Advanced Computational Methods in Multi-view Medical Imaging

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Liron Yatziv

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

Guillermo Saprio

September, 2012

Acknowledgements

First and foremost, I would like to thank my adviser, Professor Guillermo Sapiro, for his advice and mentorship throughout the past nine years. His willingness to let me work in his lab during my Master studies has helped me refine my own research interests and later convinced me to pursue a Ph.D as an external student.

I appreciate the effort of committee members, Professor Nikolaos Papanikolopoulos, Professor Emad Ebbini, and Professor Daniel Boley. I acknowledge also Professor Mitchell Luskin who participated in my prelim.

I would like to thank Dr. Chenyang Xu from Siemens AG for his support to pursue a Ph.D and his valuable guidance in what achievements matter most. Numerous interactions with my colleagues from Siemens Corporate Research including Dr. Rui Liao, Dr. Alexander Brost, Dr. Yiyong Sun and many other, that as well inspired me and my work. In particular, special thanks to Dr. Hari Sundar for his research guidance and collaboration with the cardiac and respiratory cycle but most of all being a great friend. I thank Dr. Norbert Strobel for providing me the clinical background. I also thank Dr. Saurabh Datta, Julian Ibarz, and Mathieu Chartouni for their collaboration on our joint publications. I recognize Siemens AG for providing me the medical images and data that is prerequisite for this work.

Finally, I would like to thank my loving wife for her unconditional support during my graduate studies. To my two little girls that inspire me everyday. To my sisters for being there for moral support and motivate me to complite this work. To my parents, I thank you for your ever-present advice and encouragement. Without your support, this work would not have been possible. This thesis is for you.

Abstract

Modern automatic imaging methods can help physicians to visualize physical details in the patient's body. Such methods must be efficient and reliable to allow making the right decisions during clinical intervention. This thesis presents three medical challenges involving multi-view fluoroscopy and proposes corresponding innovative computational methods.

The first method addresses challenges of multiple catheter detection and tracking in fluoroscopic guided interventions. The proposed algorithm incorporates the clinical setup constraints and uses a novel computationally efficient geodesic properties framework that traces the catheter from the sheath to the tip. The method was validated on 1107 fluoroscopic images, demonstrating robust multiple catheter tracking.

The second method helps avoid serious complications during a cardiac ablation procedure by extracting esophageal silhouettes and reconstructing a 3D model from fluoroscopic views. The thesis describes the key algorithmic components and suggests how the result could help the clinical workflow.

The third method presents automatic image-based cardiac and respiratory cycle gating of image sequences. We developed a novel algorithm for an automatic detection of the moving organ directly from images. The algorithm was validated on X-Ray angiographic and intra-cardiac echocardiography images. The method is very fast and its effectiveness was demonstrated on 3D intra-cardiac ultrasound sequences.

All three methods are computationally efficient, robust and represent the state-of-the-art of medical imaging challenges.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Many modern medical techniques and processes create image of the human body for clinical purposes. Those images are referred to as "medical images" and are in many cases synthetically generated to help clinicians visualize the data which represent physical details and properties of the body. As medical imaging technology advances, the variety of techniques of medical images increases: Projection radiography (X-ray), Tomography (CT scans), Ultrasound, Magnetic resonance imaging (MRI) and Positron emission tomography (PET) are just examples of both 2-dimensional and 3-dimensional images.

Currently, most images are not processed automatically and the physicians use mostly their brain power to interpret them. Yet, as talented as they are, they have limited ability to accurately analyze some of those images, whether those image are 2D images or 3D/4D scans. Some of the 2D images are a projection of the treated organ which is of course 3D. Those images are transparent in nature which makes it difficult to estimate the depth of objects. The 3D scans (e.g. CT, MRI, PET) generate data volume which is difficult to visualize in its raw format. Advanced rendering technology can help visualize them, yet this is typically an interactive process and requires some skill to manage to observe the desired layer of the volume. Those are just a few examples where the information is there but us human need much time to recover it out of the image accurately. Therefore, the field is in great need for automatic computation methods, tailored for the clinical needs which would analyze automatically medical images. This would help physician to get the most out of the images, accurately and promptly, allowing them to make the right decisions doing a procedure or while the patient is

there.

Due to the nature of the field, any method developed must be accurate, quick and reliable in order to avoid misleading the physicians. Even with the advance on computer processing power and GPUs, achieving usable accurate computational methods is a challenging problem. With such a new variety of data there is a great need to come-up with innovative practical solutions.

In clinic environment, I found only a few imaging applications based on smart computational methods to process the images. Here are three examples of the few successful technologies that made it eventually products in hospitals:

1. 3D volume reconstruction from sequence of X-Ray images acquired when rotating the X-Ray system around the patient [1] and [2].

2. 3D segmentation of the left Atrium given a single point in a CT or MRI volume [3].

3. Rigid registration between two volumes such as CT-CT or CT-MRI [4].

Yet, apart for a few, most available methods just don't make it producible due to the tough requirements of both accuracy and reasonable run-time. The run-time requirement depends on the application and the time of use in clinic. For example, for diagnostic applications, it would be acceptable to get a result within a minute, while interventional applications require results to be available within a few seconds, if not real-time.

Image processing techniques can easily be too computational expensive for real world applications. Researchers in image processing and computer vision are constantly searching for more efficient mathematical models and computational algorithms, however, new methods are yet to be developed for practical applications; for example, one most popular problems in computer vision is face recognition, yet although there is a great need for security and many of algorithms developed in the area, no practical method is both reliable and computationally practical to be an bases for a successful application.

A common use of x-ray in hospitals other than visualization of bone is the visualization of tissue and guidance of medical instruments. Fluoroscopy is the common name for

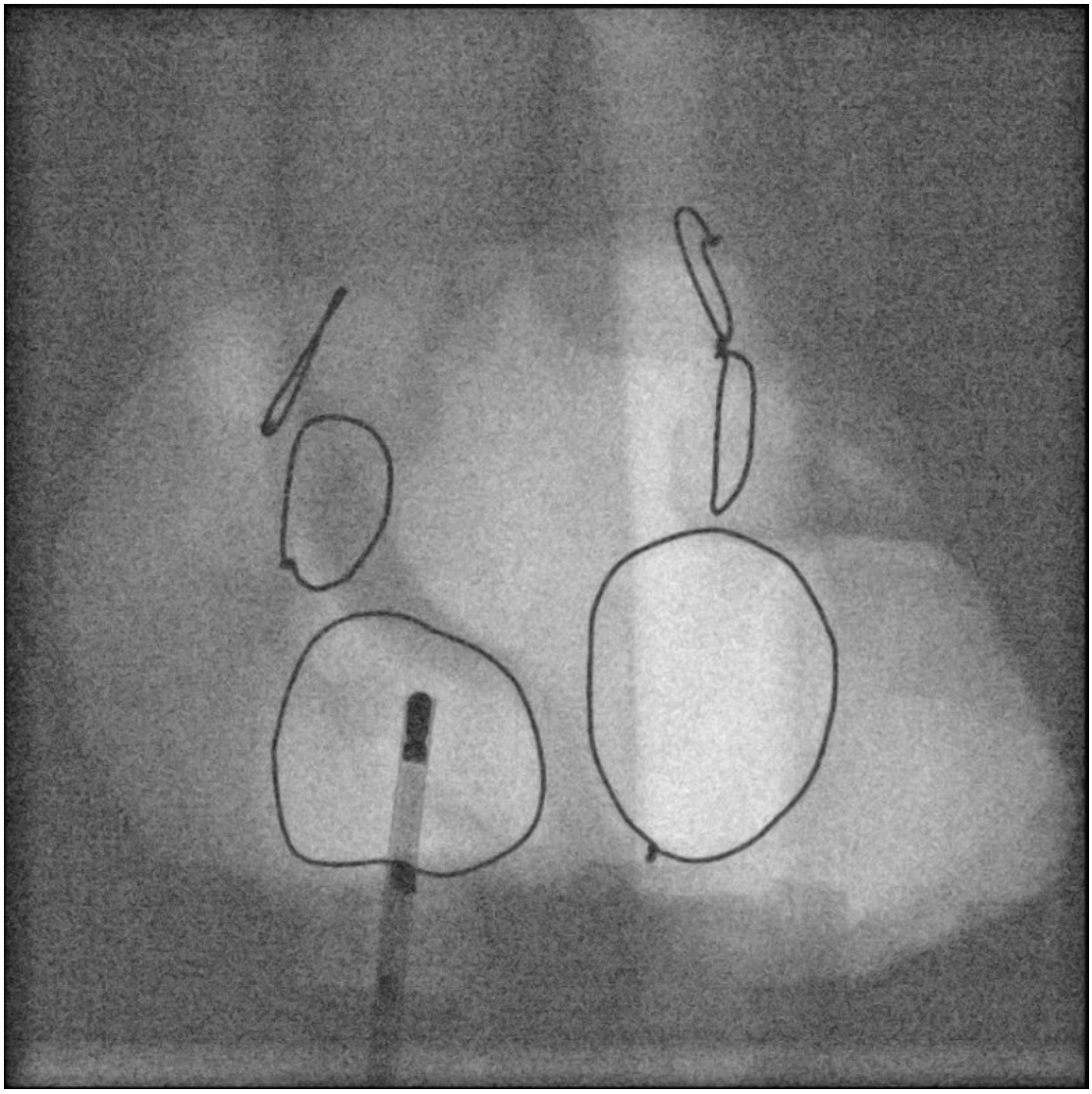


Figure 1.1: Sample fluoroscopic image. This image of a synthetic simple heart model. Because the model is rather uniform it is easy to see the quantum noise, which is typical to fluoroscopic images using low dose of radiation

production of real-time x-ray images, instantaneously made visible during examination. The images generated from fluoroscopy are called fluoro images. They come handy in the clinic, since they are created in real-time and the physician can easily rotate and change the position the X-ray field during the examination. This makes fluoroscopy a very powerful diagnostic tool. However, due to the length of the fluoroscopic examinations, the radiation exposure rate must be kept very much lower than in common X-Ray usage. The visible fluoroscopic image is therefore formed using much less X-ray photons. This will result in images that suffer from a much higher level of quantum noise making them harder to automatically process as can be seen in Figure 1.1.

In medical setting, an X-Ray system that is commonly used for generate x-ray images is referred to as a "C-Arm". It consists of a X-Ray source and on one side and a detector that records the image on the other, connected using a robotic "C" shaped and therefore its name, see Figure 1.2. The arm can flexibly rotate around the patient and change the distance between the source and the detector, allowing views from many directions without the need to move the patient.

The major companies that provide modern systems are General Motors, Siemens, Philips, and Toshiba. The device can be mobile or fixed to the ground and the ceiling. When two C-Arms are joined, two fluoro images can be taken nearly in the same time from different angles. Such a system is referred here as a "bi-plane" system.

Siemens has kindly agreed to provide me bi-plane fluoro images for my research. Therefore, the research would focus on the Siemens system. Fluoro images are stored as special medical DICOM type files. Those files contain detailed information on the physical setting of the system during acquisition of the fluoro which include zoom, angles and distance from source to detector. That information can be used to calculate bi-plane geometry between images providing the projection matrices, since the X-Ray source can be considered a pin-hole camera.

Catheters are routinely inserted via vessels to cavities of the heart during fluoroscopic image guided interventions for EP procedures such as ablation. EP clinical applications can benefit from fast and accurate automatic catheter tracking in the fluoroscopic images. The typical low quality in fluoroscopic images and the presence of other medical instruments in the scene make the automatic detection and tracking of catheters in clinical environments very challenging. Towards the development of such an application,

Figure 1.2: A photo of C-arm system. The system has a C shaped arm that can be electronically rotated around a center point (called the isocenter) using a robotic. On one end of the arm is the source that generates the x-ray. On the other end is the detector which is used to generate the image.

a robust and efficient method for detecting and tracking the catheter sheath is developed. The proposed approach from our publication [5] is presented in chapter 2 exploits the clinical setup knowledge to constrain the search space while boosting both tracking speed and accuracy, and is based on a computationally efficient framework to trace the sheath and simultaneously detect one or multiple catheter tips. The algorithm is based on a modification of the fast marching weighted distance computation that efficiently calculates, on the fly, important geodesic properties in relevant regions of the image. This is followed by a cascade classifier for detecting the catheter tips. The proposed technique is validated on 1107 fluoroscopic images acquired on multiple patients across four different clinics, achieving multiple catheter tracking at a rate of 10 images per second with a very low false positive rate of 1.06%.

Cardiac ablation involves the risk of serious complications when thermal injury to the esophagus occurs. Chapter 3 proposes to reduce the risk of such injuries by a proactive visualization technique, improving physician awareness of the esophagus location

in the absence of or in addition to a reactive monitoring device such as a thermal probe. This is achieved by combining a graphical representation of the esophagus with live fluoroscopy. Toward this goal, an automated method is presented to reconstruct and visualize a 3D esophagus model from fluoroscopy image sequences acquired using different C-arm viewing directions. In order to visualize the esophagus under fluoroscopy, it is first bio-marked by swallowing a contrast agent such as barium. Images obtained in this procedure are then used to automatically extract the 2D esophagus silhouette and reconstruct a 3D surface of the esophagus internal wall. Once the 3D representation has been computed, it can be visualized using fluoroscopy overlay techniques. Compared to 3D esophagus imaging using CT or C-arm CT, our proposed fluoroscopy method [6] requires low radiation dose and enables a simpler work-flow on geometry-calibrated standard C-arm systems.

In chapter 4, presents our proposed method, as published in citesundar2009, to detect the current state of the quasi-periodic system from image sequences which in turn will enable us to synchronize/gate the image sequences to obtain images of the organ system at similar configurations. The method uses the cumulated phase shift in the spectral domain of successive image frames as a measure of the net motion of objects in the scene. The proposed method is applicable to 2D and 3D time varying sequences and is not specific to the imaging modality. Experiments demonstrate its effectiveness on X-Ray Angiographic and Cardiac and Liver Ultrasound sequences. Knowledge of the current (cardiac or respiratory) phase of the system, opens up the possibility for a purely image based cardiac and respiratory gating scheme for interventional and radiotherapy procedures.

Finally, Chapter 5 summarizes the key results outlined in this thesis and provides some concluding remarks.

Chapter 2

Towards Multiple Catheters Detection in Fluoroscopic Image Guided Interventions

2.1 Introduction

Electrophysiology (EP) involves diagnosis as well as correction of cardiac arrhythmias using procedures such as catheter ablation. The underlying cause of cardiac arrhythmia is an abnormal electrical activity in the heart, often resulting in irregular rhythm or beats, affecting millions of people around the world [7]. Ablation is a treatment that burns (cauterizes) cells to alter electric conduction pathways, resulting in the correction of rhythm abnormalities in patients. During an electrophysiology procedure, the electrophysiologist threads special electrode catheters (long, thin, flexible wires) into the heart, routinely through a vein or artery of the upper leg. Once the area of the heart responsible for the arrhythmia has been determined, a special ablation catheter that can deliver radio frequency energy is used to burn or ablate the target site. Catheter ablation is an outpatient procedure that normally takes only a couple of hours to complete and has few complications.

In some cases, a routine treatment involves ablation of 3 or 4 specific areas in the left atrium near the openings of the 4 pulmonary vein, Figure 2.1. The abnormal electrical

activity usually originates at the pulmonary veins location, and electrically isolating these veins with ablation keeps the cardiac arrhythmia from occurring [8].

Physicians navigate the introduced catheters based on x-ray or fluoroscopic images, Figure 2.2. Fluoroscopy used in these procedures use a low radiation 2-D x-ray images that allows catheters visualization, while the anatomical details of the heart are barely visible, making it hard to know the location of the catheters in reference to the anatomy. Software and visualization products such as Biosense Webster CARTO©, Medtronic LocaLisa®, Stereotaxis NAVIGANT™, and St. Jude EnSite NavX™ have been developed for interventional procedures that assist and guide physicians. Such products rely on the location of the catheter for navigation, electrocardiac mapping and other activities. In current clinical practice, a variety of methods are available for finding the catheter location in 3-D. One such method is to use magnetic tracking, being the main drawback of this approach the high cost associated both with equipment and catheters. Catheter localization directly based on the fluoroscopic images allows for a low-cost solution, without the need for additional equipment. From the location in 2-D images, a 3-D location can be found by analyzing distances on the visible catheter [9]. In some clinics, two real-time fluoroscopy images are available from two different angles. Such a setup allows to triangulate and achieve millimeter accuracy in 3-D space [10].

In order for a tracking method to be usable as an interventional tool, it must be both computationally efficient and accurate. Routinely, fluoroscopic images in EP are acquired at about two to ten images per second, depending on the system settings. Therefore, tracking at the rate of several images per second would be considered as an acceptable efficiency. The accuracy requirements depend on the specific application, and millimeter accuracy is roughly comparable with the currently available magnetic navigation systems [11, 12, 13] while others [14, 15], have found 2 millimeter accuracy to be sufficient. For ablation catheters, the most vital information to be tracked accurately is the catheter's tip, since it performs the burning (cauterizing). A very low false positive rate is also important to reduce any risk of inaccurate or insufficient treatment. The work presented here was specifically aimed at such application where catheter ablation is used to correct arrhythmia. It is a common practice that more than one catheter are used simultaneously, therefore, ability to track multiple catheters is also an important requirement satisfied by our developed system. For example, tracking the coronary

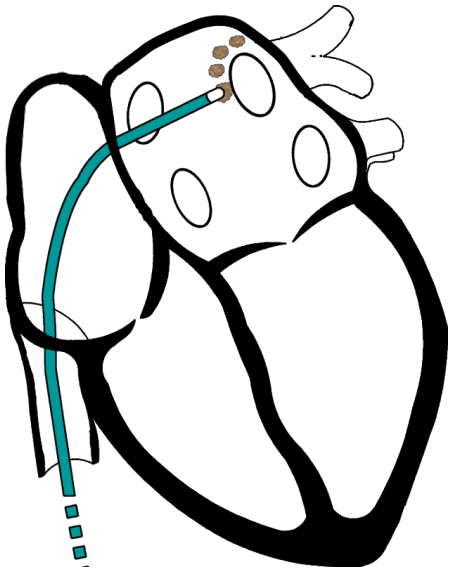


Figure 2.1: Illustration of the heart and the ablation catheter entering the left atrium. In left atrium ablation cases, one way of inserting the ablation catheter is through a vessel in the groin area. The catheter is then threaded via the inferior vena cava to the right atrium. Finally, the catheter passes through an atrial septal puncture (done using a needle) and reaches the left atrium with access to the pulmonary vein. In atrial fibrillation patients the tissue around the pulmonary vein is cauterized using the ablation catheter as shown here.

