as a reference and a set of fiducial points are defined (a), as well as a model for thorax vessels centerlines (c). The set of fiducial points is located on another patient image (d), which induces a space-transform (e) that is used to adapt the model (f).
tently with the front evolution method by modifying the expression of the speed of propagation $F$.

![Diagram](image)

**Figure 4.13**: After the model has been registered, potential functions are evaluated on the domain. The potential function along the red line is reported on the graph. Collisions of potentials serve to define zones where the speed of propagation $F$ is decreased. This permits the creation of the missing boundaries as depicted by the dash line.

This is realized by defining a vessel potential function, $P(\bar{X})$, as

$$P_i(\bar{X}) = \frac{R_i}{D_i(\bar{X})} \quad (4.23)$$

where $i$ refers to the vessel section being considered, $R_i$ is the mean radius for the vessel section, $D_i(\bar{X})$ is the Euclidian distance between $\bar{X}$ and the nearest point on the centerline of vessel $i$.

From the potential function, we derive the vessel interior indicator, $V(\bar{X})$ as

$$V(\bar{X}) = \begin{cases} 1 & \text{if } \max_i(P_i(\bar{X})) - \max_{j \neq i}(P_j(\bar{X})) > \lambda \varepsilon \\ \text{otherwise} & \end{cases} \quad (4.24)$$

with $i_{\text{max}} = \arg(\max_i(P_i(\bar{X})))$. $\lambda$ is a parameter ($\lambda > 0$) that can be used to adjust the reconstructed boundaries width, and $\varepsilon$ a positive value near zero.

The vessel potential is used to modify the speed of propagation of the active front:

$$F'(\bar{X}) = F(\bar{X}) \cdot V(\bar{X}). \quad (4.26)$$

The function $V(\bar{X})$ is effectively built in such a way that it is near zero when potentials of neighboring vessels are the same (potential collision). This has the effect of recreating the possibly missing vessel boundaries where $V(\bar{X})$ imposes the boundary. On the contrary, when other vessels are not in its immediate surroundings, then, the image imposes the boundary exactly
as it does when the model is not used. This is in contrast to the method proposed by Kitasaka et al. [50], as it gives more reliance to the image for boundary reconstruction when that information is present.

In practice for the sake of computing efficiency, $P_i(\bar{X})$ is computed on a sub-sampled 3-D grid (by a factor of 8) using the fast marching algorithm for each vessel. The algorithm is initialized as described in section 4.2.1 on Bézier splines. Two values of $P_i(\bar{X})$ are stored on that grid; these are the largest ones when iterating over vessels in the model.

When the value of $V(\bar{X})$ is required for calculating the speed of propagation $F'$ on the regular grid (not the sub-sampled one), $P_i$ values are piecewise tri-linearly interpolated from the eight surrounding values at the considered position $\bar{X}$. The slice marching algorithm simply uses $F'$ instead of $F$, and missing boundaries are created.

The implementation of $F'$ in C programming language is provided in Appendix B for illustration purpose.

### 4.4.5 Segmentation Takeoff

The segmentation is ignited with a seed point that is chosen inside the main trunk of the PA, in the proximity to the heart. Under a constant speed of propagation, the contour grows similarly in all directions, including the one leading back to the heart. As most thorax vessels and the heart are generally opacified, the contour would eventually segment all thorax vessels, which is to be avoided.

To handle that situation, the co-registered model is used. The speed of propagation $F$ is modified to prevent it from going backwards, towards the heart. The equation of the plane that is orthogonal to the model and passes through the seed point $S$ is established (see Figure 4.14), and all voxels on the wrong side of that plane are attributed a residual speed.

$$F_{\text{init}}(\bar{X}) = F'(\bar{X}) \quad \text{if} \quad (\bar{T}_{\text{model}} \cdot (\bar{X} - \bar{S}) > -N_{\text{offset}}),$$
$$\varepsilon \quad \text{otherwise},$$

where $T_{\text{model}}$ is the normalized tangent to the registered model of the PA at its first control point, $S$ is the seed point, and $N_{\text{offset}}$ an offset value whose role is to translate the plane backwards by a few voxels. That way, we prevent the removal of voxels too close to the plane which could be responsible for stopping the propagation because of rounding errors, as illustrated in Figure 4.14.

---

6 the wrong side of that plane is the one where the heart is.
Figure 4.14: The initial speed of propagation is modified to prevent going back to the heart. The image on the right shows that the segmented region (in green) is constrained by a plane at its start.

This modification of the speed must only occur at the start of the segmentation. This justifies an initialization state in the segmentation state machine.

4.4.6 Switching Off the Model

The model can be used where vessels topology or shapes are not dependent on the patient. This is verified for most thorax vessels, as the aorta, vena cava and the main trunk and first bifurcations of the pulmonary arteries and veins. These last two vessels present a higher variability in shape as you go farther from the heart. For this reason, we have to stop modeling these past a given order of bifurcations. Switching off model usage is accomplished, and rules are added to keep control of contour evolution past this point. In particular, we want to prevent the contour from going back to the heart through the veins from a contact with a segmented artery.

Figure 4.15: The segmentation goes through multiple states, where the expression of the speed of propagation changes.

Three states are defined, as represented in Figure 4.15, which are used to select the expression of the speed of propagation. The transitions between states are evaluated by analyzing the vessel section estimates on successive
4.4 Prior Knowledge Model Extension

slices. The first transition arises on the first section decrease in successive slices. This is because the seed point initially grows spherically (section between successive slices is about the ratio of the corresponding spheres, that is $S_{k+1}/S_k = (R + \Delta)^2/R^2 > 1.0$) until it reaches the vessel border. At that point, the section decreases $(S_{k+1}/S_k < 1.0)$ and the state is changed to state 2, where the speed constraint is released.

The maximal vessel section is kept as the front propagates in the PA, after it has entered state 2, and is used to determine the second transition. Vessel sections are filtered (low-pass) on successive slices, and the second transition to state 3 is fired when $f(S_k)/S_{\text{max}} < R_{\text{thres}}$ (a value of 0.15 was used, which corresponds to a radius ratio of about 0.39 for circular sections).

As we have found out, contacts between the PA and the PV sometimes exist farther in the hierarchies of these vessels. As the model is not anymore used in that state, another mechanism is used to prevent the contour from going back towards the heart inside the veins. Indeed, section increases are banned in state 3, by marking the voxels of these slices as frozen. This does not guarantee that the contour will not move to a vein, it just means that if it moves to an adjacent vein, it will only segment it in the direction of larger-to-smaller sections. An approach where opposite $\bar{D}_k$ were searched for (equation 4.20), as a sign of entering another vessel and to discard a subtree of the slices, but with not much success.