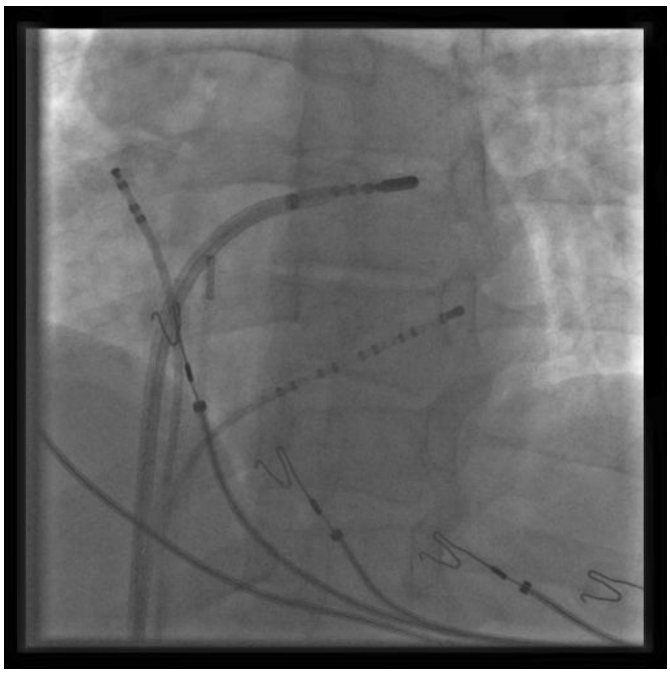


Figure 2.2: Example of a fluoroscopic image which contains several catheters coming from the bottom of the image (inferior vena cava). Catheters often have a metallic tip and therefore appear darker. The upper catheter facing to the right is the ablation catheter which is in the left atrium. Below it, also facing to the right, is the Coronary Sinus (CS) catheter. In this image, there is another catheter facing to the upper left and ECG wires with 'y' shape tip which are attached to the patient's skin.

sinus catheter can be used for transseptal puncture guidance [16] and respiratory motion correction [17]. There are several studies on the feasibility and significance of such a method for fluoroscopic based guidance and mapping. Phillips has a product called the "EP Navigator" that requires a user operator to indicate the position of the ablation catheter on the fluoroscopic image (so called "Point Tagging"). It was evaluated during a catheter ablation [18], while [19] found it feasible to use their automatic catheter tracking method in a clinical environment.

Many tracking methods are based on temporal assumptions that constrain the frame by frame detection problem by limiting the search domain. Such assumptions are typically Markov model based, where the previous known locations or results are combined with trajectory assumptions to limit the tracking domain. In the EP clinical scenario handled in this manuscript, assumptions based on a recent location would not hold due to rather long time gaps between two successive fluoroscopic sequences. A typical clinical workflow involves many short sequences taken at about two to ten frames per second and for only a few seconds. Continuous fluoroscopic acquisition is avoided to minimize the radiation exposure, both to the patient and the interventionist [20]. Therefore, a successful tracking method still needs to include a localization method at the initial frames of each sequence, as here developed. Once a location is available, temporal constraints could be used if needed.

De Buck et al., [21], introduced a method which required manual segmentation of the catheter. Fallavollita et al., [22], developed a catheter tip detection algorithm based on thresholding the fluoroscopic images. Their method performed poorly on low radiation noisy fluoroscopic images commonly used in the clinic. Frenken et al., [23], proposed a method to automatically detect the catheters in fluoroscopic images by using steerable tensor voting [24], in combination with a catheter-specific multi-step extraction algorithm. Their method managed to find the tip in 80% of their evaluation set. However, the authors reported that the implementation was too slow for clinical use. More recently, De Buck [19] tracked the ablation catheter tip using a fixed template-based registration method in conjunction with a Kalman filter, requiring periodic manual initialization. In [25], the authors propose to track a lasso catheter (a circular mapping catheter seen in Figure 2.9e) to estimate the 3-D respiration motion assuming an elliptical lasso. Molina et al., [26], used active contours (snakes) to reconstruct catheter

paths from bi-plane angiograms. In their method, the user interactively creates the 2-D splines. Ma et al., [17] using an electrode detector and a cost function to track only the CS catheter. Brost [27] tracked the CS catheter based on a model. Several methods, e.g., [28, 29, 30], were introduced in the related topic of guidewire tracking. Guidewires differ from catheters in appearance, their usage, clinical environment, and their limited motion, and therefore those methods can not be assumed as an obvious extension to the catheters detection problem here addressed, which, e.g., at times includes occlusion. In addition, we found no work involving simultaneously tracking more than one catheter.

In this work we propose a method to address challenges of multiple catheter tracking that can be used in a clinical setting. The method focuses specifically on the tip of the catheter, the tip location being the most vital information needed for electrophysiology applications, since the burning (cauterizing) happens at the tip. The proposed algorithm incorporates the clinical setup constraints and uses a novel *computationally efficient geodesic properties framework* that traces the catheter body all the way to the tip. As mentioned above, due to the nature of the short fluoroscopic acquisition, the proposed detection technique does not rely on previous detection results, and therefore can handle tracking independently of frame rate or the elapsed time from the previous sequence.

The rest of this chapter is organized as following: In Section 2.2 we explain what geodesic properties are and their importance in detecting and tracking catheter tips. In Section 2.3 we introduce the proposed *computationally efficient geodesic properties framework* which computes such properties with the same (linear) complexity required to compute the actual geodesic curve. Based on this framework, in Section 2.4 we introduce the proposed catheter tracking method. In Section 2.5 we present the method's exhaustive evaluation. In Section 2.6 we discuss the results, possible limitations, and compare it with the state-of-the-art.

2.2 Geodesic Properties

In the image domain, a geodesic curve defines the shortest path between points in the presence of a metric F. The metric defines the cost along a path in the domain. More formally, the geodesic path $P(s)$ between point a and point b on the image plane, with parametrization s satisfying $\frac{\partial P(s)}{\partial s} = 1$, is the planar curve connection a and b that

minimizes

$$\min_{P} \int_a^b F(P(s))ds. \tag{2.1}$$

Intrinsic properties of geodesics are often useful beyond the path itself. The most trivial property is the total travel time/cost, which is the value the geodesic path minimizes. For a (global) geodesic $P(s)$, the travel time between a and b can be expressed as

$$T_P(a,b) := \int_a^b F(P(s))ds. \tag{2.2}$$

Other interesting properties are the Euclidean path-length and the average travel speed, respectively,

$$L_P(a,b) := \int_a^b P(s)ds, \tag{2.3}$$

$$A_P(a,b) := \int_a^b \frac{F(P(s))}{P(s)}ds. \tag{2.4}$$

Such properties are typically used in ratio weight cycles methods [31, 32], and can be useful for limiting the search domain [33]. As we will later see, they are also important for detecting and tracking the catheter tip. Before that, let us see how we can efficiently compute such properties.

2.3 Efficient Geodesic Properties Computation

The proposed algorithm for the efficient calculation of geodesic properties computes such properties for all geodesics starting from given/selected points to all other grid (image) points. It is based on the upwind fast marching method, and therefore it only supports properties that can be calculated accumulatively. The advantage of the framework is evident when massive calculations of properties are needed for geodesics reaching numerous grid points in the search domain. Otherwise, the calculation of the properties for a single geodesic may not require this framework.

2.3.1 Fast marching algorithm and efficient implementations

The proposed computational framework is based on the fast marching algorithm [34, 35]. This algorithm solves the Eikonal equation $\|\nabla T(x)\| F(x) = 1$, where F is the speed/weight function and T is the arrival time function defined above. The algorithm

is closely related to Dijkstra's shortest path algorithm (it is actually a numerically consistent version of it), and computes the grid solution, starting with the source grid points where $T(s) = 0$, and traversing the domain increasing values of T. The "advance" step of the algorithm uses the unassigned grid points surrounding those for which T has already been computed, and analytically assigns a new T value to them. Figure 2.3 illustrates how the linear wavefront is estimated at this step.

The run-time complexity of the original fast marching algorithm on a grid is $O(N \log N)$, where N is the number of grid points. Some of the more recent efficient implementations, e.g., [36, 37, 38], improved the run-time complexity. Those techniques are based on using a pre-defined sweep strategy, replacing the use of the heap priority queue, to find the next point to process, and thereby reducing the overall complexity to $O(N)$ (though often with a large constant, since all grid points are processed). In prior work, [39], we proposed a new implementation of the fast marching method which also reduces the computational cost to linear complexity $O(N)$ (with the same low constant as the original implementation), using bucket sort style priority queues. Even though the new queue may introduce additional (quantization) errors, the overall error is still within the same magnitude of the original implementation when F is bounded.

2.3.2 Geodesic properties computations via the extended fast marching algorithm

The proposed framework extends the fast marching algorithm to calculate the desired geodesic properties. While we could have also used a fast sweeping approach, the modified fast marching method described in [39] is also $O(N)$, with a lower constant (as detailed above); and can also enjoy early termination. Only the above mentioned "advance" step of the original algorithm is modified to perform the needed calculations. The arrival time is initially calculated and then used to evaluate the properties.

Figure 2.3 illustrates the advance of the linear wavefront estimation, where the arrival times T_1 and T_2 corresponding to grid points P_1 and P_2 were already assigned, and the arrival time T_3 corresponding to grid point P_3 is the one being calculated. Once the arrival time T_3 is known, the wavefront arrival *direction* can be found and used to calculate the geodesic properties. The framework is best explained using the length property (Equation (2.3)), which is the Euclidean length of the geodesic reaching P_3.

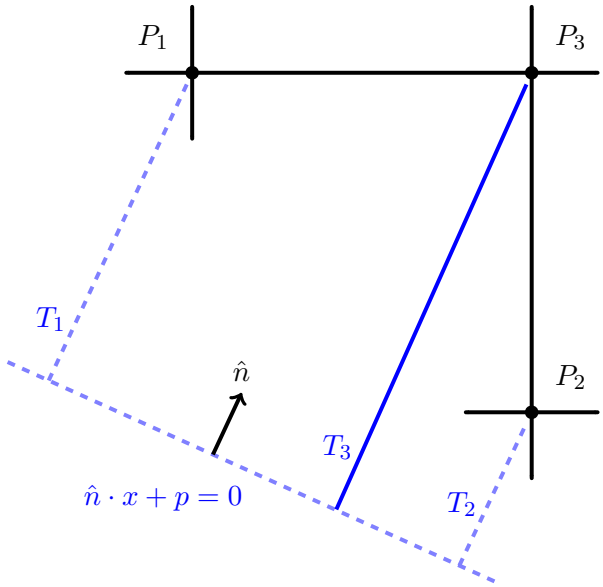


Figure 2.3: The linear approximation of the wave front motion during the "advance" step of the fast marching method. Given two grid points P_1 and P_2 with already computed front arrival times T_1, T_2, and given $F(P_3)$, the arrival time T_3 at grid point P_3 can be estimated. The assumption is that the wave front is planar $\hat{n} \cdot x + p = 0$, passing P_1, P_2 and reaching P_3. The knowledge of T_1, T_2 and $F(P_3)$ allows to determine the parameters of the plane $\hat{n}$,p and solve for T_3.

The geodesic reaching P_3 is estimated to be a straight line inside the $P_1P_2P_3$ triangle of Figure 2.3. However, the geodesic and wavefront direction prior to this triangle may be different. Figure 2.4 illustrates how the length property is estimated. The length property at L_3 describes the Euclidean length of the geodesic reaching P_3. Since the path inside $P_1P_2P_3$ is straight, the Euclidean length of the segment $P_3'P_3$ is added to the geodesic length L_3' at the path entry point P_3' to $P_1P_2P_3$. P_3 and the segment length $P_3'P_3$ are found using geometric calculations based on the wavefront direction. Let $\hat{n} = (n_r, n_s)$ be the direction of the path in $P_1P_2P_3$, with a 45 degree rotated coordinate system, where component n_r is in the $\overrightarrow{P_1P_2}$ direction and component n_s is perpendicular to n_r. Using the solution of the Eikonal equation $T(P_3)$, we compute $\hat{n}$ as

$$\hat{n} = \frac{1}{\sqrt{2}}(\bigtriangledown T(P_3) \cdot (1,-1), \bigtriangledown T(P_3) \cdot (1,1)). \tag{2.5}$$

So explicitly,

$$\hat{n} = \frac{F(P_3)}{h\sqrt{2}}(T_2 - T_1, 2T_3 - T_1 - T_2), \tag{2.6}$$

h being the grid spacing. The entry point P_3' is obtained by linearly interpolating along the line P_1P_2,

$$P_3' = \frac{1}{2}\left[1 - \frac{n_r}{n_s}\right] P_1 + \frac{1}{2}\left[1 + \frac{n_r}{n_s}\right] P_2. \tag{2.7}$$

The geodesic path outside of $P_1P_2P_3$ could be different and therefore L_3' is estimated based on a linear wavefront of the length property, which may or may not have a different direction than inside $P_1P_2P_3$,

$$L_3' = \frac{1}{2}\left[1 - \frac{n_r}{n_s}\right] L_1 + \frac{1}{2}\left[1 + \frac{n_r}{n_s}\right] L_2. \tag{2.8}$$

Finally, the length property L_3 at grid point P_3 is given by

$$L_3 = L_3' + \|P_3'P_3\|. \tag{2.9}$$

Computation of the Euclidian path above is rather simple and here used for illustration purposes only, yet in a similar fashion, more complex properties can be calculated.

The average travel speed property mentioned in Equation (2.4) could be simply calculated as

$$A(P_3) = \frac{L(P_3)}{T(P_3)}. \tag{2.10}$$

Similarly, other geodesic properties, as the estimated kinetic energy spent (e.g., fuel consumption), assuming a fine grid, can be similarly computed:

$$\begin{aligned} KE(P_3) &= \int_0^{T(P_3)} F^2(s)ds \\ &\approx KE(P_3') + F^2(P_3) \cdot (T_3 - T_3'), \end{aligned} \tag{2.11}$$

where

$$T(P_3') = \frac{1}{2}\left[1 - \frac{n_r}{n_s}\right] T_1 + \frac{1}{2}\left[1 + \frac{n_r}{n_s}\right] T_2. \tag{2.12}$$

Other examples could be the average F, standard deviation of F, and moments along the geodesic path.

Having concluded the description of the efficient computation of geodesics and geodesic properties, we now proceed to use this for the task at hand of catheter tip detection and tracking.

2.4 Catheter Tracing Method

The proposed method for catheter detection uses a setup assumption to constrain the general problem and search space. Modern x-ray systems provide wide enough field of view of the heart, and the catheter sheath[1] is also seen when it is entering the heart from a vessel, see for example ablation catheter in Figure 2.1 and in Figure 2.2. The vessel is anatomically stationary and therefore limits the sheath motion. This information constrains the problem in such a way that the catheter always passes through a certain location in the image space. In the beginning of the intervention, the user provides the approximate location x_1, in the image domain, where the catheter exits the sheath. That location is rather easy to observe since once the catheter is out of the sheath, it periodically moves due to the heart beat. In case of multiple catheters entering from different locations, the user may choose more than one entry location $x_1, ..., x_n$.

[1] A sheath is a rigid plastic wrapper tube that facilitates insertion and proper placement of the actual catheter. It is also used for access and bloodless exchange of catheters and protection of the vessel. The catheters, which are made out of a more flexible material, are inserted into the sheath and threaded to the arteries of the heart.

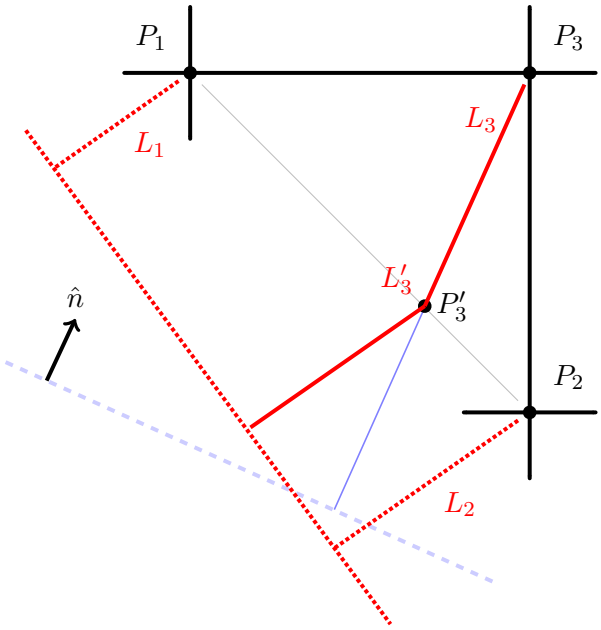


Figure 2.4: Based on the already calculated arrival times T_1, T_2 and T_3, the arrival time wavefront propagation direction $\hat{n}$ and path inside the triangle $P_1P_2P_3$ are known. The path direction when entering $P_1P_2P_3$ could be different than its direction within $P_1P_2P_3$, therefore the length property at entry point P_3' to $P_1P_2P_3$ is calculated based on the linear wavefront corresponding to the length property. Inside $P_1P_2P_3$, the actual Euclidean length is added (segment $P_3'P_3$).

Prior knowledge of the x-ray system setup magnitude and patient position leads to the simple non-limiting assumption that the visible catheter width does not change significantly with depth. This is reasonable as the catheter is bounded by the heart cavaties, which depth is relatively smaller compared to the distance from the x-ray source.[2] The width w_p in pixels can be estimated with acceptable accuracy using the known projection parameters and the estimated distance to the patient's heart,

$$w_p = \frac{SID}{SOD} \cdot \frac{w}{p_s}, \tag{2.13}$$

[2] Magnification range depends on the system parameters and patient size. Using typical numbers, the magnification roughly ranges in the ±10% interval.

where SID is the distance between the x-ray source and the detector, p_s is the image pixel spacing, w is the actual catheter width (in millimeters), and SOD is the distance between the system iso center (center point of C-Arm rotation) and the x-ray source, and is provided by the x-ray system. This is an estimate for the distance between the heart and the x-ray source since it is common practice to place the organ of interest at the iso center in order to make it visible in the center of the fluoroscopic images for any direction.

In addition to the streaming fluoroscopic images I, the proposed method then takes as input the visual catheter width in pixels w_p, the catheter entry locations $x_1, ..., x_n$, and the number of catheters in the scene, N_c. Both $x_1, ..., x_n$ and N_c are manually set by the user, and, ideally, need not be modified during the intervention. This is a likely scenario in EP interventions since the patient, if at all conscious, is discouraged from moving and N_c does not change often during the intervention.

The proposed computational method can be separated into three steps:

- *Preprocessing* - Image scaling and generation of a catheter cost function.
- *Geodesics computation* - Using the geodesic properties framework to generate geodesics and their properties starting from the $x_1, ..., x_n$ locations.
- *Tip classification* - Identifying catheter tips by using a cascade of classifiers.

We will detail each of these steps next.

2.4.1 Preprocessing

The goal of the preprocessing step is to convert the input intensity image I_t at time t to a cost function θ_t in the same domain as the image. The values of $\theta_t(x)$ represent the likelihood (or inverse cost) of catheter presence at a position x in the image at time t. The process initially scales I_t to obtain I_t^* (see Figure 2.5). The resulting scaled image is used to build a background image B_{t+1} and to compute the cost θ_t used in the geodesic computation method (as part of F, see Equation (2.1)).

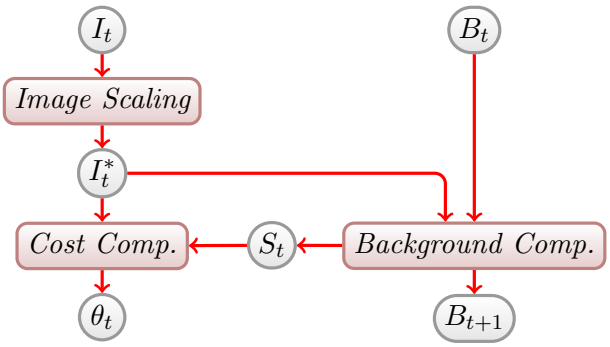


Figure 2.5: Steps in the preprocessing component of the method. At time t, I_t is the input fluoroscopic image and B_t is the background image computed so far. I_t^* is the scaled fluoroscopic image, S_t the subtracted image, θ_t the output catheter cost function, and B_{t+1} the updated background image to be used at time $t+1$.

Image scaling

Fluoroscopic images are often digitized with a surrounding black border. The border size may change on different system modes and also when the physician adjust the collimator, which reduces the field of view, in order to lower the amount of radiation. To avoid edge artifacts, the black border around the image is detected and not taken into account in the following steps. Next, using the computed projected catheter width w_p, the image is downscaled such that a catheter is four pixels wide ($w_p^* = 4$). This reduces computational cost and avoids the need to handle the scale in each step of the method, while at the same time ensuring the catheters connectivity is not lost. The image resulting from this step is denoted as I_t^*. We experimentally selected $w_p^* = 4$ as the optimal value based on a training set (lower values created discontinuities in the cost function, leading to misses of the tip, while higher ones did not improve the overall performance). We kept this value stable all for our experiments.

Background computation

The *background* in this work is defined as the image that encompass everything but the catheters. Catheters appear as dark(er) curves on the image in comparison to the

background that contains tissue and bones, which are less x-ray opaque and therefore usually brighter, see an example in Figure 2.6a.

In EP interventions, the catheters inside the heart are in constant motion due to the cardiac activity. Since the moving tissue is barely visible, the background is assumed to be static and the missing gaps in the background image B_{t+1} are filled by combining fluoroscopic images overtime, $I_0, ..., I_t$. This is done based on an exponentially decaying process. A time decay constant r determines how the insensitivity of the background image B_{t+1} is to change due to each new frame I_t. We define η as

$$\eta = 0.5^{(\tau_{1/2} f)^{-1}}, \tag{2.14}$$

where f is the frame arrival rate in Hertz, typically 2Hz to 15Hz in EP interventions, and $\tau_{1/2}$ is the decay half-life in seconds. Based on the assumption that the human heart rate in rest is between 60 and 80 beats per minute, which corresponds to a cycle duration of 0.75 to 1 seconds, we choose $\tau_{1/2} = 2sec$. Therefore, changes in the fluoroscopic images that don't stay for at least two heart cycles have little effect on the background image.

To compute B_{t+1}, we first compute an intermediate "subtracted" image S_t, Figure 2.6c,

$$S_t = I_t^* - B_t. \tag{2.15}$$

Then, B_{t+1} is computed,

$$B_{t+1} = B_t + S_t \begin{cases} 1 - \eta^2 & \textit{if } S_t(x) > 0, \\ 1 - \eta & \textit{otherwise.} \end{cases} \tag{2.16}$$

Thus, the background image has double the half-life in response to negative changes. This is done to further reduce the effects of moving catheters which have lower intensity values than B_t. We experimented with the frame rate effect on the background image via simulations (sub-sampling the data in the temporal domain). The background image is in general not sensitive to the frame rate, as its decay rate is based on the actual frame rate. An exception happens if frame rates are very slow, close to the ECG signal rate. In such a case, catheters may appear almost static on the image, causing the background image to start containing traces of the catheter in it. This of course hurts the proposed method. Fortunately, the great majority of clinics we observed choose not to use such low frame rates.

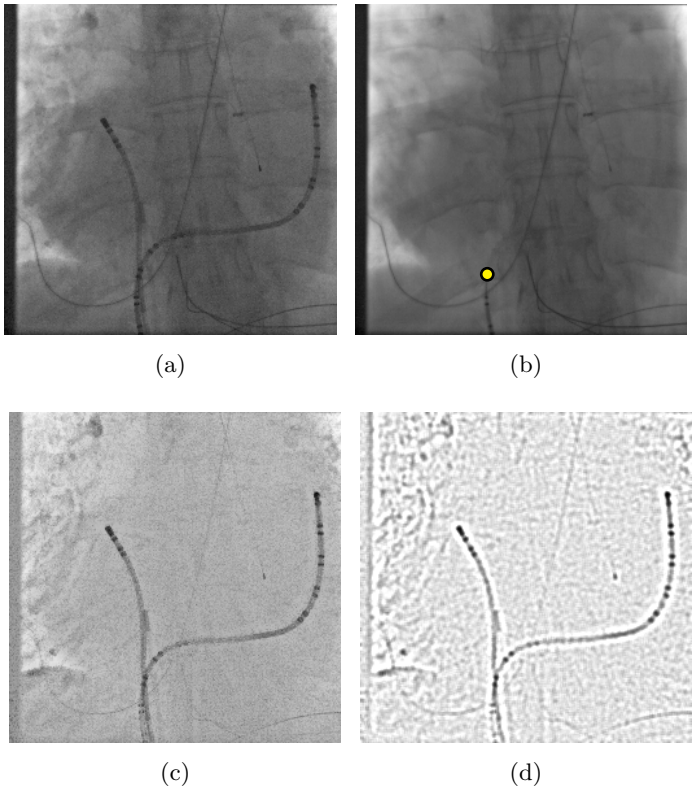


Figure 2.6: Preprocessing result on a sample image. (a) shows the scaled fluoroscopic image I_t^*. (b) shows the computed background image B_t, where $t \gg 0$. Note that the motion of the catheters at the lower part of I_t^* is constrained by the sheath and therefore ended up visible in the background image. This phenomenon may even aid in the selection of the entry points $x_1, \ldots, x_n$. The overlaid circle represents a selected starting point x_1 just where the catheter motion is significant enough such that they are not visible. Other medical instruments outside the heart that lack sufficient motion may also end up visible in the background image, e.g., the ECG wires outside of the patient. (c) shows the subtracted image S_t according to Equation (2.15). (d) shows the resulting cost θ_t.

The background computation process is initialized by averaging the initial fluoroscopic images of the case under study. Since catheter tip detection is not relevant during the rather long time it takes to setup a case, background initialization has fairly low significance. By the time the detection method is used, it can be reasonably assumed there is a stable background image.[3] In an online system, it would make sense to verify background stability, and re-initialize if needed, a process which we did not include as part of this current work.

Cost function computation

The goal of this step is to compute a cost function θ_t, in the same domain as I_t^*, representing the pixel-dependent likelihood (cost) of a catheter presence on I_t^*. A bandpass filter is used to filter out both low frequencies and very high frequencies, which are not characteristic of projected catheters in I_t^*. This filter is built by subtracting two Gaussian filters, one with a variance smaller than the catheter width and the other one with a variance larger than the catheter width. The filter can be implemented in the spatial domain or Fourier domain. Our experiments showed that the proposed method works well even with a rough Gaussian variance estimation. Therefore, a bandpass filter is implemented by subtracting the mean of two square windows, one of size three pixels and one of size nine pixels (recall the catheter is about 4 pixels wide after image scaling). When a stable background image B_t is available, which is expected in EP cases, the band pass filter is applied to S_t (Equation (2.15)) rather than to I_t^*. Finally, the result is normalized to be between zero and one. A value of zero indicates the most likely presence of a catheter, and a value of one represents the opposite. Figure 2.6d shows an example result of the cost function.

While other efficent filters could be used, we found that at best, they lead to a slight improvement of the system as a whole. Directional filters were found to be problematic for example at the intersection of multiple catheters. As described above, the filter we eventually used, which is a type of blob-detector, is both fast and has the desired side effect of enhancing high cost catheters, keeping the propagating front from making false short-cuts. While there may be a better choice of filter, we intenionally picked a

[3] Stability of the background image can easily be measured and verified, e.g., $\sum_x |B_{t+1} - B_t| < threshold$.

simple one to demonstrate that the proposed system does not depend on a sophisticated filtering design.

2.4.2 Geodesics computation

The preprocessing steps established the cost function θ_t that is used for geodesic measurements. Catheters are long and thin and therefore θ_t should have a low intensity valley along the catheters. These intensity valleys have many branches and depth changes, and a method that uses only local information is not sufficiently robust. Our proposed approach traces those valleys, using the framework presented in Section 2.3 to generate geodesics $g_1, \ldots, g_\ell$ that start from $x_1, \ldots, x_n$, thereby integrating the local information. The cost function θ_t is used as the metric F (Equation (2.1)). In order to reduce the sensitivity to the user provided starting locations $x_1, \ldots, x_n$, improved starting locations $x'_1, \ldots, x'_n$ around $x_1, \ldots, x_n$ were used. For each x_i, the location x'_i is determined as the minimal value of $\theta_t(x'_i)$ in a square window of $2w_p$ pixels size centered on x_i. The framework propagates from all $x'_1, \ldots, x'_n$ simultaneously. Figure 2.7b shows an example of the T function (time of arrival or geodesic distance).

For each of the m catheter tips $y_1, ..., y_m$ on I_t^*, there is a geodesic (with the proper metric) reaching it from one of the $x_1, \ldots, x_n$. A geodesic, being the fastest path between two points, represents the most likely path between a tip and one of such starting locations. Nevertheless, it cannot be assumed that those geodesics accurately follow the catheter centerlines (even though it is often the case). To reduce the chances geodesics find shortcuts, we limit the search domain, similarly to [33]. The authors proposed a rough Euclidean length estimation of the geodesic, during the propagation, to determine whether to continue the propagation of the geodesic or not. If the Euclidean length of a geodesic reaching a frontier grid point is much shorter than the lengths of the geodesics reaching the other frontier grid points, then no further propagation is performed on that grid point. In addition to improving the computational efficiency, this forces the geodesics to stay close to the catheter. The downside of this method is that only one path is assumed to be correct. For example, if the front is following two catheters, just because the front is following one catheter faster, the second catheter might get left-out of the search domain. To handle such situations, the relative front speed constraint is replaced with a *recent average* speed constraint, which is a local

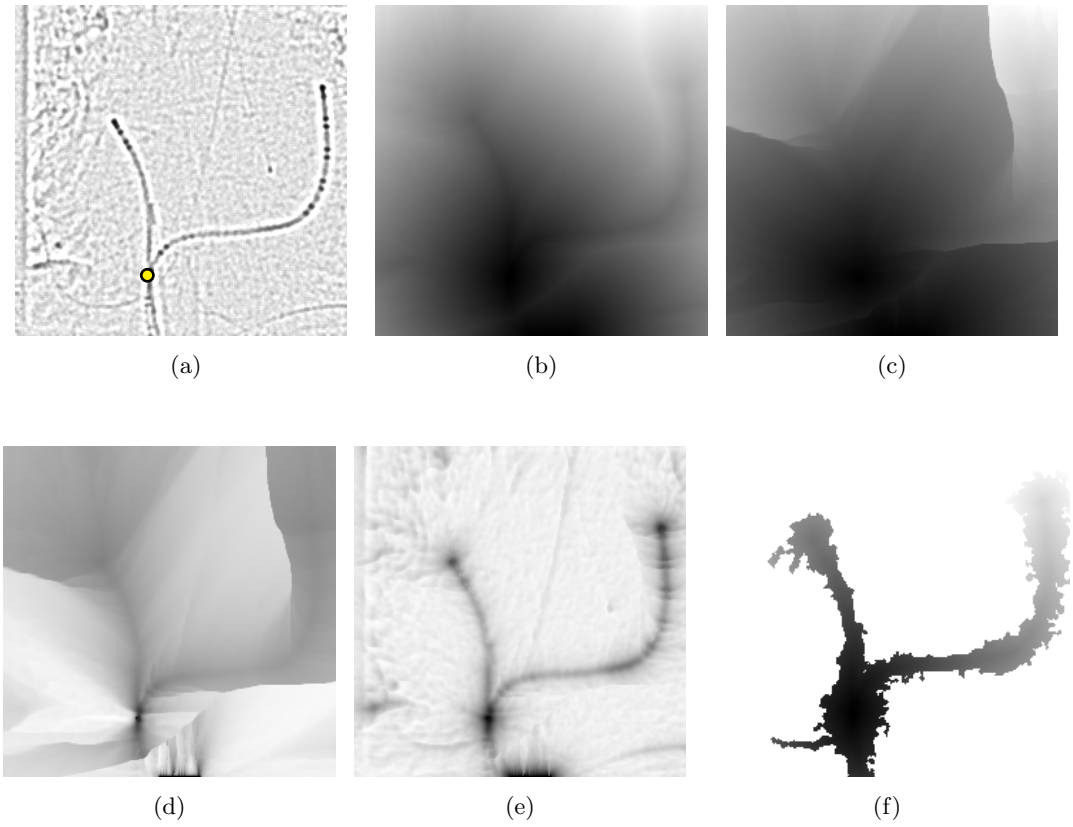


Figure 2.7: Geodesics and their properties computation. (a) shows the preprocessing result of the cost function θ_t. The overlaid circle represents the starting point x_1 as determined in Figure 2.6b. (b) shows the outcome arrival time function T (Equation (2.2)). (c) shows the geodesic Euclidean length property L (Equation (2.9)). (d) shows the average travel speed property A (Equation (2.4)). (e) shows the decaying average speed property D (Equation (2.19)). For demonstration purposes only, the search domain was not limited based on the decaying average for (b)-(e). (f) shows the outcome arrival time function (Equation (2.2)), here the search domain is limited based on the decaying average function.

constraint and is not relative to the rest of the front. Consequently, if a geodesic average travel speed for the past several iterations is too slow, that geodesic will not be propagated any further. It is important to check the speed of several recent advances since just looking at a single advance would be too vulnerable to noise. An ideal approach would involve iterations to average the advance speed of the j last advances for each geodesic, and then check if the average is not above some threshold. However, such a dynamic programming implementation would be costly in terms of run-time. By taking advantage of the *Geodesic Properties Computation Framework* described in Section 2.3, the recent average speed can be efficiently estimated. A *decaying average speed (weighted by the Euclidean distance)* property D is then defined, which is similar to the decaying process used for the background image in Section 2.4.1. Unlike in Equation (2.14), the decay factor here is a variable ζ that depends on the Euclidean distance of the geodesic length increment $\|P_3'P_3\|$,

$$\zeta = 0.5^{-\|P_3'P_3\|\lambda_{1/2}^{-1}}, \tag{2.17}$$

where $\lambda_{1/2}$ is the Euclidean length half-life constant measured at grid spacing h. The gaps in the cost valley rarely go beyond one catheter width, therefore $\lambda_{1/2}$ was selected as one catheter width w_p^*. Similarly to the length property, Equation (2.8), we have

$$D_3' = \frac{1}{2}\left[1 - \frac{n_r}{n_s}\right] D_1 + \frac{1}{2}\left[1 + \frac{n_r}{n_s}\right] D_2, \tag{2.18}$$

where, unlike in Equation (2.9), a decaying process weighted by Euclidean distance was applied,

$$D_3 = \zeta D_3' + (1 - \zeta)F(P_3). \tag{2.19}$$

An example of the decaying average speed property can be observed in Figure 2.7e. When limiting the search domain by applying the decaying average speed constraint, the arrival time function T stays constrained by the catheter presence in θ_t, as seen in Figure 2.7f.