4.4.7 Illustration

The model-based segmentation of the PA is illustrated in Figure 4.16, where the algorithm is first run without the model (using $F$), and then with the model (using $F'$) on the same datasets. What is observed is that the model actually solves the ambiguity of vessel boundaries and prevents the contour from moving to adjacent vessels despite the lack of a visible vessel border in the image. The re-created vessel boundaries are not always at the location one would expect, but generally shifted toward the interior of one of the vessels. This is explained by the fact that there is also a variability in vessels radii, and also by the fact that co-registration of the model to the image is not perfect at every location in the volume.

4.4.8 Influence of Emboli on Segmentation

Recalling the initial goal of emboli detection, we have to see how the presence of clots influences the segmentation process. Being darker than the contrast-enhanced blood-flow, the segmentation front simply goes around them as already evoked earlier; they are outside of the segmentation. This can be
Figure 4.16: Results of model-based segmentation. The left column (a) shows 3 datasets that are erroneously segmented (Aorta, vena cava, or PV are also segmented). (b) shows the same datasets correctly segmented with prior model information, where the front does not move to the aorta or vena cava. Notice the displacement of the reconstructed boundary (top-right).
observed in Figure 4.16 for instance, where dark spots are outside of the blue region.

Two situations that may happen are that the clot either fully obstructs the vessel or does not, and, in the former case, the segmentation of the PA just stops. This typically occurs in smaller vessels but is also possible in larger ones for massive embolism. For emboli which do not completely obstruct the vessel, the front typically goes around it and continues its path inside the vessel. As one could intuitively suppose, complete blockage has an impact on the detection method that relies on this segmentation.

4.5 Detection Algorithms

The topic of thrombi detection in the PA was not much studied, probably because of the recent shift towards the CT modality. Two different approaches were found in the literature, the first being based on prior segmentation of the PA developed by Masutani [67, 68] and the second one without segmentation by Liang et al. [70].

A first approach based on the segmentation of the PA is experimented, as a logical way to exploit our previous efforts. It is very similar to [67], as it looks for holes or other inward defects in the PA mask. As we found out, this method works well for emboli that do not totally obstruct the artery, even when it is in contact with the border of the vessel. However, clots that cause total blockage pose problem to this method. This justifies the development of a second detection method that does not make any assumption about the PA but relies directly on image data.

4.5.1 Emboli Detection based on PA Segmentation

This method is referred to as M1 later in the text.

Summary

The idea behind this method is to come back to the definition of a clot, that is, a dark spot inside the PA. Starting from the segmentation mask of the PA, a combination of morphological and thresholding operators is used. The block diagram for this method is depicted in Figure 4.17.

\[\text{a mask is a structure that stores a boolean value (1 or 0) at every voxel position.}\]
Figure 4.17: Method M1 diagram. PA segmentation is computed from the CT data, then fed into a morphological closing. The difference between the closing and the original segmentation gives the clots candidate mask, whose intersection with the gray mask is computed.

**Processing**

By nature, the segmentation technique leaves the clots out of the resulting segmentation mask, because of their lower contrast. Depending on the clot position in the vessel, two cases show up:

- there is a contact between the clot and the boundary of the vessel,
- or there is no contact.

In the former case, the segmentation mask which corresponds to the blood-flow region will generally show a concavity for acute embolism, and sometimes but not always for chronic embolism. In the latter case, a hole is present in the mask. This distinction only makes sense when considering the neighborhood of a chosen region in the vessel, because in general, an embolus may have a very elongated shape, comparable to a tube inside a vessel and whose boundaries may touch locally. What has to be considered as well is that not all concavities in the segmentation mask do represent a clot, although this method will classify them as such if they have the same density as emboli (false alarms). The left/right bifurcation of the PA demonstrates that situation.

The first step in this approach is to decrease acquisition noise, yet preserving edges. Substantially, we want to get homogeneous regions in terms of voxel intensity for the different relevant classes (white, gray, black), thus avoiding salt and pepper kinds of noise. This is done through bilateral filtering [18] for each 2-D image in the volume. Median filtering could have been used as well.

The second step is to perform the **closing**, in the sense of mathematical morphology, of the segmentation mask to retrieve its concavities and holes.
Although it makes sense to have a variable size for the SE \[74\], depending on the size of the nearby vessels, we did not experience it. The shape of the structuring element is a 3-D box, providing shorter processing time, and its size is $12 \times 12 \times 12$ voxels (which was chosen subjectively and corresponds to a 12mm cube). The chosen size must be bigger than the largest embolus to be detected.

In the last step, we combine the difference between the original segmentation mask and the closed one, with the mask identifying the interesting gray regions using a voxel-wise AND operator. Both thresholds, low and high, are easily chosen because a contrast product is used, but may depend on scanner brand; we used 0.125 and 0.9 respectively.

The detection result of method \textit{M1} can be superimposed to the original image, showing the voxel-wise detection. This is illustrated in Figure. 4.18.

**Limitations**

This method has several limitations. Firstly, as it is based on PA segmentation, it cannot detect clots where the segmentation fails. Segmentation can fail for various reasons, in particular if one artery is completely obstructed by a clot. Secondly, another problematic case is when going farther in the arterial tree (smaller vessels), clots may no longer show up as concavities in the segmentation mask. We overcome these by method \textit{M2}.

### 4.5.2 Standalone Detection

This method is referred to as \textit{M2} later in the text.
Methods

Summary

The idea behind this method is to go through the entire image volume and find clots based on shape and density properties. Segmentation of the lung volume is performed to limit occurrences of false positives. Method diagram is in Figure 4.19.

![Method M2 diagram](image)

**Figure 4.19:** Method M2 diagram. Gray thresholding is performed on the CT data. The volume is then swept through from the 3 reference planes, performing connected component analysis. The three results are added, and morphology opening is performed (clean erroneous isolated detections).

Emboli Detection

If we go back to the definition of a clot in image terms, we could model its shape and color as a *gray cylinder* in 3-D space, at least locally. This definition remains correct even if the embolus completely obstructs the vessel. Of course, this means that there is an additional constraint in the medical protocol, that is that not only the arteries must be opacified, but also the veins and the other vessels of the thorax, otherwise they will be interpreted as clots because of their densities.

Such a cylinder has at least one intersection with one of the three reference planes (XY, YZ, or ZX) whose section is an ellipse with an axes ratio in $[1/\sqrt{3}, 1]$, the value of one corresponding to a circle. Connected gray shapes that meet this ratio criterion are sought in all the planes slicing the volume that are parallel to the reference planes. Practically, clots are not perfect ellipses and their sections have an upper bound (the size of the biggest artery) and this must be taken into account.