The *Geodesic Properties Computation Framework* does a single pass on θ_t, in which all geodesic properties are calculated simultaneously. Therefore, the *Euclidean length* and the *average path speed* (or cost) properties, as defined in Equation (2.4), are also computed since they are required by the next step of the method.

2.4.3 Tip classification

Computational efficiency is very important to achieve a clinically usable detection method. A computationally inexpensive and efficient classifiers is constructed based on the geodesic properties computed in the previous step and further considering local image characteristics. The approach of cascade classifiers in [40] is followed, where the most efficient classifiers, with low false negative rates, are called before the more costly classifiers that achieve low false positive rates. A training set was used to experimentally determine both the classifiers and the template (see below). This set included short sequences from patients excluded from the test set, from two of the same clinics as the test case (Cleveland and Swedish). About 200 frames were used for this. Changing the order of the classifier 1 through 8 detailed below only affects the processing speed.

Each classifier determines if a candidate at position x in the image plane is a catheter tip. The cascade of classifiers we used is shown in Table 2.1. Each classifier is applied to all the remaining candidates before continuing to the next classifier. The order of the cascade was empirically determined based on the overall computational cost of the cascade. Classifiers nine and ten involve dynamic programming and therefore left to the end.

The first classifier, $A(.)$ *at* x *is a local minima*, was found to be extremely efficient. It works well because the catheter tip, which is a dark region on the x-ray image, gets high likelihood or low value/cost of θ_t, while the background area beyond the tip contains higher values of $\theta_t(x)$. Therefore, the average cost of the geodesic reaching the tip is expected to be the lowest compared to other surrounding geodesics.

The second classifier, $A(x) < mean(\theta_t) - STD(\theta_t)$, checks that the average cost along the geodesic is lower than the mean of the cost function over its domain. The third classifier, $L(x) > 4w_p^*$, verifies that the *Euclidean length* property at x is not shorter than four times the catheter width. The catheter sheath is normally visible in the image and therefore some constraints on its length are acceptable. The fourth classifier, $I_t^*(x) > mean(B_t)$, ensures that the fluoroscopic image intensity at the candidate point is darker than the average background intensity. The fifth classifier, $\nabla\theta_t(x) \cdot \nabla L(x) > 1/\sqrt{2}$, keeps only the candidate x with a cost gradient in the same direction than the geodesic at x. The sixth classifier, $|\nabla I_t^*(x)| > STD(I_t^*)$, constraints the cost gradient at x. The seventh classifier checks that x is not on the image edge or on the dark border.

1. $A(.)$ at x is a local minima
2. $A(x) < mean(\theta_t) - STD(\theta_t)$
3. $L(x) > 4w_p^*$
4. $I_t^*(x) > mean(B_t)$
5. $\nabla\theta_t(x) \cdot \nabla L(x) > 1/\sqrt{2}$
6. $|\nabla I_t^*(x)| > STD(I_t^*)$
7. x not on image edge
8. $1/w_p^* \sum_{a=0}^{w_p^*} S(x + \nabla\theta_t(x)a) < STD(I_t^*)$
9. $x \notin$ { geodesics to other remaining candidates }
10. $x \in$ { N_c best template correlation scores}

Table 2.1: Classifier cascade used to classify a grid point x as a catheter tip. Each classifier determines if a candidate at position x in the image plane is a catheter tip and is applied to all remaining candidates before continuing to the next classifier. $A(.)$ is the average cost function as defined in Equation 2.4.

Figure 2.8: Template used in the template matching step. The template is smoothed to reduce local minima during the gradient descent.

The eighth classifier, $1/w_p^* \sum_{a=0}^{w_p^*} S(x + \nabla\theta_t(x)a) < STD(I_t^*)$, test is passed if the value of I_t^* in the direction of $\nabla\theta_t(x)$ similar to the background image.

The final two classifiers achieve low false positive results. The ninth classifier requires x not to be part of a geodesic of any remaining candidate. This is implemented by sampling the geodesic reaching x and removing any other candidates found on the way. This classifier effectively removes candidates on electrodes, which are not tips, along the catheters. The electrodes are the metallic rings placed along the catheter and are visible in the fluoroscopic image as dark dots on the catheter. Locally, electrodes may look like a catheter tip and therefore may get high correlation score in the last template matching classifier.

The last classifier, $x \in \{\ N_c$ *best template correlation scores*$\}$, is a correlation based template matching. All final candidates are compared using a small generic smooth template, Figure 2.8, to correlate with the shape of a catheter tip. The template matching is initialized based on the orientation of the geodesic at its final arrival point, from there a gradient decent method translates and orients the template to the best local matching position. If more than N_c candidates reach this step, only this first best N_c candidates are considered, based on their correlation score. As a final step, the correlation score is thresholded by a conservative threshold which is interactively determined as half of the best correlation score computed so far. The normalized correlation score is sensitive to the image quality and contrast; therefore the best score observed so far allows to generate such a threshold. This is an heuristic designed to detect noisy candidates that managed to pass all classifier. Any remaining candidate is reported as a positive result based on the final template location.

2.5 Evaluation

The method was evaluated on 1107 anonymized fluoroscopic images taken from four EP cases from four different clinics,[4] each with sufficient recorded frames to test our method. Cases involving ablation catheters and CS catheters. Each image size is 1024×1024 with a pixel spacing of 0.184 *mm*. The estimated catheters width w_p was set between 15 to 17 pixels. The maximum number N_c of catheters to be detected was set between 1 and 3 depending on the number of relevant catheters visible.

For each patient, all frames were combined into a single sequence and tested independently apart for the background image and common start locations $x_1, \ldots, x_n$. Start locations were consistently selected based on the background image as shown in Figure 2.6b. A total number of 1977 ablation catheter tips and CS catheter tips in those frames were manually annotated by two experts as ground truth for evaluation. Tips that were not visible to the annotators were not included in the ground truth (e.g., tip out of the image or motion blur). Any tip not detected was counted as a false negative and any wrong detection was counted as a false positive. A correct detection is counted when a detected tip is within $w_p^*/2$ (two pixels on I_t^*) away from the ground truth tip location. This is a reasonable accuracy for real clinical applications as the typical ablation catheter diameter is around 2 mm (7 French).

The method yielded an overall 3.97% of false negatives and a 1.06% of false positive, with associated per patient standard deviations of 1.96% and 0.82%. Per patient maximum false negative was 4.51% and false positive 2.58%. Table 2.2 shows the break down of the basic statistics. Figure 2.9 shows from the evaluation data set. Figure 2.10 shows sample results for an ablation catheter over consecutive frames.

Run-time was evaluated on an Intel(R) Xeon(TM) CPU 2.99GHz, 2.00 GB of RAM system. It achieved a tracking frame rate of 10 Hz (fps), with an average computation time of 98 msec per image. The *Geodesic Computation* measurements (Section 2.4.2) take 38 msec, while the classifiers, including template matching (Section 2.4.3), take 42 msec, and the *Preprocessing* (Section 2.4.1), including overhead, takes about 18 msec. The *Geodesic Computation* measurements and their required *Preprocessing* operations

[4] Cleveland Clinic at Cleveland, Ohio, USA; Swedish Medical Center at Seattle, Washington, USA; Krankenhaus Barmherzige Brüder at Regensburg, Germany; Hospital of University of Pennsylvania, Philadelphia, USA.

	Classifier Applied		
Average per catheter	#1	#1-#9	#1-#10
Detected Tip Locations	11.6	1.74	0.97
False Positives	1070%	77.4%	1.06%
False Negatives	0.86%	3.57%	3.97%
Average per image			
Detected Tip Locations	20.9	3.12	1.74
Totals			
Detected Tip Locations	23,022	3,437	1,919
False Positives	21,058	1,530	21
False Negatives	17	71	79

Table 2.2: Performance evaluation of the proposed method. The first column shows statistics after applying only the first classifier. The second column shows the same data after applying the first nine classifiers. Finally, the last column shows the performance after all ten classifiers are applied. The average number of detected tip locations, false positives, and false negatives per catheter, are reported as first group of rows. Subsequently, the average number of detected tip locations per image is also reported, since there is usually more than one catheter per image. In the last section of the table we report the absolute values for the tested dataset.

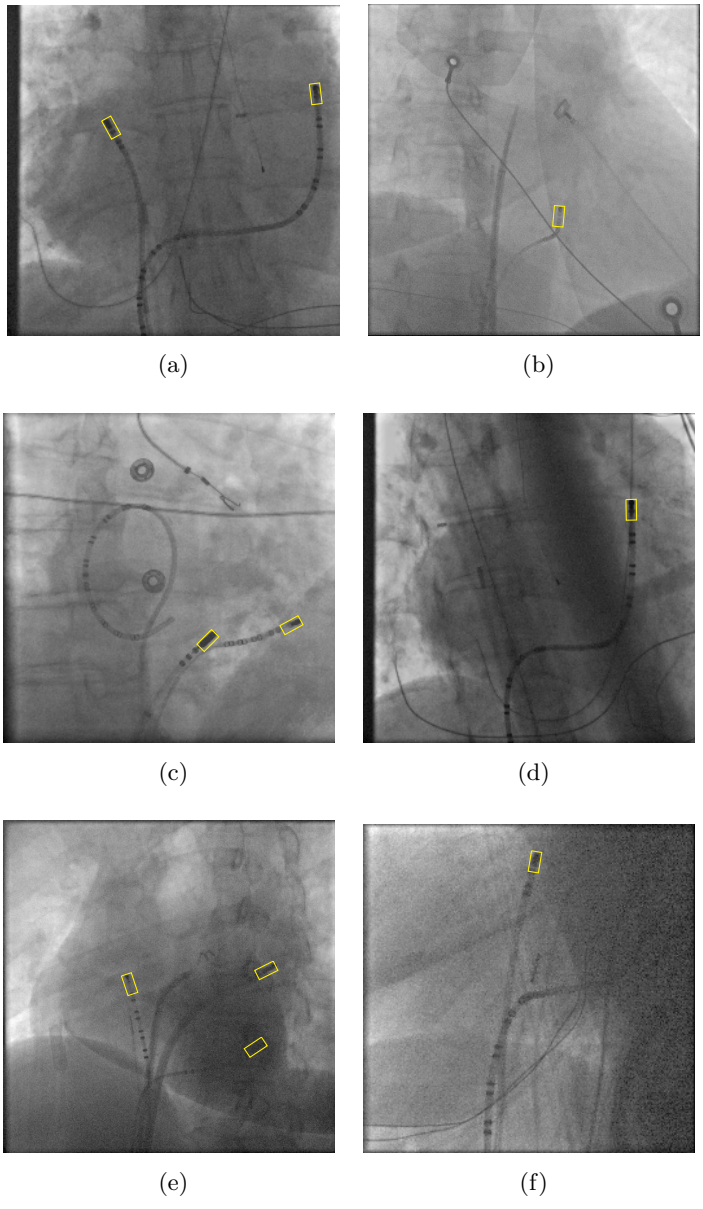
take over half of the computation time, preparing the data for simple classifiers. The method was implemented using C++ and OpenCV library to run on the CPU single thread, yet performance maybe further improved if parallelized or even implemented using the GPU.

We investigated the effectiveness of each of the classifiers. As can be observed from Table 2.2, the initial classifier is very effective and leaves only a handful of candidates out of the search domain, while at the same time keeping a low false negative rate. On the other hand, the template matching keeps the false positives low. The other classifiers are not as effective individually but are relatively cheap to compute and help reduce the number of (more computationally demanding) template matching required operations. The ninth classifier helps in removing electrodes along the catheters that might mislead the template matching, as they locally may resemble a catheter tip.

2.6 Discussion

While the overall performance is very good, the possible reasons for wrong detections were investigated, being the most common reason for a false negative the obstruction of a catheter tip with another catheter, Figure 2.11a. This causes the tip candidate to be dropped by classifier #9. False positives did not have such a common cause of failure, though in many cases the reason is a misidentification by the template matching (classifier #10), Figure 2.11b. None of the false positives occurred on the catheter tips. For evaluation purposes, a generic template appropriate for our experiments was used, while a stronger array of templates may yield an even lower number of false positives and offer the possibility of tracking other kinds of catheters (see additional comments below on the lasso catheter). Therefore, a current limitation of the proposed method is the scenario where a catheter tip is obstructed by another catheter for a long duration.

The state-of-the-art method proposed in [19] uses template matching and Kalman filtering to reduce the search space. They achieved tracking at the rate of 3 Hz for two simultaneous planes which may be comparable to 6 Hz for a single plane. They reported average 2-D accuracy of 3.0 ± 4.2 pixels, and did not define a failure criterion, therefore no false positives rate was reported. According to the authors, the tracking was disturbed by other catheters in 18% of their evaluation sequences. Their method



(a) (b) (c) (d) (e) (f)

Figure 2.9: Sample results from the evaluation data set.

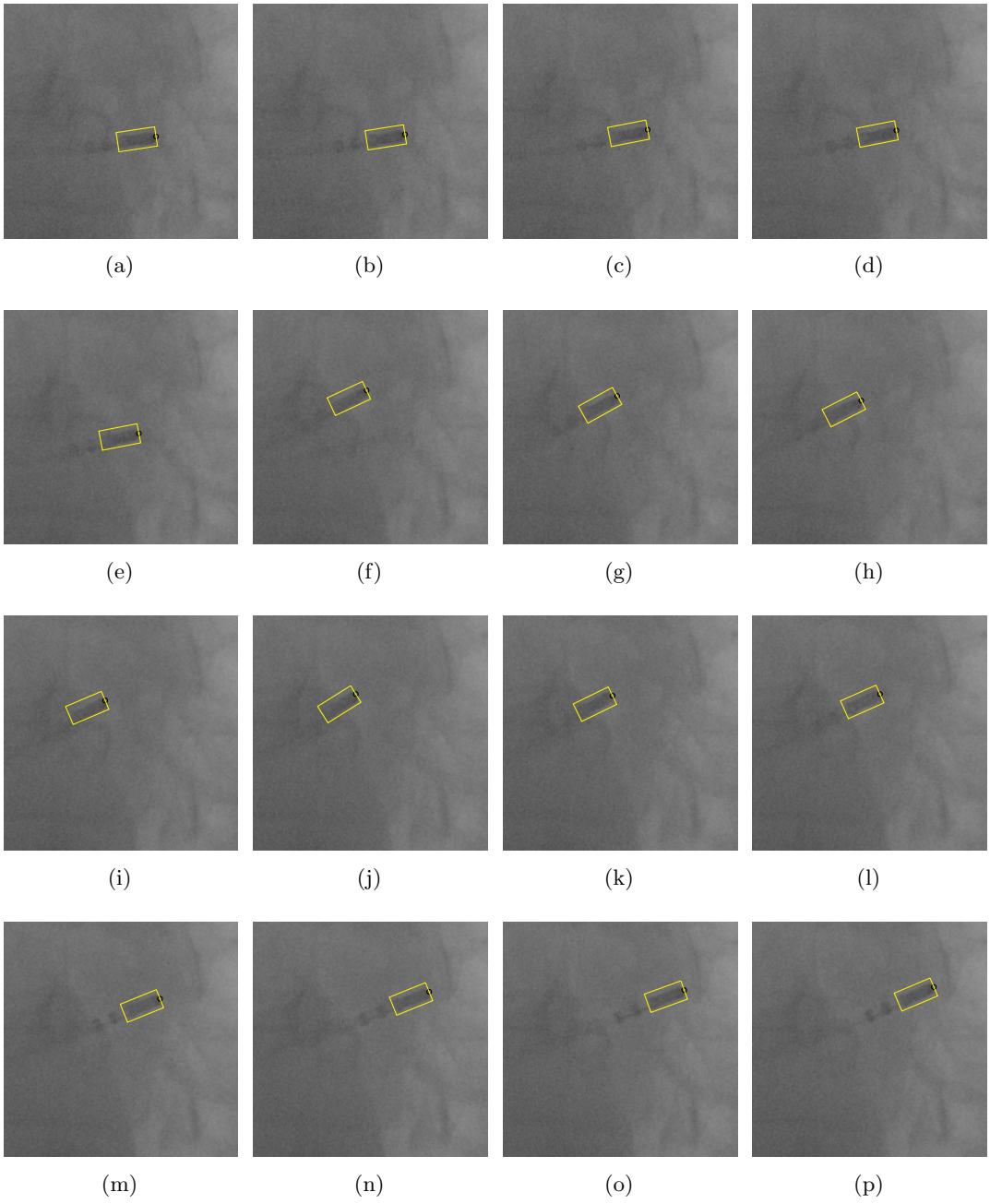
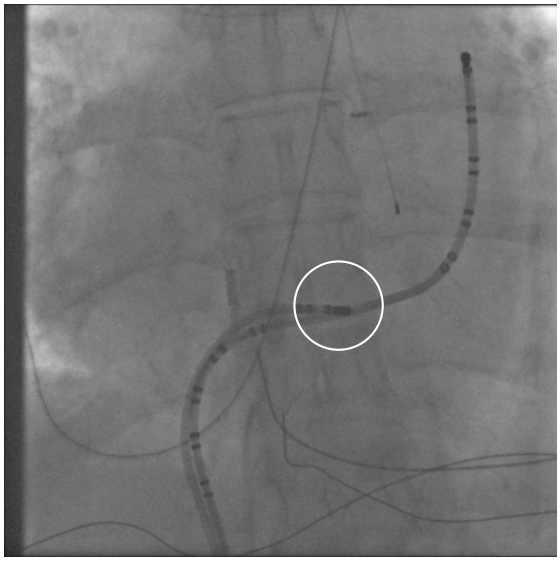
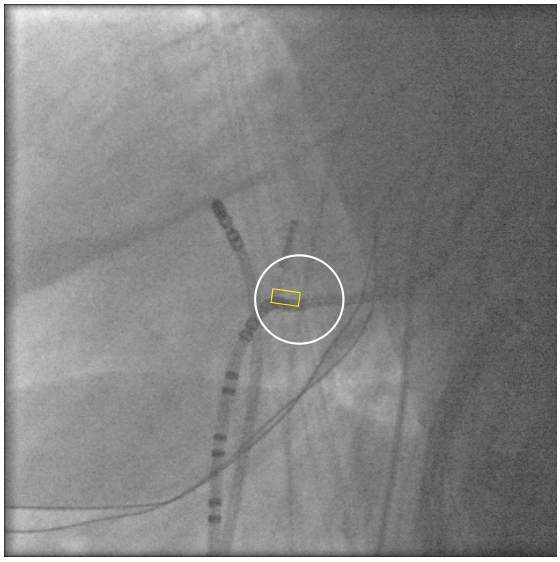


Figure 2.10: Sample results for consecutive frames.



(a)



(b)

Figure 2.11: An example of (a) a false negative due to catheter overlap, indicated by a circle; and (b) a false positive to miss identification, highlighted by a circle.

requires manual initialization and re-initialization each time the tracking is lost and could not recover. In comparison, the proposed method allows for robust multiple catheters tracking at faster rates. Yet, it is important to note that the evaluation data differs for the two methods. The work [19] focused on ablation catheter tracking when it is in proximity to a lasso catheter during PV ablation, which was the probably prime source of the disturbance, while our evaluation set included a lasso for only one of the patients. Since during pulmonary vein ablation it is common practice for a lasso catheter to be present, we further analyzed the effects of the lasso catheter, and found that a thin lasso catheter is not expected to affect the detection, as it is mostly lost following the preprocessing described in Section 2.4.1, while a thicker lasso might increase mostly the false negatives when the tip of the ablation catheter overlaps the lasso catheter or any other catheter. Even when such an occlusion happens, the proposed method may still be successful if the ablation catheter tip is darker than the lasso. If a bi-plane setup is available, it is unlikely that the ablation catheter would be occluded from both planes at the same time, thereby allowing for the proposed method to detect the ablation catheter from at least one of the planes. It could still be beneficial in the future to consider Kalman filtering or some form of time constraint to further improve speed and robustness.

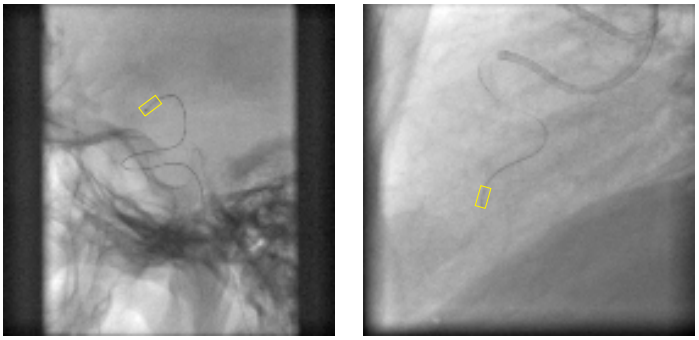
The ideal setup for the proposed method is a bi-plane system that provides two viewing angles and therefore rotation of c-arms does not usually happen during the ablation phase. In case of a mono-plane system, the c-arm is rotated a number of times during the case to help navigate the catheter. In such a scenario, the operator would need to set a new entry point or points after each rotation. If the physician uses standard c-arm angles, an online system can be extended to support two or more states in which both the entry points and background image are automatically remembered and set. The worst case scenario happens if the physician decides not use standard angles. In such a case an operator might need to reset the entry points often, which would require a dedicated medical stuff to operate the system as also required by alternative products. Yet, we believe that this scenario won't happen in practice since it is not uncommon that the operating physician that uses a mapping/navigation technology changes his/her workflow to adjust to the technology requirements. Alternatively, the method could also be improved to automatically detect the entry points, which is the

subject to future research. This may also benefit the case where there is an organ movement beyond respiratory motion and cardiac motion, which does not happen often but when it does may need resetting of the entry points.

The proposed detection method and geodesic properties framework could be extended to other applications. The method showed potential when tested on guidewire detection and neurological catheter detection, see Figure 2.12. However, modifications in the classifiers will be required to achieve satisfactory accuracy in such applications.

The scope of the work does not attempt to explicitly identify the detected catheter. If an application requires this, the result of our method leaves only a few possibilities. An additional component could be added to identify the result using for example an user supplied template.

We now roughly compare x-ray navigation/mapping methods in general with other established products; Biosense Webster CARTO© and St. Jude EnSite NavX™. CARTO© uses a special catheter that senses a magnetic field which is used to determine the location of the catheter. NavX™uses a specialized catheter to detect electric signals from patches placed on the patient body to determine the 3D location in a similar way a GPS works. The main advantages of x-ray based guidance system include the low cost, that no new specialized hardware is required, no special catheter, and that registration to a c-arm CT is trivial. The main disadvantages are the additional radiation (requires fluoroscopy image to navigate), and the requirement of either a bi-plane c-arm system or some constraints on the physician workflow. Advantages of CARTO© and NavX™include radiation free navigation and free integrated electric signal recording and mapping. The main disadvantages are the high cost per use, the requirement of specialized hardware and catheters, and the fact that it has no reference to other modalities, since registration to fluoroscopy or other 3D modalities such as c-arm CT is not trivial.



(a) (b)

Figure 2.12: Selected results when the detection method is tested on other applications. (a) neurological catheter. (b) guide-wire.

Chapter 3

Esophagus Silhouette Extraction and Reconstruction from Fluoroscopic Views for Cardiac Ablation Procedure Guidance

3.1 Introduction

Over the past few years, intraoperative radiofrequency (RF) ablation has become an important therapeutic option for patients suffering from atrial fibrillation (AF). A rare but very serious complication is atrio-esophageal fistula formation, which is associated with a high mortality rate. Reducing the ablation temperature and power settings along the posterior left atrial wall has been proposed in order to reduce the risk of this complication [41]. Alternative approaches include esophageal temperature monitoring and the use of intracardiac echo [42]. However, there is a general discomfort and risk associated with inserting a probe in the esophagus, especially if the patient has to remain awake during the procedure. Identifying the esophagus location, and avoiding lesions on the left atrium by directly overlying the esophagus model in the visualization, has been proposed as well [18]. A particularly cost-efficient approach is to use orally administered barium to visualize the esophagus under fluoroscopy during a procedure,

see Figure 3.1. While this is a simple and inexpensive technique, the main difficulty with this approach is that the paste does not remain in the esophagus during the entire ablation procedure. There is a need therefore to have a graphical representation of the barium location in the esophagus that would be persistent throughout the entire ablation procedure. Unfortunately, due to esophagus motion, a static esophagus representation cannot be fully trusted [43]. An esophagus model can, however, be useful when combined with conscious patient sedation, as it can guide the physician when he or she approaches the sensitive esophagus region. The closer the physician gets to this region, the more important it is to monitor the patient for signs of pain. Monitoring pain during conscious sedation works well as patients are very sensitive to esophageal warming [44]. It has been shown that a patient's pain perception can be used as a form of (simple) temperature monitoring. However, such an approach is not exact, and it carries the risk that the patient is not providing feedback soon enough. Confronted with such situations, physicians often desire the benefits of a visual representation of the esophagus, even if it is not exact. This gives the physician an opportunity of alerting the patient to pay close attention to any pain they may feel during the next steps of the procedure. While the optimal technique for avoiding injury to the esophagus has not yet been determined [8], the graphical representation proposed here provides a possible solution while preserving the benefits of real-time 2D esophagus monitoring, as suggested by [43], and at the same time expanding visualization to 3D using multiple views reconstruction on a geometry-calibrated C-arm system. The 3D surface produces a useful graphical overlay in case the C-arm system changes the viewing angle.

Prior work in this relatively new topic is fairly limited. Studies such as [45] reconstructed the esophagus using an electro-anatomic catheter for tagging the esophageal borders. Such a method is time consuming and complex in comparison to the use of barium, and it may not be suitable for atrial fibrillation procedures. Works such as [46, 47] segmented the esophagus from a CT, yet [48] concluded that organ movement makes pre-acquired CT inadequate for atrial fibrillation treatment use.

In this paper we first propose a method to extract the 2D esophagus silhouette by compounding the barium flow location from a single 2D fluoroscopic sequence.[1]

[1] If the physcian chooses not to move the C-arms during the procedure, then the 2D silhouette already has clinical value, and further steps are optional.

Then we develop a method that takes multiple 2D esophagus silhouettes extracted from fluoroscopic sequences, acquired from different views, and reconstructs a 3D surface which estimates and represents the internal wall of the esophagus during swallowing. This 3D model can be overlaid on the live fluoroscopic images, essentially serving as an augmented-reality substitute for the esophagus once the barium has disappeared. The esophagus can also be visualized as part of a 3D scene comprising CT rendering or other objects such as a model of the left atrium, ablation points, and catheter models.

The proposed method extends any existing work-flow, with the optional requirement of simultaneous x-ray from two views:

1. Administer, by swallowing, the barium into the patient's esophagus.
2. Acquire simultaneous x-ray sequences from two views,[2] capturing the barium flowing down the esophagus.
3. Reconstruct the esophagus model from the image sequences.
4. Visualize the result, either in a 3D setting or overlaid on the fluoroscopy.

We next describe the key algorithmic and clinical components of this proposed work-flow.

3.2 Methods

Our approach involves two major steps. First, the outline of the esophagus is extracted based on a sequence of x-ray images acquired during a barium swallow recorded on a bi-plane C-arm system whose projection geometry has been calibrated (such systems are manufactured by a number of vendors, and high precision calibration is possible [10]). Then the 3D shape of the esophagus is reconstructed based on the known projection geometry of each view.

[2] If simultaneous x-ray is not available, the method can still produce a 2D esophagus representation but not a full 3D surface reconstruction. The 2D representation is valid as long as the C-arm system does not change the viewing angle.

Extraction of 2D esophagus silhouette

When the barium contrast agent has been swallowed by the patient, it flows down the esophagus. A single still fluoroscopic image from each direction does not capture enough of the esophagus, since it represents only a snapshot in time, see Figure 3.1. However, a sequence of fluoroscopic images can be used to combine esophagus position information over time. This finally enables us to extract the esophagus' silhouette.

Although the barium is usually darker than any of the patient's anatomy, some challenges still remain to get an accurate outline of the esophagus due to the presence of other dark regions and objects in the x-ray images, such as catheters. Since the dark regions of the image represent lower pixel values, the following operations are applied. First, a maximum image is computed by evaluating the per-pixel maximum along the time dimension of each fluoroscopic sequence, Figure 3.2(d). Since the moving barium is darker than the surrounding tissue, it is usually not picked up by a temporal per-pixel maximum operation. As a result, the barium is usually not visible in a maximum image. The moving catheters are also darker than tissue, and they also vanish in a temporal per-pixel maxium image due to their motion. For each image (frame) in the sequence we also perform a morphological closing [49], with kernel slightly larger than half the catheter width.[3] This replaces the catheters with a brighter gray scale intensity, Figure 3.2(b). As the swallowed barium is a wide smooth blob of contrast, it keeps its shape. Then we generate the per pixel (time) minimum image over the resulting frames from the closing operation, Figure 3.2(c). Afterwards, we multiply the minimum image by a factor of two before we subtract the maximum image obtained in the previous step, Figure 3.2(e).

The resulting image integrates the whole time lapse and highlights the locations reached by the barium, while other dark background regions and medical instruments are obscured. The closing operation is necessary to remove the moving medical instruments as demonstrated in Figure 3.3. It may occur that two close-by catheters form a dark region which is wider than twice the closing operation kernel size. This would cause some minor remnant in the processed image, not large enough to interfere with the overall method. Such a remnant can be found in the lower part of Figure 3.2(e) towards

[3] Since the catheters move within the left atrium, a small heart chamber, far enough away from the x-ray source, the projected catheter width does not change significantly.

the middle of the image.

The subtracted image is used to extract the esophagus silhouette, Figure 3.4. An histogram is created over the subtracted image intensities. Then, a gray-level threshold is determined based on Ostu's method [50], Figure 3.4(a-c). The method assumes that the image to be thresholded contains two classes of pixels (barium and background), and then calculates the optimum threshold separating these two classes, so that their combined spread (intra-class variance) is minimal. This method works well on the subtracted image due to the sharp change in histogram values. Next, all segmented structures are found and classified as an esophagus according to their shape and size. In nearly all tested cases, we obtained a single segmented structure. When more than one structure was found, as in Figure 3.4(c), we used a simple classification method based on a bounding box. The shape of an esophagus is vertically long and narrow, and therefore the bounding box is constrained to have a height to width ratio of at least 3:1, with a minimum width of $5mm$ (clinically set value).[4] While this classifier, which is based on anatomical knowledge, worked well for our data, a more sophisticated method might be needed if the fluoroscopic acquisition direction is with an oblique craniocaudal (CC) angle, though these C-arm angulations are rarely used in electrophysiological ablation procedures. The resulting segment is the esophagus silhouette, Figure 3.4(c,d). This shape is then fed into the 3D reconstruction algorithm, as explained next.

3D Esophagus reconstruction from multiple 2D esophagus silhouettes

This step reconstructs a 3D surface of the esophagus from multiple 2D esophagus silhouettes. The 2D esophagus silhouettes are extracted from fluoroscopic sequences acquired under at least two directions.

For each 2D esophagus silhouette, the left and right edges are found and smoothed using a standard geodesic active contours approach [51]. As a result, each silhouette produces two 2D curves as shown in Figure 3.4(e). Each pair of silhouettes provides four 2D curves (two per silhouette), which are used to compute the corresponding four

[4] Other structures may be due to motion of organs when breathing. This is why patients are encouraged to hold their breath after the barium swallow during the barium imaging. Ideally, imaging is preformed with the diaphragm, which is the most interfering structure, below the region of interest, i.e. the heart. When standard electrophysiology C-arm viewing angles (RAO, LAO) are used, the esophagus projections do not overlap.

3D curves. This is done based on the known projection geometry between the two fluoroscopic sources and detector. The clinical C-arm view directions are such that epipolar lines are close to horizontal, while the esophagus is a predominantly vertical structure. This enables us to use simple triangulation methods available in the computer vision literature for this computation [52]. Each 2D curve on the left image is sampled and for each such sample the epipolar line is computed on the right image. The intersection points of the epipolar line with the 2D curves on the right are used to triangulate 3D points. Those points are used to interpolate four 3D curves.

The resulting 3D curves outline the visual hull, [53], of the esophagus. Depending on the end-user application, the visual hull may already be a sufficient estimation of the esophagus. The accuracy of the esophagus structure could further be improved when prior information is available. For example, if there are CT images of the patient, then the esophagus could be segmented from this 3D data set and matched to the enclosing visual hull. Without using prior information, an heuristic approach can be used to further improve the estimate of the esophagus structure.

For atrial fibrillation (Afib) radio-frequency catheter ablation procedures, the goal is to define an awareness zone around the esophagus, in order to avoid heat damage and esophageal fistulas. On bi-plane C-arm systems two views are available, usually spaced approximately 90 degrees apart. Therefore, standard stereo vision methods, [54], that assume locality in 3D, are invalid. As a solution to this problem, we applied an algorithm that takes into account that the esophagus is a mostly smooth organ, enclosed by the visual hull and its tangent to it [53, 54]. This proposed procedure estimates the esophagus surface as an Hermite curve inscribing the visual hull, as described next.

Initially, the direction that describes the esophagus length direction (head to feet direction) is found. Similarly to CT, the fluoroscopy system has a standard patient coordinate system. Not surprisingly, its $\hat{z}$ (head-feet) axis is parallel to the esophagus length direction. In case such information is not available, the head-feet direction can be found using principal component analysis [55]. The found direction is sampled such that each sample plane intersects the four curves, resulting in four points on each plane. Center points are determined between each two neighboring points as illustrated in Figure 3.5. The center points are used as control points for an Hermite curve, resulting in 3D curves describing the esophagus internal surface.

Note that for radio-frequency ablation cases in the heart, the angle between x-ray projections is about 90°. In addition, the distance from the esophagus to the x-ray source is large. Therefore, the projection geometry is close to a parallel-beam geometry. This explains why the 2D shape created by a slice of the visual hull resembles a quadrilateral region. Since an ellipse can model an esophagus slice [46], we could in principle estimate the esophagus as the largest ellipse (in area) that can fit the computed visual hull inside this slice. Yet, it is important that the reconstructed shape projects back to the fluoroscopic image as the silhouette that created it. For this to happen, esophagus shape must have tangent points at each vertex of the quadrilateral, but the largest ellipse may not necessarily satisfy this condition. We therefore use a more flexible Hermite curve as an estimate which can always satisfy this condition.

3.3 Results and validation

The proposed method was tested on six pairs of fluoroscopic sequences (from six different subjects). The data was acquired over several atrial fibrillation cases at the Swedish Medical Center, Seattle, WA, USA, using an AXIOM Artis dBC biplane C-arm system (Siemens AG, Healthcare Sector, Forchheim, Germany).

For two of the patients, a pre-acquired CT was also available. They were used as a reference for a first qualitative assessment to find out if our method produces reasonable results for the application at hand. To this end, the reconstructed esophagus was manually registered using a rigid transformation to the CT images based on the centerlines of the reconstructed model and its counterpart in the CT data set. The esophagus as seen in the MPR slices of the CT data set was then visually compared to the 3D model, see Figure 3.7. As expected, the esophagus in the CT is slightly thinner, since there was no barium swallow during the CT acquisition. For an application targeted at providing a safety region around the esophagus for protection, however, it should be acceptable or even desirable if the esophagus is reconstructed in a more active state. In these data sets, the reconstructed esophagus matched surprisingly well to the pre-acquired CT data sets. Note, however, that the esophagus may shift by more than 15 mm over the course of an Afib treatment, mostly due to organ movements [48]. This suggests that the safety region around the centerline of the initially reconstructed

esophagus should have a radius of at least this size, preferably more.