This is implemented by first segmenting interesting gray regions (as expressed in 4.5.1) and then by performing connected component analysis. For each obtained component, we measure its perimeter and area and accept it if both its area is smaller than the largest accepted clot section, and its compactness ($4\pi.\text{area}/\text{perimeter}^2$) is big enough (an ellipse with axes ratio of
1/\sqrt{3}$ corresponds to a compactness value of about 0.896. Three traversals of the image volume are made, one for each reference plane, and the results are combined with OR operator to provide the detection mask.

This detection technique obviously considers many shapes as emboli candidates, because the associated hypothesis is easily met. What we could observe is that many of the false positives were lying outside of the interior volume of the lung, especially in the thoracic bones. Segmentation of the lung volume is thus investigated to further constrain the detection and to decrease the false alarm rate.

**Lung Segmentation**

Many techniques can be considered to perform the lung segmentation task. Looking at the images, we see that air voxels of the lungs have low densities (black voxels in the window that is used), and a method based on thresholding and connected component analysis is able to handle that. The block diagram of the method that we used is presented in Figure 4.20. It extracts the lung region on each slice of the image volume (2-D processing). Two classes are defined, black and non-black voxels, and a simple region growing method is started from the outside border of each 2-D image. The reached voxels are removed from the black class, and connected component analysis is performed on that modified image. Only the components that correspond to the black class are kept, which are the air regions of the lung. Finally, a closing operator is applied with linear structuring element of infinite length for two orthogonal and oblique directions. That process is performed successively on each slice, to retrieve the entire lung volume segmentation (illustrated in Figure 4.21).

Oblique SE is used because the spine and the sternum must be avoided as much as possible, and morphology based on a horizontal SE would keep much larger parts of these structures.

**4.5.3 Fusion**

The two envisaged methods do not detect the same features. Combining both detection masks is made by a logical OR operator to preserve all information, at the expense of keeping all the false positives as well. This is dictated by the conservative requirements of a CAD in the health domain, where non-detections should be avoided. A comparison of the detection results of both methods is depicted in Figure 4.22.

---

*Flood-filling is used, the region continuously grows while the voxel density remains close to a reference value (black in this case).*
Figure 4.20: Diagram of lung segmentation method. Bilateral filtering is used to smooth the image, black regions are segmented and searched for connected components. Connected components having a contact with image borders as well as the one not meeting a size criterion are discarded. Closing is then applied to retrieve thorax vessels and organs.

Figure 4.21: Results of method $M2$. The first row shows the lung segmentation (axial, coronal, sagital slices), the second row the original image, and the last row the superimposed emboli detection mask.
Figure 4.22: Both detection methods are compared. Method \( M1 \) correctly identifies a clot that partially obstructs the artery in the bottom part of this image, but that \( M2 \) does not detect. To the contrary, the bigger clot above is only partially detected by \( M1 \), while \( M2 \) detects it accurately.

4.6 Validation Scheme

Assessing the quality of a detection method requires the intervention of experts, in this case, radiologists. Different levels of assessment can be considered, that range from looking at the exam as a whole, and reporting that a patient is affected by PE or not, to a voxel-wise classification (embolus or not), telling for each voxel of the dataset whether it is part of a clot or not. Our detection method in essence provides that voxel-wise information, but this may vary for other methods. Although it could be interesting to obtain reference data from radiologists as a voxel-wise classification, this is not practical due to the huge amount of work it requires (hours of manual labeling for a single dataset).

The problem of emboli detection is a recent one, hence no validation scheme have emerged lately. Some researchers discuss their results in terms of false alarms and non-detections [69, 70], and they seem to mark the clots by expert. However, the way this is made is not given.

For this reason, we decided to propose a new scheme for assessing PE detection methods [78]. This scheme should provide a compromise between high accuracy and being practical to realize with another requirement that the measures it gives should be easily understood by radiologists.
Figure 4.23: Creation of ground truth. A 3-D grid is superimposed on the image data, the radiologist uses the mouse to highlight boxes of voxels whenever a clot (at least partially) belongs to it.
4.6 Validation Scheme

4.6.1 Ground Truth Definition

Clots are structures that cannot be summed up to a location and size, but they can be elongated instead, and their size may extend from one millimeter in width and a few millimeters in length to a few centimeters in width and over 10 centimeters in length. This implies that one cannot introduce a scheme where reference data is represented as a list of emboli locations, as such locations do not make any sense. Because of this, our approach is essentially to partition the dataset into regions, or *boxes*, and to mark them individually for the presence of emboli.

In practical terms, the ground truth is defined as a partition of voxels boxes of an identical chosen size, which are individually marked by a boolean as containing a clot or not. The size of these boxes is an important parameter with lower sizes giving more precise information at the cost of more work needed to realize the ground truth and their size also has an effect on the results. We arbitrarily chose to use $16 \times 16 \times 16$ boxes. Practically, the radiologist navigates through the volume on which a 3-D grid is painted, and marks the boxes that contain clots (either entirely or not) with a tool running on his workstation. This is illustrated in Figure 4.23.

This presence mask is constructed from the output of the detection method on boxes of the same size as the one chosen in the ground truth. Both results, ground truth and detection algorithm, can then be compared.

Inter- and intra-radiologist variability has to be considered. Although it was not studied as part of this work, we can assume that this variability has different causes:

- Interpretation of image. Different radiologists interpret images differently.

- Border considerations. The delineation of clots along box borders can be ambiguous.

The first point is inherent to the detection process performed by experts. The second one is linked to the use of the tools. Proper marking for clots only slightly touching boxes of voxels is depicted in Figure 4.24. One possible way to address it can be done by providing a *score* or confidence instead of a boolean value stating the presence of clots.

---

9this term designates the known reference that is created by an expert.
4.6.2 Measuring Sensitivity and Specificity

In traditional binary classification, *false positives* (FP) are defined as the cases where the detection method detects a clot while the expert does not. A *false negative* (FN) is the dual situation where experts detect a clot and the method does not. *True positives* (TP) and *true negatives* (TN) are for the cases when both agree, respectively detecting a clot (positive) and not detecting a clot (negative). Sensitivity ($S_n$) and specificity ($S_p$) are defined by the following formulas:

$$S_n = \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad S_p = \frac{N_{TN}}{N_{TN} + N_{FP}}$$

where $N_{TP}$, $N_{TN}$, $N_{FP}$, and $N_{FN}$ respectively are the number of occurrences of true positive, true negative, false positive, and false negative classes.

We note that both concepts do not have the same importance depending on the context in which they are used. In the medical diagnosis context, non-detections should generally be avoided as much as possible, as this can directly lead to health related consequences and a bad specificity may impact the value of a CAD tool.

Another remark is that the sensitivity and specificity concepts are generally used patient-wise in medicine. This is not the case here and it should not be confused, as these concepts apply here to the boxes of voxels. They qualify the capacity of the method to detect emboli for each box of voxels and not for each patient.

These values can be computed when both the ground truth and detection results are available, and even used to compare various parameters of the same method. A qualitative method based on visualization is also possible by color coding the boxes of voxels depending on the correspondence of the method and the ground truth. This has been implemented and is shown in Figure 4.25.

Figure 4.24: The voxel boxes has to be highlighted even if the emboli is only partially contained. The difficult cases are marked by an arrow, where only a few voxels have lower densities and can easily be missed. The third image shows a case where a small vessel with a clot is across two boxes.
Figure 4.25: Comparison between the results of the method and ground truths on multiple patients. Voxel boxes are color coded, green is used when both the GT and the method agree, blue for false positives and red for non-detections. The upper-left image shows a combined view of axial, coronal and sagital planes, the right column is a set of coronal slices, and the rest are sagital slices.
Chapter 5

Experimental Results

5.1 Data Collection

Data for testing the various methods were collected at two sites, namely the Cliniques universitaires Saint-Luc in Brussels, UCL, Medical Imaging Department with Prof. Coche, and at the Centre hospitalier régional d’Orléans (CHRO), Radiology Department with Dr. Richard.

Both departments use different brands of CT scanners, Philips is used in Brussels, while Siemens is used at Orléans and they present specific image features. A total of 31 datasets were collected from several patients addressed for suspicion of PE, 11 from Orléans, 20 from Brussels. CT examinations were performed using the following parameters: 16 X 0.75 mm slice thickness, 0.6 mm interval of reconstruction, 120 Kv, 150 mAs, pitch of 0.95, matrix size: 512 x 512, 12 bits per voxel. 100 mL of non ionic contrast medium was injected through an antecubital vein at a rate of 4 mL/sec with variable delay in order to opacify the PA. Out of these images, 18 are diagnosed as PE positive by radiologists. Occurrences of clots are expressed in Figure 5.1.

These images are exported from the visualization workstations to CD-ROM in the DICOM format, generally as two or three datasets per CD. These data are anonymized and imported into our tools, where the proper window is selected, and they are then converted to a more convenient file format.

Emboli detection ground truths have been manually realized by Prof. Coche for all datasets and the detection algorithm is evaluated against them.
5.2 Results & Analysis

5.2.1 Overview

Comparison between the ground truth and the detection method id reported in Table 5.1 for the 18 PE-positive datasets. In terms of sensitivity and specificity, the method performs as follows: sensitivity of 88.4% is achieved on average, as well as a specificity of 98.9%, with respective standard deviations of 10.4% and 0.39%. The large standard deviation of sensitivity is the result of important non detections for a few cases (3 cases have their sensitivity below 80%). This is discussed in the next section. On the majority of cases, the sensitivity is high, and while it can still be improved, it already makes the method well suited for its original goal. The specificity is a little misleading, as it suggests that there are not many false alarms. This is due to the fact that the number of true negative boxes is large because it applies to a 3-D volume. Another way of considering the results is to look at how much data the radiologist has to examine if he only considers the boxes marked as positive by the method. This is the ratio between positive-detected boxes to the total number of boxes, that we called region of interest (ROI), and in our datasets, its value was in the range of $[0.85\% , 2.34\%]$. The higher the number of false alarms, the more often the radiologist will have to manually discard erroneous detections provided by the method. The entire
Table 5.1: Table of Detection Results Figures

<table>
<thead>
<tr>
<th>Dataset</th>
<th>PA</th>
<th>Lung</th>
<th>Sens. (%)</th>
<th>Spec. (%)</th>
<th>ROI (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>chro1</td>
<td>OK</td>
<td>OK</td>
<td>79.3</td>
<td>99.4</td>
<td>1.31</td>
</tr>
<tr>
<td>chro2</td>
<td>OK</td>
<td>OK</td>
<td>94.9</td>
<td>98.6</td>
<td>1.68</td>
</tr>
<tr>
<td>chro3</td>
<td>OK</td>
<td>OK</td>
<td>91.7</td>
<td>98.2</td>
<td>2.02</td>
</tr>
<tr>
<td>chro5</td>
<td>OK</td>
<td>OK</td>
<td>82.2</td>
<td>99.3</td>
<td>1.08</td>
</tr>
<tr>
<td>chro11</td>
<td>OK</td>
<td>OK</td>
<td>93.3</td>
<td>98.2</td>
<td>2.34</td>
</tr>
<tr>
<td>ucl1</td>
<td>OK</td>
<td>OK</td>
<td>93.3</td>
<td>98.6</td>
<td>2.06</td>
</tr>
<tr>
<td>ucl9</td>
<td>OK</td>
<td>OK</td>
<td>92.6</td>
<td>99.1</td>
<td>1.08</td>
</tr>
<tr>
<td>ucl10</td>
<td>OK</td>
<td>OK</td>
<td>57.9</td>
<td>99.3</td>
<td>0.97</td>
</tr>
<tr>
<td>ucl11</td>
<td>OK</td>
<td>OK</td>
<td>93.3</td>
<td>99.2</td>
<td>0.85</td>
</tr>
<tr>
<td>ucl12</td>
<td>OK</td>
<td>OK</td>
<td>100</td>
<td>98.9</td>
<td>1.15</td>
</tr>
<tr>
<td>ucl13</td>
<td>OK</td>
<td>OK</td>
<td>73.1</td>
<td>99.0</td>
<td>1.39</td>
</tr>
<tr>
<td>ucl14</td>
<td>OK</td>
<td>OK</td>
<td>84.2</td>
<td>98.7</td>
<td>1.37</td>
</tr>
<tr>
<td>ucl15</td>
<td>OK</td>
<td>OK</td>
<td>92.8</td>
<td>99.3</td>
<td>0.93</td>
</tr>
<tr>
<td>ucl16</td>
<td>OK</td>
<td>OK</td>
<td>92.8</td>
<td>98.5</td>
<td>1.59</td>
</tr>
<tr>
<td>ucl17</td>
<td>OK</td>
<td>OK</td>
<td>100</td>
<td>99.1</td>
<td>0.95</td>
</tr>
<tr>
<td>ucl18</td>
<td>Not OK</td>
<td>OK</td>
<td>96.9</td>
<td>99.0</td>
<td>1.04</td>
</tr>
<tr>
<td>ucl19</td>
<td>OK</td>
<td>OK</td>
<td>89.3</td>
<td>98.6</td>
<td>1.65</td>
</tr>
<tr>
<td>ucl20</td>
<td>OK</td>
<td>OK</td>
<td>84.0</td>
<td>99.3</td>
<td>0.93</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>–</td>
<td>–</td>
<td><strong>88.4</strong></td>
<td><strong>98.9</strong></td>
<td>–</td>
</tr>
<tr>
<td><strong>Std. Dev.</strong></td>
<td>–</td>
<td>–</td>
<td><strong>10.4</strong></td>
<td><strong>0.39</strong></td>
<td>–</td>
</tr>
</tbody>
</table>
processing is visualized on three datasets (ucl9, ucl15, and chro1) in Figure 5.2.