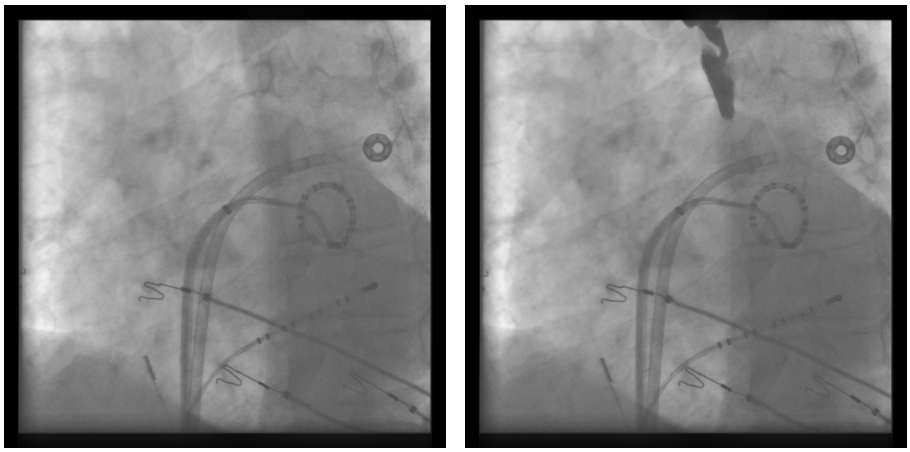
A more quantitative analysis was performed as well. During this experiment, the barium region was segmented using the proposed method. For comparison, an expert annotated the esophagus silhouette on the fluoroscopic images as well. We found that the automatic segmentation adequately defined the barium region in all six cases. This can be appreciated when the edges obtained by automatically segmenting the barium regions are overlaid as shown in Figures 3.6(a-b). Per pixel comparison of the proposed automated segmentation with expert annotation (averaged by patient) resulted in 2.9% false positives with 0.8% standard deviation, and 0.3% false negatives with 0.09% standard deviation. Measuring the segmentation edge distance among pixels in the automated segmentation edge to the expert annotation edge (nearest Euclidean distance) yields per-patient average over-segmentation error of 0.48 mm (about 2.6 pixels), and standard deviation of 0.16 mm. The under-segmentation yielded 0.20 mm (about 1.1 pixels) average error, and standard deviation of 0.16 mm. The automatically segmented edge consisted of 66% over-segmentation (recall that having a conservative protective barrier is often desirable), 12% under-segmentation, and 22% was found pixel-accurate. Overall the obtained per patient edge distance error is 0.36 mm (about 2 pixels) with standard deviation of 0.13 mm.

Due to the cone-beam magnification factor of 1.5, these 2D values correspond to an average 3D error of about 0.24 mm and standard deviation of 0.09 mm. A systematic triangulation error due to imperfect system calibration adds about 1 mm of uncertainty [10] (the same calibration has been used in this work). Assuming the angle between views is 90 degrees, which is common for cardiac ablation cases, we then expect an average 3D error of $1 + 0.24\sqrt{2} \approx 1.3$ mm and standard deviation of $0.08\sqrt{2} \approx 0.12$ mm.

3.4 Discussion

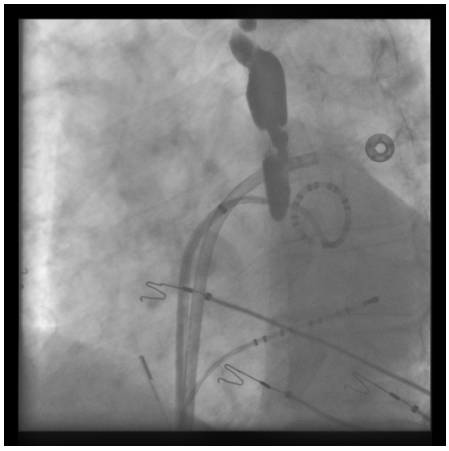
This work proposes to reduce the risk of atrio-esophageal fistula formation by superimposing a graphical esophagus representation onto live fluoroscopy, Figure 3.8. There are potential limitations to the reconstruction concept introduced here that will be addressed in future developments. For example, the wall thickness of the esophagus

cannot be accurately captured when relying on contrast inside the esophagus. This issue is probably of lesser importance in the esophagus avoidance application, since the esophagus of a patient who swallows barium widens while doing so. In addition, fluoroscopic sequences are acquired over a breath-hold, As a result, the esophagus may look different when the patient breathes freely. Safety margins may be added to the esophagus contours to account for this [56]. Since the barium swallow is imaged over multiple heart cycles, we believe that our esophagus model may already capture deformations due to heart beat. Nevertheless, we advise to increase our safety margins even further to also account for heart motion we may have missed. Lastly, the esophagus is a mobile structure, and a 3D model generated at the beginning of the ablation procedure may not accurately reflect the position and width of the esophagus at the time of ablation.

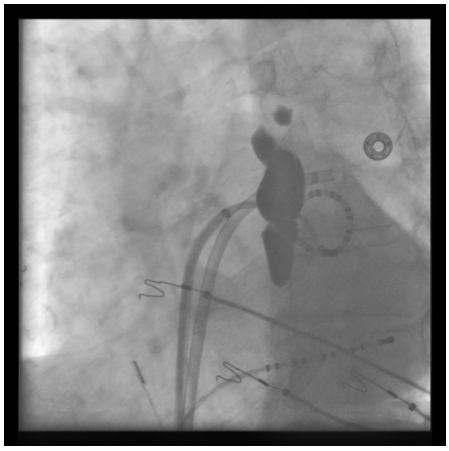


(a) T=0

(b) T=.5 sec



(c) T=1 sec



(d) T=2 sec

Figure 3.1: Selected images from a fluoroscopic sequence (at 15 frames per second) illustrating a barium swallow. The first image on the left shows the scene prior to the barium reaching the field of view. Subsequent images display the barium entering the field of view and flowing down the esophagus. The x-ray images have been acquired on an AXIOM Artis dBC biplane system (Siemens AG, Healthcare Sector, Forchheim, Germany).

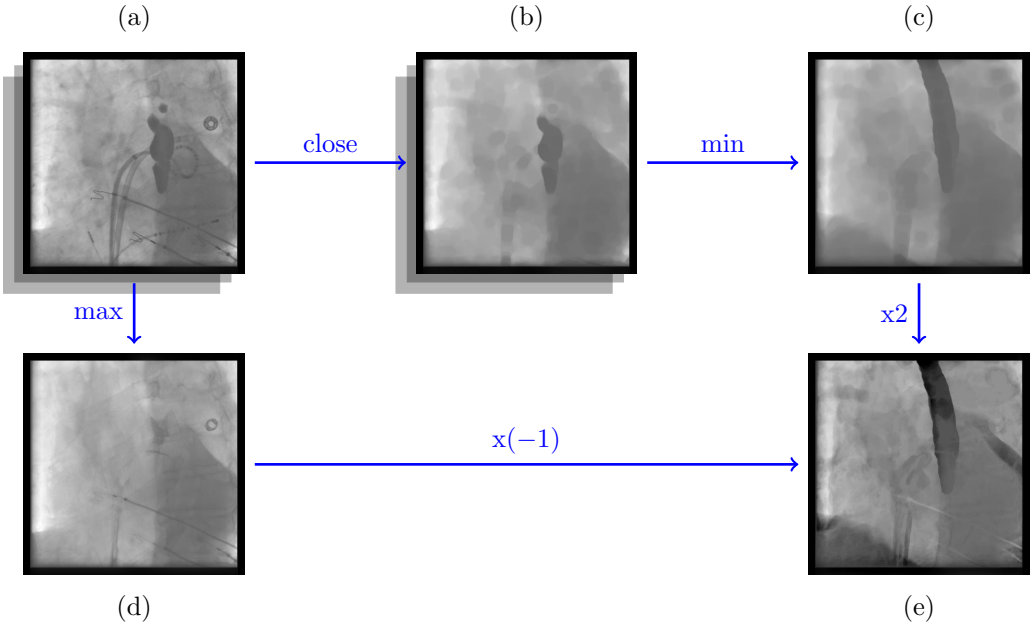
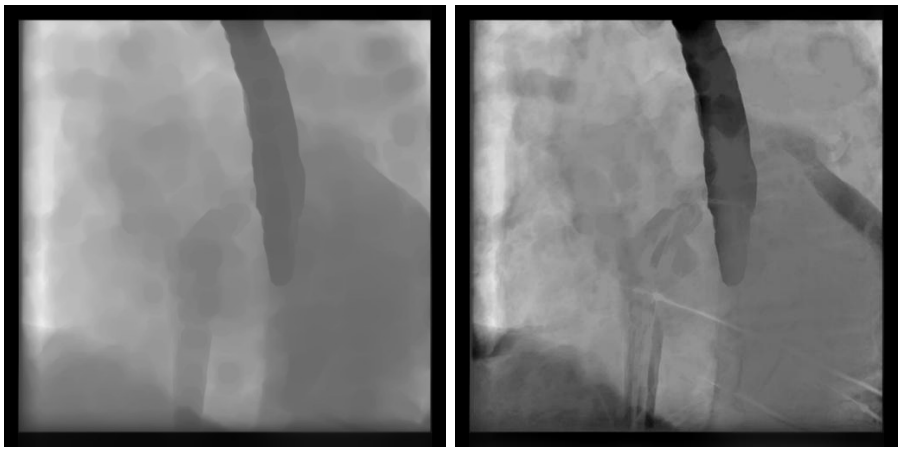
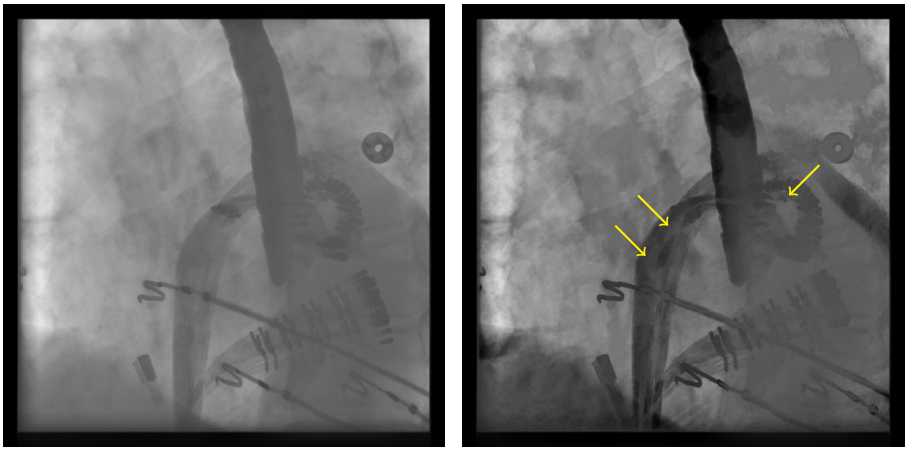


Figure 3.2: This diagram explains how the barium location is collected and enhanced over the fluoroscopic sequence. (a) Sample input frame, used for demonstration, out of an image series, (b) the result of the per-frame closing operation, (c) the per-pixel minimum image over the full sequence, (d) the associated per-pixel maximum image, (e) result obtained after multiplying the minimum image (c) with a factor of two before subtracting the maximum image (d), highlighting the barium location for all the images in the sequence. The time series of X-ray images taken during the barium swallow is used to detect the entire esophagus coated by this contrast medium.



(a)



(b)

Figure 3.3: Illustration of the relevance of the morphological closing operation on the merge step of the proposed method. Using the same sample sequence as in Figure 3.2. (a) The minimum image (left) and the resulting difference image (right) as proposed by the method. (b) Corresponding result if the morphological closing operation is not applied. Note the artifacts caused by the catheters as indicated by the arrows. This example illustrates how the closing operation successfully removes the catheters while preserving the visible barium region.

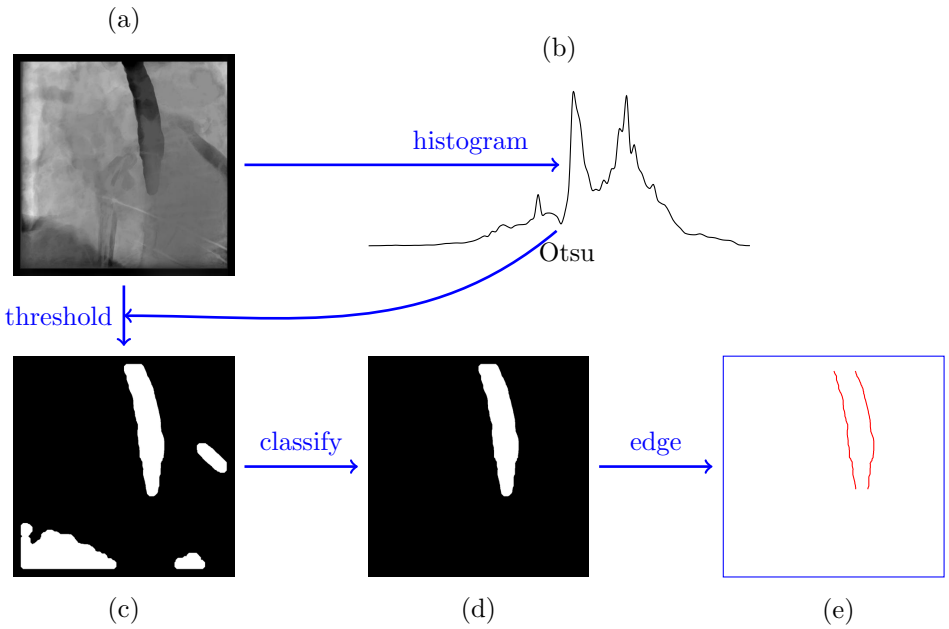


Figure 3.4: Silhouette extraction. (a) Resulting subtracted image from the previous *merge* step; (b) an histogram is generated (ignoring the dark frame of zero pixels around the x-ray image border), and a threshold is determined based on Otsu's method; (c) the thresholded subtracted image; (d) the segment in the thresholded image that is classified as an esophagus shape is selected. For the 3D reconstruction step, the left and right edges of the silhouette are extracted and smoothed using a standard geodesic active contours approach.

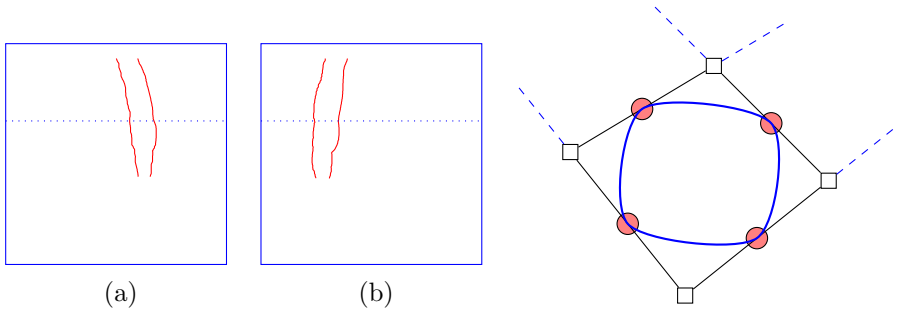


Figure 3.5: 3D Esophagus reconstruction from multiple 2D esophagus silhouettes. (a), (b) Left and right silhouette edges seen from two different C-arm viewing directions. The dotted lines on the edge show a slice selected for reconstruction in this example (the clinical C-arm view directions are such that epipolar lines are close to horizontal). (c) The actual slice comprising a quadrilateral. The quadrilateral represents the intersection of the reconstructed points by triangulation. The red circles are the mid-points between each neighboring reconstructed point. The four mid-points are connected with an Hermite curve, estimating the esophagus wall within that slice.