As previously stated in this text, comparing the results to those of Liang et al. [70] and Masutani [69] is not possible at this time and worthless because the validation schemes that are used in their application are unknown. However, access to the raw detection data of other algorithms could be used for evaluation purpose in this proposed scheme.

5.2.2 In-depth Look

The co-registration of the model to the patient image is the first step that must be accomplished. Locating the set of fiducial points is a straightforward process for the radiologist, and the thin plate spline evaluation always converged in our test cases. It naturally depends on the choice of fiducial points and a regular spread in the volume of the lung, around the vessels that are of interest, gives good results. A possible case of divergence may arise due to instabilities when fiducial points are too close to each other, but that was avoided in our experiments. The vessel centerlines of the model generally represent the actual centerlines in the patient image, with a possible inaccuracy of a few voxels when going farther from the fiducial points. That, in turn, may produce a shift towards one vessel or the other in the reconstructed boundaries in the segmentation step, but to an acceptable extent.

The segmentation of the PA is made after the co-registration of the model, and takes the seed point as a parameter. The fast marching propagation scheme could handle variable luminosity (irregular opacification) in the image without trouble and could recreate the missing boundaries between touching vessels in all cases, thanks to the model. In one of the 31 cases though, the segmentation of the PA failed, as shown in Figure 5.3. The heart and other vessels were segmented as well. Looking closely at the segmentation result and the iso-T surfaces (and their curvature) reveals that the contour first entered the vena cava through a segmental ramification of the PA, after the model usage was switched off. The contour, in turn, entered the aorta and the heart. False alarms for M1 were consequently increased, as most gray regions between vessels were detected as thrombi, yet the overall sensitivity was not affected.

Lung segmentation that is used to limit the false alarm rate of method M2 was found to be robust as no segmentation error occurred in our experiments. We could see however that some inaccuracies exist around the spine in this segmentation mask that are due to the morphology closing algorithm
Figure 5.2: The entire processing chain on three datasets. (a) one slice of the original dataset, (b) the registered model, (c) PA segmentation, (d) the lung segmentation, (e) $M1 + M2$ emboli detection mask, and (f) the comparison to the ground truth.
Figure 5.3: Segmentation failed in this case, because of a contact between a bifurcation of the PA and the vena cava.

that is used, leading to parts of the spine being segmented erroneously as lung volume.

In the detection step, we generally see that $M1$ properly detects the obvious clots within the main section of the PA without many false positives, but misses more and more clots as it goes farther in the PA. This is because the clot takes a bigger part of the section of the vessel when the size of the vessel decreases, and it does not create a concavity anymore. Moreover, the segmentation may stop early in vessels that are obstructed by clots, that present opacification defect, or due to limited resolution of CT, and thus leaving $M1$ useless for the remaining of the vessel. In contrast, method $M2$ detects clots everywhere as long as their shape are compatible with its hypothesis. $M2$ gives more false positives and fewer non-detections. Yet, $M1$ can detect clots that form concavities that $M2$ cannot detect when its shape criterion is not met, and this increases the overall sensitivity when both methods are combined. Detection rates for each methods considered separately actually depend on where the clots reside in the PA, and their ranges are given for illustrative purpose: $[0\%, 51\%]$ for $M1$, $[57\%, 100\%]$ for $M2$. In reality, $M2$ accounts for most of interesting detections for radiologists. This is because clots in the main parts of the PA are generally obvious to them, while peripheral clots generally require higher attention to detect because of their lesser visibility, and are generally out of reach for $M1$.

When looking more closely at detection results individually, the worst sensitivity of $57.9\%$ shows up for a case of massive embolism ($uc\ell1\ell0$ depicted in Figure 5.4), where most of the PA is obstructed. In that case, the segmentation stops in the right PA (left on the image) because the clots apparently completely obstruct the vessel. Moreover, hypothesis made by $M1$ and $M2$ about the maximal size of the clot are not met for bigger clots. We clearly see that the clot in the right PA is detected until its dimension becomes larger than the threshold, and the one in the left PA is completely missed.
Figure 5.4: Case with massive embolism, original slice (a), PA segmentation (b), and PE detection (c). (c) shows that the clot is correctly detected until its size becomes bigger.

for the same reason. In the context of a detection aid, this issue is probably not an important one, as these clots are easily seen by the radiologist. If that specific case is to be handled by the method anyway, thresholds may be adapted or other features may be used.

5.3 Implementation

5.3.1 Platform and Tools

The whole system is implemented in C/C++ languages for the algorithmic part and Objective-C/C++ for the user interactions part, on a regular computer, that is, a dual processor workstation equipped with two IBM PowerPC 970 running at 1.8GHz, with 2 GB of memory and specific OpenGL-capable display.

5.3.2 Computational Efficiency

Efficiency of computer softwares are often expressed in terms of processing time and memory requirements.

Processing Time

The inherent complexity of the algorithm as well as an efficient implementation of it was an important topic. Processing time is an important characteristic of the method, as it determines its suitability to clinical applications. Special efforts were made throughout the research and development phases to ensure that it would stay within acceptable limits. This is true especially in this application, when considering that PE can lead to death just a few hours after admission to the hospital.
Preliminary optimization in computer science is generally bad and must be avoided, but what is crucial is the choice of an algorithm whose complexity is compatible with the requirements. The fast marching algorithm was chosen exactly for this reason, and an efficient implementation was realized, still leaving room for further optimizations.

Processing times of the various algorithms are given in Table 5.2 and their value is proportional to the size of the dataset, expressed here for a 400-slice dataset. The total processing time is of the same order of magnitude as the radiologist time (10 minutes), making these methods suitable for clinical application.

### Memory Requirements

Special attention to memory requirements were taken into account from the start to keep memory allocations as low as possible, and computationally efficient structures were introduced to store the variables of various algorithms. In particular, the fast marching algorithm was implemented in a way that it allocates memory for its processing structures and the final solution only where the contours pass through, instead of the whole domain as is usually done. Indeed, allocating 5 double-precision floating points on a domain of $512 \times 512 \times 512$ requires 5GB of memory, which is more than the extent of the address space of 32-bit computers. In practice, the required memory in the implemented scheme was generally below the 500MB threshold.

### 5.3.3 Software Architecture

A summary of the software architecture is given here and illustrated on the diagram of Figure 5.5. External libraries are used for specific purposes, and specific development is made for others. The Visualization Toolkit (VTK),

---

1 which is a conservative estimate for the needs of the slice marching algorithm.

---

**Table 5.2: Processing Times for each algorithm**

<table>
<thead>
<tr>
<th>Method</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA Segmentation</td>
<td>35 seconds</td>
</tr>
<tr>
<td>$M1$ Detection</td>
<td>7 minutes</td>
</tr>
<tr>
<td>Lung Segmentation</td>
<td>12 seconds</td>
</tr>
<tr>
<td>$M2$ Detection</td>
<td>6 minutes and 45 seconds</td>
</tr>
</tbody>
</table>
provided by Kitware Inc., is an open source computer library for and is used here for visual, interactive rendering which is possible thanks to its OpenGL backend. That library is also used for importing DICOM data into our own data structures (IPCommon). The architecture of the software is implemented with Cocoa\(^2\) and is based on the MVC\(^3\) paradigm, and each tool is implemented as a plug-in of the main application. The IPCommon library is TCTS Image Processing library that has been augmented with the features required by this development. LSLib has been specifically developed as a fast marching and level set implementation and OpenCV is the Computer Vision library developed by Intel, which was used for the bi-lateral filtering and connected component analysis. XML is used for storing meta data for the algorithms, such as the segmentation seed point or model definition.

\(^2\)Collection of re-usable classes for application development provided by Apple
\(^3\)Model-view-controller paradigm that enforces code re-usability by separating data from visualization
5.3.4 Visualization and User Input

A graphical application, iMed Studio, was created for the visualization of the images and entering the various parameters required by the algorithms. Most visualization methods are based on VTK and the application graphical user interface (GUI) itself was prototyped and built with the Cocoa framework. The visualization of volume data is made with the re-slicing classes of VTK, providing sagital, coronal, and axial planes. The radiologist navigates in the volume with a computer mouse, and he can zoom in/out when accuracy is needed. A 3-D view combining the three planes is also rendered, including the results of methods, and the wireframe model can be superimposed to it to better figure out its relation to the dataset.

Parameter input is realized with the mouse, by locating a 3-D cross inside the volume. This serves for instance to initialize the seed point in the segmentation algorithm, to create the control points of the model, or to locate the fiducial points. Marking of the ground truth boxes is performed in a similar way, by holding shortcut keys while navigating in the image. An image of the running application is given in Figure 5.6.
Figure 5.6: The iMed Studio application is shown. Visualization is performed in the main window, while 3-D navigation and processing tools are configured in the right pane.
Chapter 6

Conclusion

6.1 Definition of the Subject

The subject of this work is restated here:

In the context of the recent shift of the modality used for PE detection towards CT scanners, the goal of this work has been to develop a semi-automated method for the detection of pulmonary embolism in opacified CT images, as well as to provide a scheme for evaluating its performance from expert information.

Special considerations about the computation time required by the method were crucial, recalling the emergency under which this condition may appear in hospitals and the quick response that has to be provided. Additional considerations include that the tools should be designed with practical concerns in mind, in terms of required equipment, and usability. Moreover, the method is not presented as a competitor to the expertise of radiologists, but it is instead offered as a complement to it, more in a gain of time perspective, especially in this current situation of insufficient number of radiologists.

6.2 Achieved Results

To achieve that goal, multiple approaches were considered. The first one, which consists of performing the segmentation of the pulmonary arteries, is investigated. The segmentation task itself is based on a proposed region growing algorithm named slice marching and is built upon the fast marching algorithm, advancing as slices inside vessels. Applications of this method show that it can handle segmentation of opacified vessels, but demonstrated
an incorrect shift towards touching vessels. An extension is then introduced to handle that specific case, which makes use of a model of a priori knowledge of vessel topology. This enables correct segmentation of the arterial tree, but at the price of manual, fiducial point-based co-registration of the model to the image of the patient.

The result of the segmentation algorithm is then used as a starting point for a detection method, which presents good results for larger parts of the vessels, but is found not to work properly for smaller, peripheral vessels, in which clots do not obstruct them in a concave way. For this reason, another detection method is proposed that makes simple assumptions on shape and color of the clots, and does not use the segmentation mask.

That second method is very sensitive and presents an important false alarm rate, especially for structures outside of the interior volume of the lungs. Segmentation of the lung is then performed and combined to the detection mask to successfully decrease the false alarm rate.

Finally, an evaluation method is needed to assess the performance of the method. No existing scheme for measuring the accuracy of PE detection methods was found in the literature, leading to a new one being proposed as part of this study. Ground truths, or reference data, are created that consist of marking a partition of the image volume as containing clots or not, by a radiologist. These ground truths are then used to derive false alarm and non-detection rates of the method, as well as its sensitivity and specificity. Results are analyzed both quantitatively and qualitatively thanks to specific visualization tools and a mean sensitivity of 88.4% and a specificity of 98.9% are attained. The non-detections, being generally rare, are identified as due to detection hypotheses that are not met, and the false alarm rate attributed to the simple hypothesis of method $M_2$. These encouraging results already set the method as appropriate for marking suspect regions in the volume.

6.3 Applications

Possible applications for this work include the development of a computer-aided diagnosis tool (CAD) into imaging workstations, an annotation system for pulmonary clots, a possible quantitative analysis of PE, as well as a tool for vessel analysis.

6.3.1 CAD Tool

The main application strongly suggested by the title of this study is to provide a computer aid to the radiologist. In the context of the decreasing
number of radiologists, the frequency of occurrences of PE, and its high mortality rate, such a tool could naturally find its way into the workstations or PACS.

As discussed in the introduction, this detection system should be used appropriately either as a preliminary detection means, or as a post-exam step (safeguard). Because the system is very conservative by design, the detection results could be passed directly to the radiologists as a color superposition for instance to help them to find clots in case of an oversight and that second reading could lead them to an improved final diagnosis. The case of small peripheral clots is the most interesting one, and the superposed diagnosis aid is illustrated in Figure 6.1. This can be thought of as an augmented display, or head-up display, for the second review. Although such an augmented view of the image will show all the false alarms along with the correct detections, this will catch the attention of the radiologist to suspect areas he/she might have missed on the first review. Moreover, this second review step may be performed much more quickly than the first one.

![Figure 6.1: Same CT slice shown as a regular image in (a), and augmented with detection results in red in (b). A second augmented review by the radiologist could easily be performed, drawing attention to suspect zones in the image. The true positives are marked with an arrow.](image)

Also, as stated before, this tool should not be used to decide when ambivalence exists between experts as the information it uses to decide is much more limited than the amassed experience of radiologists.

### 6.3.2 Annotation System

Another possible application is to use outputs of these methods to complement and qualify the images, and to store that information with the images themselves. The various by-products and results provided in these
algorithms plus manual input and verifications by radiologists constitute valuable knowledge that should not be lost. That annotation capability may then serve various purposes as the quick criterion-based retrieval and analysis of past exams, or even as a way to quantitatively compare exams \textit{a posteriori}, and track the evolution of the PE condition in patients.