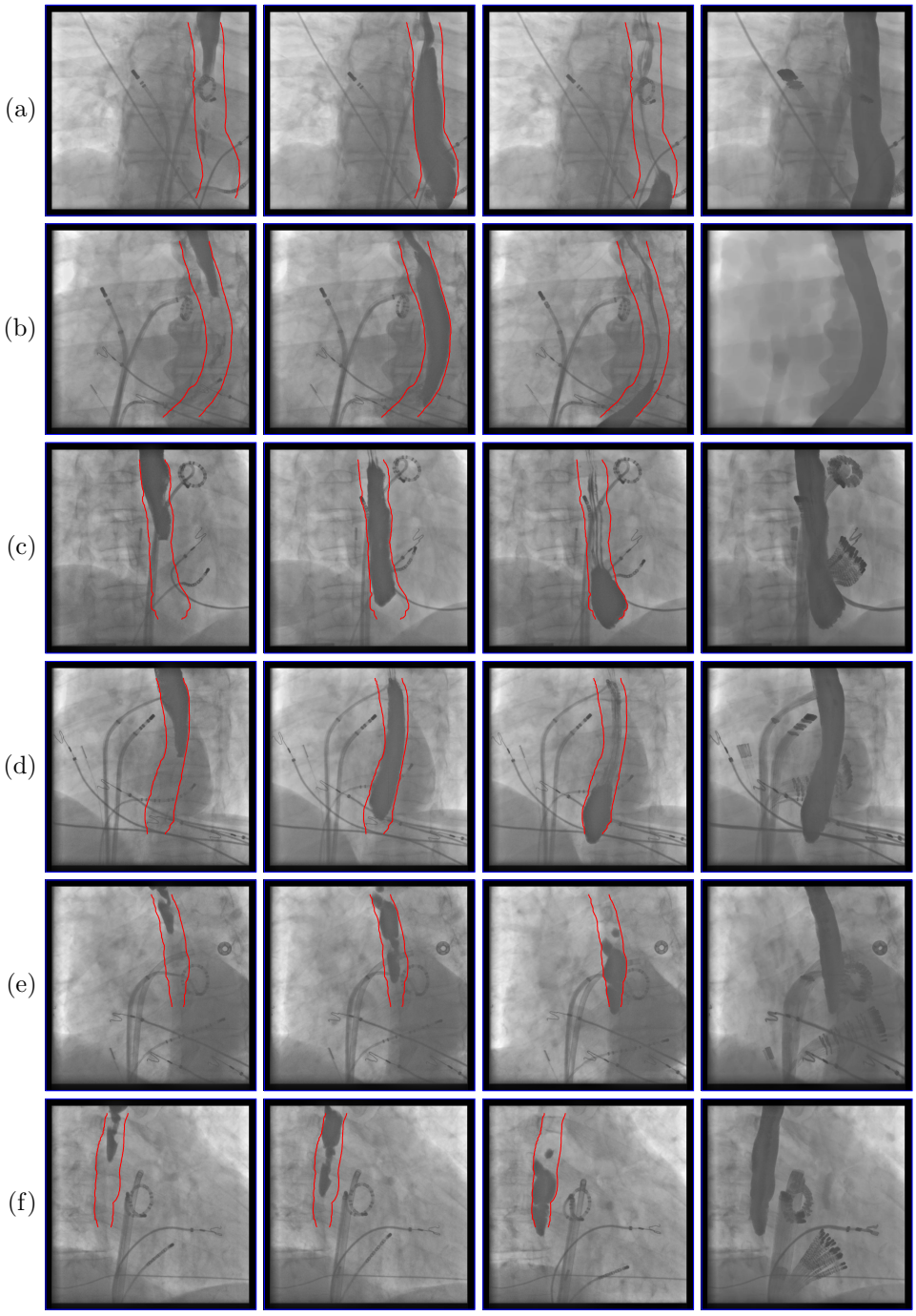
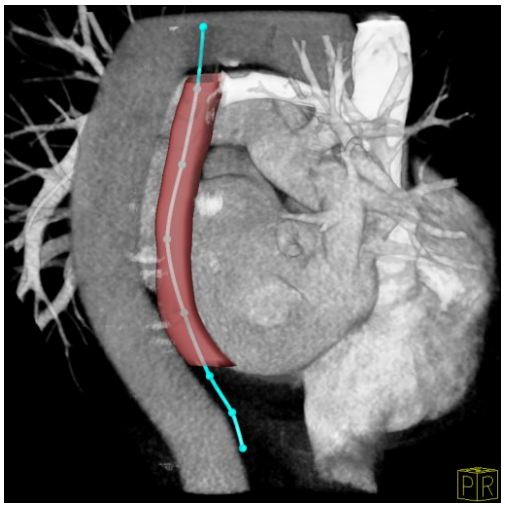
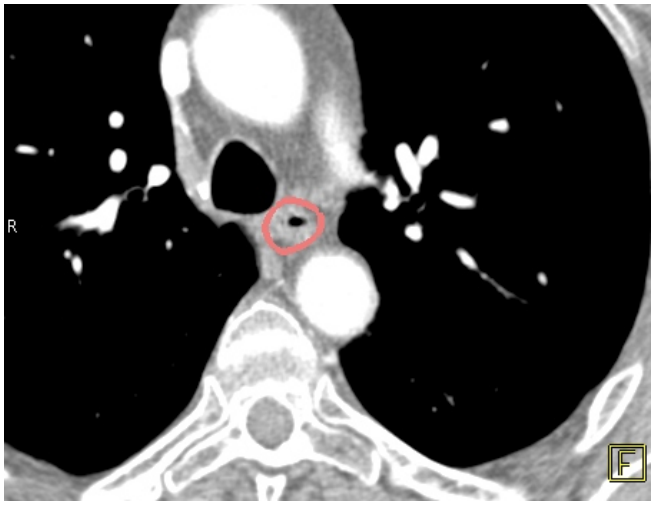


Figure 3.6: Sample results of the proposed method. (a)(c)(e) Selected images from a fluoroscopic sequence with overlaid segmentation results. The last image in each of these rows shows the per pixel minimum image over the sequence, allowing to evaluate the accuracy of the segmentation. (b)(d)(f) The corresponding images from the corresponding bi-plane view direction.



(a)



(b)

Figure 3.7: *Results validation of the proposed method. (a) The 3D reconstructed result once registered to a pre-acquired CT of the patient, rendered from the posterior right patient direction. For comparison, the cyan 3D line indicates the esophagus centerline as extracted from the CT. (b) An MPR view of the CT with a slice of the esophagus.*

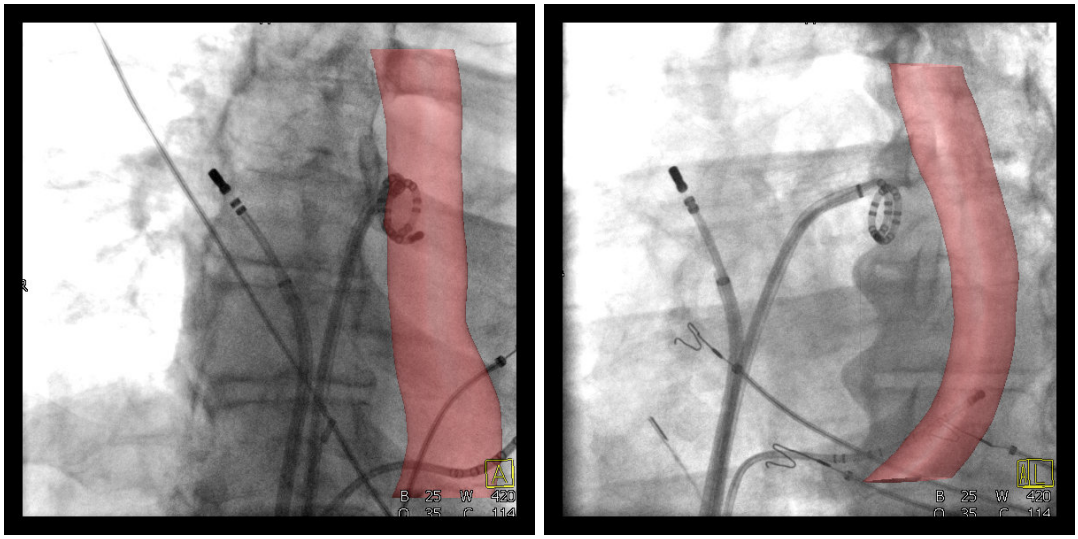


Figure 3.8: Sample superimposing of a graphical esophagus representation on live fluoroscopy. On the left, anterior view angle. On the right, left slightly posterior view angle.

Chapter 4

Automatic Image-based Cardiac and Respiratory Cycle Synchronization and Gating of Image Sequences

4.1 Introduction

Image synchronization and gating are problems faced during the imaging and subsequent processing of organs with quasi-periodic motion, like the heart. The two most common sources of organ motion during imaging are respiratory and cardiac motion. The deformations caused by cardiac and respiratory motion make it difficult to image organs in the thorax and the abdomen. This severely limits the efficacy and efficiency of interventional and radiotherapy procedures performed in this region.

Different approaches have been devised to overcome cardiac and respiratory motion. Cardiac motion is usually handled by ECG gating whereas respiratory motion is usually handled by the use of markers placed on the patients body [57]. The problem with these approaches is that the ECG requires additional hardware, long setup times and in most cases there is a delay between the ECG signal and the image acquisition which makes it hard to make them synchronized. As far as the detection of respiratory

phase is concerned, the placement of markers is usually impractical in a clinical setting; furthermore it is difficult to set up and prolongs the overall acquisition procedure. In addition most laboratories and existing image databases do not contain either ECG or respiratory information. Additionally, since the end goal is to perform image gating, detecting the phase using the images should be more reliable and robust compared to cases where an external signal is used.

Another strategy for reducing respiratory motion is to acquire images using breath-hold techniques [58]. Although this reduces breathing motion by relatively simple and natural means, it is, nonetheless, restricted by the patient's ability to perform a supervised breath hold during the treatment [59]. A third class of strategies addresses the problem of respiratory motion correction by incorporating suitable motion models. Manke et al. [60] proposed a linear parametric model describing the relation between the variation of the diaphragmatic position and the respiratory-induced motion derived from image-intensity based registration for cardiac-triggered 3D MR imaging. King et al. [61] present an affine model which is based on the tracked motion of the diaphragm to compensate for respiratory motion in real-time X-ray images. The main drawback of these approaches is that they require manual landmark selection for diaphragm tracking.

Related Work. Prior work in this area is fairly limited, with [62] presenting an approach for retrospective gating of Intra coronary ultrasound (ICUS) sequences using feature extraction and classification. The method is computationally expensive and requires processing the whole sequence together as some of the features are temporal. In [63], the authors propose a method to analyze images in the sequence and retrieve the cardiac phase by using average image intensity and absolute image difference between the consecutive frames. The method is not very robust and has not been shown to work with real datasets. In [64], the authors propose to detect the respiration (phase) using mutual information. The mutual information (MI) is calculated between the fluoroscopic image and a reference angiogram. The MI calculation is performed only on selected regions of interest (ROI) and the results are variable depending on the choice of ROIs. A similar approach using Normalized MI is presented in [65]. A review of the effects of respiratory motion is presented in [66] and highlights the importance and need for reliable detection of the respiratory phase, specifically for the target organ. In [67], Berbeco et al. detect the breathing phase information by analyzing the fluoroscopic

intensity fluctuations in the lung. However, the approach requires the selection of region of interest(ROI) in a section of the middle of the lung that does not contain the tumor.

Contributions. We propose a novel method to detect the current state of the quasi-periodic system which in turn will enable us to synchronize/gate image sequences to obtain images of the organ system at similar configurations. The proposed method is applicable to 2D and 3D time varying sequences and all imaging modalities. We demonstrate its effectiveness on 2D X-Ray Angiographic and 3D Liver and intra-cardiac Ultrasound sequences.

4.2 Methods

Let $I_t(x, y)$ represent the image at time t of a scene changing with time. According to the phase correlation technique [68], if we represent the Fourier transform of $I_t(x, y)$ by $\mathcal{F}(\xi, \eta)$, and assume that objects in the scene exhibit only translations, then for $I_{t+1}(x, y) = I_t(x - x_t, y - y_t)$, we have

$$\mathcal{F}_{t+1}(\xi, \eta) = e^{-j2\pi(\xi x_t + \eta y_t)} \mathcal{F}_t(\xi, \eta). \tag{4.1}$$

By inverse transforming the ratio of the cross-power spectrum of I_{t+1} and I_t to its magnitude,

$$\frac{\mathcal{F}_{t+1}\mathcal{F}_t^*}{\|\mathcal{F}_{t+1}\mathcal{F}_t^*\|} = \exp(-j2\pi(\xi x_t + \eta y_t)),$$

we obtain a peak at (x_t, y_t). If we consider a simple 1D example as shown in Figure 4.1, we have the original image represented by the blue sinusoidal curve, and applying a translation to the image produces a phase shift in the spectrum.

This approach was used in [69] to compensate for translational motion in coronary angiogram sequences.

It is important to realize that image features can be thought of as being a combination of signals at different frequencies, and that calculating the phase shift of the different frequencies in an image amounts to measuring the motion of image features corresponding to those frequencies. Measuring the phase shift in the frequency domain allows us to estimate the overall motion within an image without explicitly solving the problem of correspondence detection. The phase shift can be calculated in the Fourier

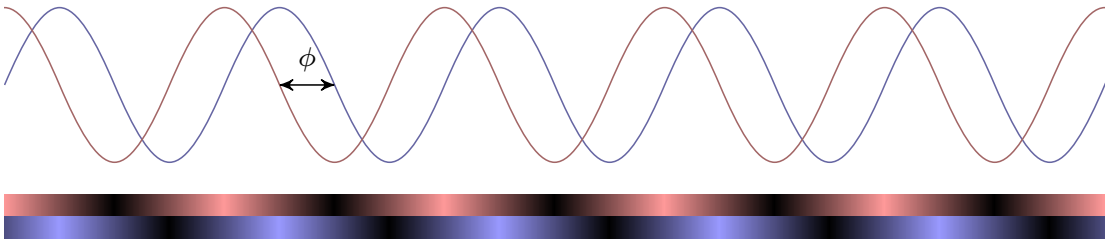


Figure 4.1: The Relationship between object motion and the phase shift (ϕ).

domain [70]. The overall change in the energy within the object being imaged can be estimated by integrating over the phase shift, which is the same as the energy change only due to the motion in the spatial domain because of Parsevals Theorem [70]. It is important to note that this energy is not affected by changes in intensity which might happen as a result of varying levels of contrast agent and differing acquisition parameters. In addition, the proposed method is relatively insensitive to noise, as long as the spectrum of the noise is similar in successive images. As in this usually the case with image sequences acquired on the same scanner during the same acquisition session, the phase detection is robust to the noise levels during acquisition.

We estimate the energy change of the system being imaged by analyzing consecutive frames. The energy change in the system is computed in the spectral domain making the system more robust to outliers that might be introduced in the scene. Common examples are catheters, contrast agents, needles etc. The energy change in the scene is given by,

$$E = \int \frac{\mathcal{F}(I_t)^* \cdot \mathcal{F}(I_{t+1})}{|\mathcal{F}(I_t)^*| \, |\mathcal{F}(I_{t+1})|} df, \tag{4.2}$$

where, $\mathcal{F}(I_t)$ represents the Fourier transform of the image I at time t.

This energy is used as an estimate of the current phase of the organ. The extrema of this signal can be easily detected to detect the cardiac and the respiratory phases. Additionally, if separate cardiac and respiratory signals are required, then a band pass filter centered at 1Hz should recover the cardiac signal, whereas a low pass filter will recover the respiratory signal. The system can be used with N-dimensional by computing the appropriate higher order Fourier transform.

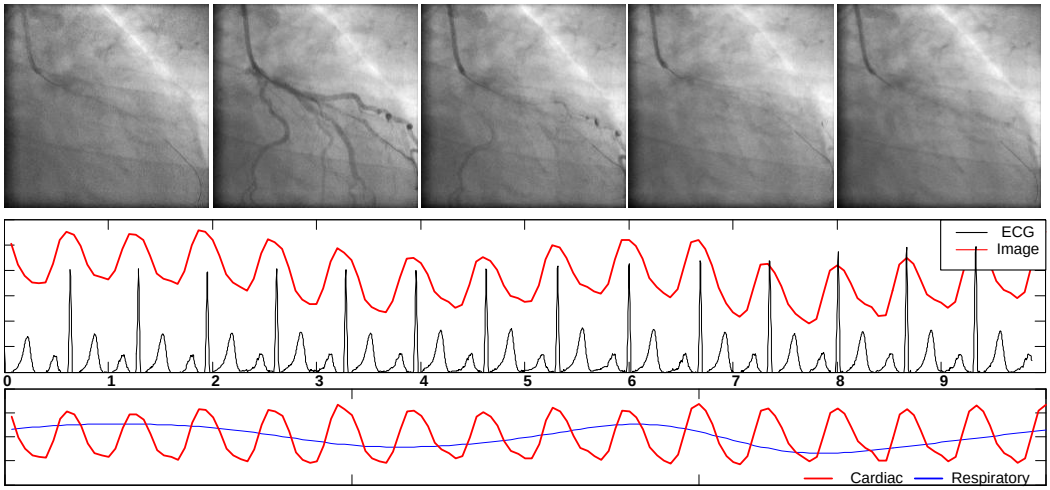


Figure 4.2: Results of Phase detection in X-Ray sequences. The estimated phase is overlaid with the actual ECG signal. The low frequency variation in the phase corresponds to the respiratory cycle. The cardiac and respiratory cycles are also shown in this figure are separated using high and low pass filters respectively.

4.3 Results

In order to test the efficacy and accuracy of the proposed method, we tested it on X-Ray and ultrasound(US) images and compared it against ground truth data. ECG signals were treated as ground truth for the detection of the cardiac cycle, and breathing phase obtained using a magnetic tracking system was treated as ground truth for respiratory motion.

Validation on Cardiac X-Ray Angiography images: We validate our method by comparing it against ECG signals embedded within 20 X-Ray angiographic and fluoroscopic images acquired on a biplane C-arm system (AXIOM Artis, Siemens Medical Solutions, Erlangen, Germany). The minima of the image-based phase correspond to end-diastole and therefore, we compared the alignment of this minima with the P-wave on the ECG signal. Additionally, for the cases where a mismatch existed, a visual comparison of images gated using the 2 methods yielded no obvious winner. An example of the detected phase overlaid on the ECG signal and with sample images is shown in Figure 4.2. Figure 4.2 also illustrates the separation of cardiac and respiratory signals

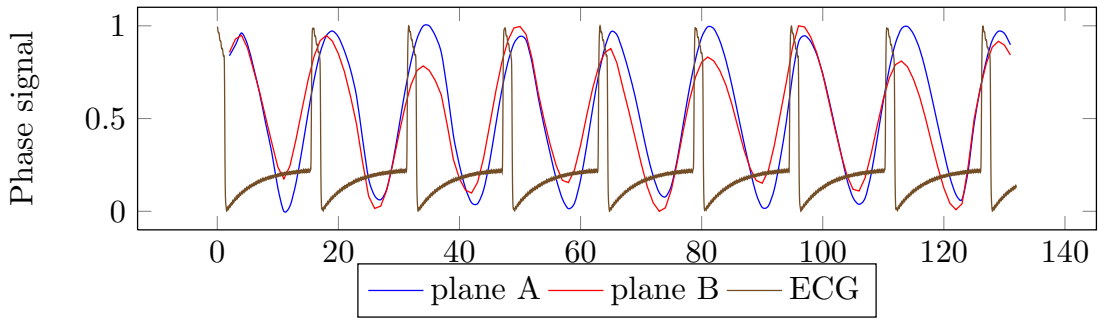


Figure 4.3: Simple validation using biplane sequences. Normalize phase signals from a biplane Xray sequence are plotted along with the normalized ECG signal. It can clearly be seen that the phase is correctly calculated from both views independently.

from the combined phase signal.