6.3.3 Computation of PE Load Scores

PE load scores (Appendix A) are a mechanism for evaluating the importance of the PE condition. In practice though, these scores are not much used because of the required laboriousness of manually estimating them. However, combining the results of automatic detection methods to the model of knowledge and to the vessel section, which is available from the segmentation step, permits the computation of such scores and could make them more practical for daily use.

6.3.4 Vessels Analysis

The proposed slice marching algorithm provides by-products that can be used as well for the analysis of blood vessels. Measuring the section of a vessel for instance can provide a good indication of aneurysm (localized enlargement) or stenosis (abnormal narrowing), and one could consider the slice marching method presented in this work for inclusion in a vessel analysis package. Indeed, navigating inside a vessel and comparing the vessel section at various location would be a valuable tool in the diagnosis of such conditions.

6.4 Perspectives

6.4.1 Embolus Features

The current state of the method only relies on geometrical features of the clots to detect them. This is a good start, as this is the more obvious point that draws the radiologist attention. But other features can be used as well, which may contain additional information, such as texture of the clot, possibly searching in other density windows to possibly get additional clues about clot presence, Minkowski functionals to express topology features, etc.

6.4.2 Classification Approaches

The next logical step when additional image-based features are defined, is to complement such a system with a classifier with these features as inputs. The
method implemented so far is good at localizing areas of potential clots, but also introduces a non negligible number of false alarms. Recent advances in the classification domain, such as artificial neural networks or support vector machines, can be evaluated for decreasing the false alarm rate, making it even more suitable as a CAD tool.

6.4.3 Validation

Moving from a research project to a product to be clinically used involves profound and thorough testing. One perspective is thus to build larger databases for the validation of the algorithm. Quantitatively studying the gain it provides to radiologists is probably the next step that may even help in refining the procedure and the tools.
Bibliography


[65] Sethian’s website. http://math.berkeley.edu/~sethian


Appendix A

PE Load Scores

PE load scores are given here. They have been extracted from [3].

Miller Score

The right PA has nine major segmental branches (three to the upper lobe, two to the middle lobe, and four to the lower lobe). The left PA has seven major branches (two to the upper lobe, two to the lingula, and three to the lower lobe). The presence of a filling defect or obstruction in any one of these branches scores 1 point. A filling defect proximal to segmental branches scores a value equal to the number of segmental branches arising distally. The maximum score is 9 for the right lung and 7 for the left lung. The maximum possible CT obstruction score is 16 for both lungs. In addition, the score originally evaluated the effect of embolism on PA flow, which currently cannot be assessed with CT pulmonary angiography.

Walsh Score

The maximum score is 18 for both lungs. Different scores are given for filling defects and obstructions and for anatomic locations. The following guidelines govern quantification of embolic abnormalities: (a) Abnormalities in a single segmental PA receive a total score that does not exceed 1, regardless of the type or number of abnormalities. (b) The total maximum score is 3 for abnormalities in a single upper lobar region, 2 for abnormalities in the middle lobe or lingula, and 4 for abnormalities in the lower lobes. (c) Obstructions in central anatomic regions receive scores according to the vessel involved. (d) If the total score for one lung is greater than 4 without considering filling defects in central regions, the central filling defects are ignored.
All filling defects in a single central region, whether single or multiple, receive a score of 3. (e) If a single vessel contains both a filling defect and an obstruction, only the obstruction is scored. (f) The sum of scores for all abnormalities in one lung may not exceed a value of 9. The maximum CT obstruction score is 18.

**Qanadli Score**

The arterial tree of each lung is regarded as having 10 segmental PAs (three to the upper lobes, two to the middle lobe or lingula, and five to the lower lobes). The presence of an embolus in a segmental PA is scored as 1 point, and emboli at the most proximal arterial level are scored a value equal to the number of segmental PAs arising distally. To provide additional information on the residual perfusion distal to the embolus, a weighting factor is used for each value (0: no defect, 1: partial occlusion, and 2: partial occlusion). An isolated subsegmental embolus is considered a partially occluded segmental PA and is assigned a value of 1. The maximum CT obstruction index is 40.

**Mastora Score**

The scoring is applied to five mediastinal PAs (PA trunk, right and left PAs, and right and left interlobar PAs); six lobar PAs; and 20 segmental PAs (three in the upper lobes, two in the middle lobe or lingula, and five in the lower lobes). The CT severity score is based on the percentage of obstructed surface of each central and peripheral PA section and uses a 5-point scale (1: 25%, 2: 25%-49%, 3: 50%-74%, 4: 75%-99%, 5: 100%). A central score (mediastinal and lobar PAs), a peripheral score (segmental PAs), and a global score (central and peripheral PAs) can be calculated. The maximum CT obstruction score is 155.
Appendix B

Computation of the Speed of Propagation

The code of the speed function is given here for illustration purpose.

The speed of propagation of the fast marching algorithm is implemented as a callback function whose commented code is given here. This function is invoked each time the main algorithm needs the value of speed of propagation at a given \((x, y, z)\) voxel. The computation of the potential function for the vessels according to Equation 4.23 is made separately.

```c
float speedCallback(void *data, int x, int y, int z)
{
    IMBlockSegmenterData *self = (IMBlockSegmenterData*)data;
    float res = 0.0;
    LSImageIndex index;

    if(self && self->speed)
    {
        index = IP_GET_INDEX(x,y,z,self->speed->param.inc);

        switch(self->speed->sampleType)
        {
            case IP_SAMPLE_TYPE_UINT8:
            {
                unsigned char *p = (unsigned char*)self->speed->data;
                res = ((float)p[index])/255.0;
            }
            break;
            case IP_SAMPLE_TYPE_FLOAT32:
            {
                float *p = (float*)self->speed->data;
                res = p[index];
            }
            break;
            default:
            GL_LOG("sample type not supported");
        }
    }
    return res;
}
```
data is a user-provided parameter that, in this case, points to a structure that holds all the necessary segmentation information. The speed of propagation that comes from the image is accessed on line 16 or 22 depending on the type that is used in the image.

```c
if(self->useModel && self->useInitialModelGuide)
{
    float p[3], f[3], t[3], p[3], i, j, k;
    // 1 voxels margin
    if(p[0] < -1.0) res = RESIDUAL_SPEED;
}
```

The segmenter is driven by a state machine, with 3 states. The boolean useInitialModelGuide is set to true when the segmentation starts, and in that case, the voxels are tested on line 37 against the offset plane equation to determine whether they are on the good side of it. If they are not, the speed is set to near zero to prevent propagation. The offset plane equation is derived from the tangent to the first Bézier segment of the model. What has to be noted at this time is that the voxel of interest has 8 surrounding voxels in the potential map.

```c
if(self->useModel && self->state != IM_STATE_SMALL_VESSELS)
{
    float f[3], t[3];
    int p[3], i, j, k;
    GLsint32 *size = self->potentialMap->param.size;
    f[0] = x; f[1] = y; f[2] = z;
    for(i=0; i<3; i++)
    {
        f[i] = IP_TRANSFORMCOORD(f[i],
            self->speed->param.step[i],
            self->potentialMap->param.step[i]);
        p[i] = floorf(f[i]);
        t[i] = f[i] - p[i];
    }
}
```

Then a check is performed on line 39 to determine whether we want to use the model and that we are not in the initial state nor in the small vessels state. Then we express the location of the voxel in the coordinate space of
the potential map (sub-sampled grid), separating the integer and decimal parts on line 52 for later computations.

```c
if(IP_NEIGHBOR_BOUND_TEST(0, p, size)
 && IP_NEIGHBOR_BOUND_TEST(1, p, size)
 && IP_NEIGHBOR_BOUND_TEST(2, p, size)
 && IP_NEIGHBOR_BOUND_TEST(3, p, size)
 && IP_NEIGHBOR_BOUND_TEST(4, p, size)
 && IP_NEIGHBOR_BOUND_TEST(5, p, size))
{
  GLsint32 *inc = self->potentialMap->param.inc;
  GLsint32 ind = IP_GET_INDEX(p[0],p[1],p[2],inc);
  MAPotentialValue *pointer = self->potentialMap->data;

  { int ves;
    float d[2][2][2][2], d1, d2;
    if(pointer[ind].potID[0] == 0
      || pointer[ind].potID[1] == 0)
      goto currentval;

    // find master id:
    int masterid = 0, nearestid, masterind, m[3],
    cind, fcount = 0, ocount = 0;
    for(i=0;i<3;i++) {m[i] = round(f[i]);}
    nearestind = IP_GET_INDEX(m[0],m[1],m[2],inc);
    nearestid = pointer[ind].potID[0];

    masterid = self->seedBranch+1;
  }

  for(i=0; i<16; i++)
    d[(i>>3)%2][(i>>2)%2][(i>>1)%2][i%2] = MAXFLOAT;

  for(ves=0; ves<2; ves++)
  {
    for(k=0; k<2; k++)
    {
      for(j=0; j<2; j++)
      {
        for(i=0; i<2; i++)
        {
          cind = ind + i*inc[0] + j*inc[1]
            + k*inc[2];
```
if (pointer[cind].potID[ves] == masterid) {
    d[0][i][j][k] = pointer[cind].potVal[ves];
    fcount++;
} else {
    d[1][i][j][k] = pointer[cind].potVal[ves];
    ocount++;
}
}
}
}  
}
}
}

Here we store the value of the potential of the closest vessel, and the next one, of the 8 surrounding voxels of the potential map. The distinction between the two is made by comparing the id’s to the masterid (PA).

for(ves=0; ves<2; ves++) {
    for(k=0; k<2; k++) {
        for(j=0; j<2; j++) {
            d[ves][0][j][k] = (1.0-t[0]) * 
            d[ves][0][j][k] + 
            t[0] * d[ves][1][j][k];
        }
        d[ves][0][0][k] = (1.0-t[1]) * 
        d[ves][0][0][k] + 
        t[1] * d[ves][0][1][k];
    }
    d[ves][0][0][0] = (1.0-t[2]) * 
    d[ves][0][0][0] + 
    t[2] * d[ves][0][0][1];
}

d1 = d[0][0][0][0];
d2 = d[1][0][0][0];

Both vessel potentials are tri-linearly interpolated to the d1 and d2 variables.

if(self->modelUsage == IM_MU_STAYNEAR 
|| self->modelUsage == IM_MU_MIXED) {
    if(nearestid == self->seedBranch+1) {

```
if(d1 > D_MAX) res = RESIDUAL_SPEED;
else
{
    if(d2 > D_MAX) res = RESIDUAL_SPEED;
}

In the _stay near_ mode, only the id of the nearest vessel is tested against the id of the PA, and the speed is set to zero when it is not the PA.

if(self->modelUsage == IM_MU_BOUNDARIES
  || self->modelUsage == IM_MU_MIXED)
{
    if(fcount != 8) res = RESIDUAL_SPEED;
    if(fcount == 8 && ocount == 8)
    {
        if(d1 != MAXFLOAT && fabs(d2-d1) < LAMBDA)
            res *= RESIDUAL_SPEED;
    }
}

In the _boundaries creation_ mode, the values of both potentials are compared and the speed is modified (to near zero) when the difference is smaller than a given amount, which expresses the recreated boundary width.

if(res < self->speedThreshold) res = RESIDUAL_SPEED;

Finally the value of the speed of propagation is returned, and the fast marching algorithm can continue to evolve the contour.
Index

acute, 8
angiography, X-ray, 10
applications, 96
bifurcation detection, 56
boundary value formulation, 34
cardiovascular system, 5
chest anatomy, 5
chronic, 8
clot, 6
clot detection, 31, 71
contrast medium, 14
CT scanner, 16
CTA, CT angiography, 11
diagnosing PE, 10
DICOM, 10
distance map, distance transform, 33
domain of dependence, 37
Eikonal equation, 35
embolism, 1
embolus, 8
entropy solution, 36
false alarms, 80
fast marching, 28, 37
fiducial points, 64
filtering, 22
freezing, 55
fuzzy logic, 26
ground truth, 79
GVF, 30

HU, Hounsfield unit, 17
image acquisition, 13
imaging protocol, 13
initial value formulation, 34
initialization of propagation, 46
interface propagation, 34
LESI, 10
level set, 28, 38
load scores, 9, 109
lung anatomy, 5
lung segmentation, 75
M1 method, 71
M2 method, 73
Mastora score, 110
microelectronics, 44
Miller score, 109
model co-registration, 61
model usage, 63
model, use of a priori, 57, 59
mortality rate, 5
narrow band level set, 41
nearby vessels, 37
non-detection, 80
nuclear medicine, 10
numerical scheme, 50
PACS, 10
PE, pulmonary embolism, 6
perfusion, 10
prior knowledge, 29, 57, 59
Qanadli score, 110
ray propagation, 26
region growing, 23
related work, 21
results analysis, 83
robotics, shortest path, 43
seed point, 67
sensitivity, 80, 84
slice marching, 53
snake, 30, 12
software architecture, 90
specificity, 80, 84
speed of propagation, 48
state of the art, 21
symptoms, 6
tobogganing, 31
touching vessels, 26, 57
TPS, 61
TPS, thin plate spline, 33
upwind scheme, 37
validation scheme, 77
ventilation, 10
vessel enhancement, 23
vessel features, 56
vessel potential function, 66
vessel segmentation, 21
Walsh score, 9, 109
wave propagation, 23
weak solution, 38
window level, 17
window width, 17