We also compared the image-based phase signals from biplane X-Ray images. Since the two images are basically projections of the same heart, the phase signals should match, in spite of differences due to the projection angles. In Figure 4.3 we illustrate one such case in which the phase signals from both planes are plotted with the ECG QRS trigger signal. As can be seen, although there are small variations in the signals from the two planes, they are in overall agreement regarding the cardiac phase.

Validation on Liver 3D Freehand Ultrasound Images: In order to test the efficacy of the system on detecting respiratory motion, we used a Siemens ACUSON Sequoia 512 ultrasound machine (Siemens Ultrasound, Mountain View, CA) with abdominal curved-array probe, and a MicroBird magnetic tracking system (Ascencion Technology Corp., Burlington VT). "Tracked ultrasound" recordings were made in free breathing, with the transducer moving arbitrarily to image any longitudinal and transversal planes of the liver. A second MicroBird magnetic position sensor was attached to the chest and was tracked. The dominant translation of the tracked sensor was treated as a measure of the breathing phase [71]. This was tested on four volunteers and in all cases the image based phase was in agreement with the tracked sensor. An example of this is shown in Figure 4.4.

Validation on Intracardiac Echocardiography images: In order to test the image-based phase detection method on highly dynamic images, we applied it

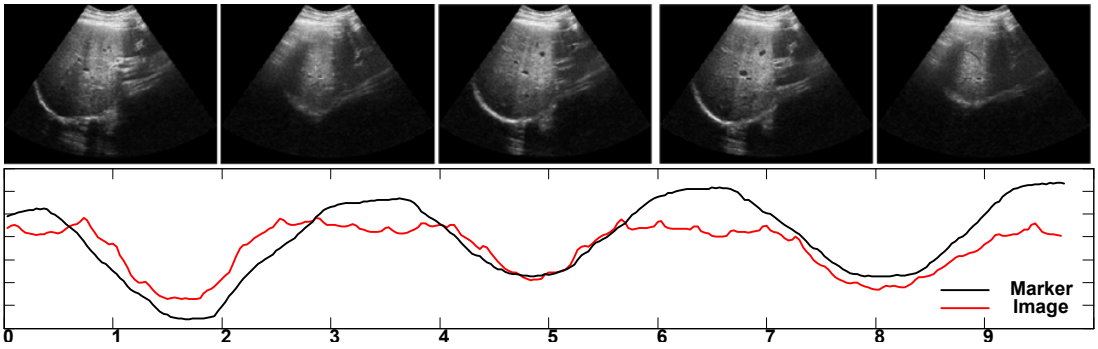


Figure 4.4: Results of Phase detection in Liver ultrasound sequences. The estimated phase is overlaid with the dominant translation of a magnetic marker placed on the chest.

Table 4.1: Correlation ratio of the detected phase against ground truth for angiography and liver ultrasound datasets.

Modality	Compared with	Correlation
2D Angiography	ECG	0.92
3D Liver US	Respiratory tracker	0.88

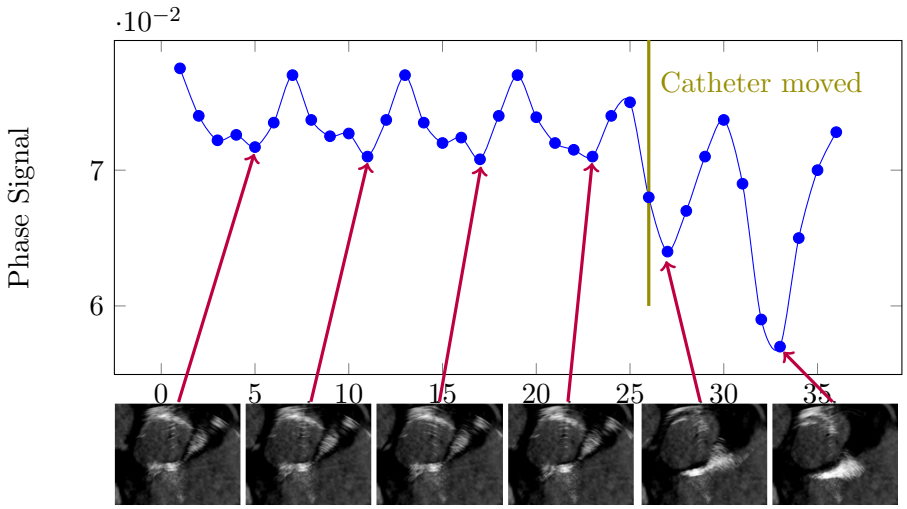


Figure 4.5: Results of Phase detection in intra cardiac ultrasound sequences. The end-diastolic image are detected using the image-based phase detection algorithm. The figure also highlights the effect of a quick catheter move on the image-based phase signal.

on Intracardiac Echocardiography (ICE) images acquired using a Siemens ACUSON AcuNav^{TM}catheter[72]. The images were acquired within the left atrium. Being inside the heart the image acquired by the ICE catheter changes rapidly and is especially pronounced when the catheter is moved. The phase detection does however manage to detect the phase correctly, although in some cases there phase is lost when the catheter is moved rapidly. An example of the phase detection on ICE images is shown in Figure 4.5. The figure also illustrates the effect of a rapid catheter move.,

4.4 Implementation and performance considerations.

The major use of the proposed algorithm will be in real time interventional applications requiring a motion phase signal. Additionally, one of the main problems in using external triggers like ECG signals or displacement transducers is that the delay between the systems can be substantial and will limit the acquisition frequency. For any image based phase detection method to be practical, performance is as important as the accuracy with which the phase can be detected. The proposed algorithm is very fast, and we tested its performance using optimized code written using the Intel Integrated

Performance Primitives for computing the Fourier transforms. We also tested using an implementation using nVidias CUDA library. The performance was evaluated on a Intel Core 2 Duo 2.2GHz notebook with a nVidia Quadro FX 5500 graphics card. We tested using a 1024×1024 angiographic sequence having 132 frames. The optimized software implementation took on an average 2.3 secs, implying a 60 fps capability with 1024×1024 images. The GPU implementation worked slightly faster on the same dataset, taking on an average 1.84 secs to process all 132 frames. Additionally, since the complexity of the proposed algorithm is dependent on Fourier transform of the input images, one easy way to speed up the algorithm further is to define regions of interest.

Chapter 5

Conclusion and Discussion

This work introduces an efficient and robust method for multiple catheter detection. The proposed technique, described in Chapter 2, exploits the clinical setup knowledge to provide search constraints and boost both speed and accuracy. The method involves user input only in the beginning of the case, and runs fully automatically for the rest of the intervention. The method is based on a computationally efficient geodesic framework to trace the sheath and to find one or multiple catheter tips. The method was validated on 1107 fluoroscopic images taken from four patients from different clinics, demonstrating robust multiple catheter tracking at 10 images per seconds while keeping a low false positive rate of 1.06%. For explicit catheter identification, a module can be added operating on the catheters detected by our proposed system, leaving very few tests to perform and thereby very low computational cost. The system's limitation include catheter tip occlusion, which may result in false negatives. We believe extending the system to bi-plane setups would help to overcome this limitation. There is still room to also investigate this method under various clinical conditions, including image quality, dose, and different systems, and this is the subject of future efforts.

This work proposes to reduce the risk of atrio-esophageal fistula formation by superimposing a graphical esophagus representation onto live fluoroscopy. Chapter 3 presents a method for extracting the esophagus silhouette from a fluoroscopic sequence acquired during a barium swallow. Subsequently, the 3D esophagus internal wall is constructed from multiple directions based on silhouettes extracted in each view. The method compounds the barium over the image sequence, segments it, and reconstructs it in 3D. The

segmentation was validated and shown to yield an average 3D error of 1.3 mm. When compared to the existing clinical method, where the barium paste washes out over the course of the procedure, the proposed technique provides both a persistent and more consistent view of the esophagus.

Chapter 4 persented a novel algorithm for the automatic detection of the phase of a moving organ system directly from images. The algorithm is able to detect cardiac and respiratory phases from images of different imaging modality. The algorithm was validated on XRay angiographic, liver US and intra cardiac echocardiography images. ECG signals were used to validate the detection of the cardiac phase, and magnetic markers attached to the chest were used to validate the detection of the respiratory phase. The algorithm is very fast and can process images of size 1024×1024 at rates in excess of 60 fps.

The algorithm should be particularly useful in cardiac and abdominal interventional procedures where cardiac and respiratory motion make localization difficult and challenging. Current attempts at image fusion for such procedures are limited for these reasons. The availability of a reliable image-based phase detection algorithm should make it possible to compensate for intra-operative cardiac and/or respiratory motion. It is possible to use the signal for not only gating but also for real-time dynamic compensation. Real-time dynamic compensation is not currently available mainly because of the high computational complexity of most motion compensation algorithms. Since a pre-operative compensation for different phases of cardiac/respiratory motion can be performed, and the results simply recalled based on the current phase of the system, the image-based phase detection should be most useful for dynamic roadmapping applications for cardiac interventions.

To conclude, I hope that the above description provides a sense of direction for future research of computational method in medical imaging. In this thesis, I have presented our approach to three relavent clinical problem. We have made some progress in this field, but full clinical solutions that help phyisicians and in turn patients are yet to be developed.

References

[1] R.R. Galigekere, K. Wiesent, and D.W. Holdsworth. Cone-beam reprojection using projection-matrices. *Medical Imaging, IEEE Transactions on*, 22(10):1202–1214, October 2003.

[2] H. Scherl, S. Hoppe, M. Kowarschik, and J. Hornegger. Design and implementation of the software architecture for a 3-D reconstruction system in medical imaging. *Proceedings of the 13th international conference on Software engineering*, pages 661–668, 2008.

[3] Matthias John and Norbert Rahn. Automatic left atrium segmentation by cutting the blood pool at narrowings. In *MICCAI (2)*, pages 798–805, 2005.

[4] Cristen Bolan. Multimodality imaging: fusing imaging for the future. *Imaging Technology News*, April 2006.

[5] L. Yatziv, M. Chartouni, S. Datta, and G. Sapiro. Toward multiple catheters detection in fluoroscopic image guided interventions. *Information Technology in Biomedicine, IEEE Transactions on*, 16(4):770–781, 2012.

[6] L. Yatziv, J. Ibarz, N. Strobel, S. Datta, and G. Sapiro. Esophagus silhouette extraction and reconstruction from fluoroscopic views for cardiac ablation procedure guidance. *Information Technology in Biomedicine, IEEE Transactions on*, 15(5):703–708, 2011.

[7] Valentin Fuster, Lars E. Ryden, David S. Cannom, Harry J. Crijns, Anne B. Curtis, Kenneth A. Ellenbogen, Jonathan L. Halperin, Jean-Yves Le Heuzey, G. Neal Kay, James E. Lowe, S. Bertil Olsson, Eric N. Prystowsky, Juan Luis Tamargo,

Samuel Wann, ACC/AHA TASK FORCE MEMBERS, Jr Smith, Sidney C., Alice K. Jacobs, Cynthia D. Adams, Jeffery L. Anderson, Elliott M. Antman, Jonathan L. Halperin, Sharon Ann Hunt, Rick Nishimura, Joseph P. Ornato, Richard L. Page, Barbara Riegel, ESC COMMITTEE FOR PRACTICE GUIDELINES, Silvia G. Priori, Jean-Jacques Blanc, Andrzej Budaj, A. John Camm, Veronica Dean, Jaap W. Deckers, Catherine Despres, Kenneth Dickstein, John Lekakis, Keith McGregor, Marco Metra, Joao Morais, Ady Osterspey, Juan Luis Tamargo, and Jose Luis Zamorano. ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association task force on practice guidelines and the European Society of Cardiology committee for practice guidelines (writing committee to revise the 2001 guidelines for the management of patients with atrial fibrillation): Developed in collaboration with the European Heart Rhythm Association and the Heart Rhythm Society. *Circulation*, 114(7):e257–354, 2006, http://circ.ahajournals.org/cgi/reprint/114/7/e257.pdf.

[8] Hugh Calkins, Josep Brugada, Douglas L. Packer, Riccardo Cappato, Shih-Ann Chen, Harry J.G. Crijns, Ralph J. Damiano, D. Wyn Davies, David E. Haines, Michel Haissaguerre, Yoshito Iesaka, Warren Jackman, Pierre Jais, Hans Kottkamp, Karl Heinz Kuck, Bruce D. Lindsay, Francis E. Marchlinski, Patrick M. McCarthy, J. Lluis Mont, Fred Morady, Koonlawee Nademanee, Andrea Natale, Carlo Pappone, Eric Prystowsky, Antonio Raviele, Jeremy N. Ruskin, Richard J. Shemin, Hugh Calkins, Josep Brugada, Shih-Ann Chen, Eric N. Prystowsky, Karl Heinz Kuck, Andrea Natale, David E. Haines, Francis E. Marchlinski, Hugh Calkins, D. Wyn Davies, Bruce D. Lindsay, Patrick M. McCarthy, Douglas L. Packer, Riccardo Cappato, Harry J.G. Crijns, Ralph J. Damiano, Michel Haissaguerre, Warren M. Jackman, Pierre Jais, Yoshito Iesaka, Hans Kottkamp, Lluis Mont, Fred Morady, Koonlawee Nademanee, Carlo Pappone, Antonio Raviele, Jeremy N. Ruskin, and Richard J. Shemin. HRS/EHRA/ECAS expert consensus statement on catheter and surgical ablation of atrial fibrillation: recommendations for personnel, policy, procedures and follow-up. *Europace*, 9(6):335–379, 2007.

[9] Pascal Fallavollita. Is single-view fluoroscopy sufficient in guiding cardiac ablation

procedures? *Int J Biomed Imaging*, 2010:631264, 2010.

[10] Alexander Brost, Norbert Strobel, Liron Yatziv, Wesley Gilson, Bernhard Meyer, Joachim Hornegger, Jonathan Lewin, and Frank Wacker. Accuracy of x-ray image-based 3-D localization from two c-arm views: a comparison between an ideal system and a real device. volume 7261, page 72611Z. SPIE, 2009.

[11] Lior Gepstein, Gal Hayam, and Shlomo A. Ben-Haim. A novel method for nonfluoroscopic catheter-based electroanatomical mapping of the heart : In Vitro and In Vivo accuracy results. *Circulation*, 95(6):1611–1622, 1997.

[12] Charlotte Eitel, Gerhard Hindricks, Nikolaos Dagres, Philipp Sommer, and Christopher Piorkowski. Ensite velocity cardiac mapping system: a new platform for 3-D mapping of cardiac arrhythmias. *Expert Review of Medical Devices*, 7(2):185–192, 2010, http://www.expert-reviews.com/doi/pdf/10.1586/erd.10.1.

[13] Fred H. M. Wittkampf, Eric F. D. Wever, Richard Derksen, Arthur A. M. Wilde, Hemanth Ramanna, Richard N. W. Hauer, and Etienne O. Robles de Medina. LocaLisa : new technique for real-time 3-dimensional localization of regular intracardiac electrodes. *Circulation*, 99(10):1312–1317, 1999, http://circ.ahajournals.org/cgi/reprint/99/10/1312.pdf.

[14] Daniel R. Elgort, Eddy Y. Wong, Claudia M. Hillenbrand, Frank K. Wacker, Jonathan S. Lewin, and Jeffrey L. Duerk. Real-time catheter tracking and adaptive imaging. *Journal of Magnetic Resonance Imaging*, 18(5):621–626, 2003.

[15] Shirley A. M. Baert, Max A. Viergever, and Wiro J. Niessen. Guide wire tracking during endovascular interventions. *IEEE Trans. Med. Imaging*, 22(8):965–972, 2003.

[16] Felix Bourier, Alexander Brost, Liron Yatziv, Joachim Hornegger, Norbert Strobel, and Klaus Kurzidim. Coronary sinus extraction for multimodality registration to guide transseptal puncture. In Thomas Kahn, Ferenc A. Jolesz, and Jonathan S. Lewin, editors, *8th Interventional MRI Symposium - Book of Abstracts*, pages 311–313, 2010.

[17] YingLiang Ma, Andy King, Nicolas Gogin, C. Rinaldi, Jaswinder Gill, Reza Razavi, and Kawal Rhode. Real-time respiratory motion correction for cardiac electrophysiology procedures using image-based coronary sinus catheter tracking. In Tianzi Jiang, Nassir Navab, Josien Pluim, and Max Viergever, editors, *Medical Image Computing and Computer-Assisted Intervention MICCAI 2010*, volume 6361 of *Lecture Notes in Computer Science*, pages 391–399. Springer Berlin / Heidelberg, 2010. 10.1007/978-3-642-15705-9-48.

[18] Sbastien Knecht, Hicham Skali, Mark D. O'Neill, Matthew Wright, Seiichiro Matsuo, Ghulam Muqtada Chaudhry, Charles I. Haffajee, Isabelle Nault, Geert H.M. Gijsbers, Frederic Sacher, Francois Laurent, Michel Montaudon, Olivier Corneloup, Mlze Hocini, Michel Hassaguerre, Michael V. Orlov, and Pierre Jas. Computed tomography-fluoroscopy overlay evaluation during catheter ablation of left atrial arrhythmia. 2008.

[19] Stijn De Buck, Joris Ector, Andre La Gerche, Frederik Maes, and Hein Heidbuchel. Toward image-based catheter tip tracking for treatment of atrial fibrillation. 2009.

[20] Lars Lickfett, Mahadevappa Mahesh, Chandra Vasamreddy, David Bradley, Vinod Jayam, Zayd Eldadah, Timm Dickfeld, Deborah Kearney, Darshan Dalal, Berndt Luderitz, Ronald Berger, and Hugh Calkins. Radiation exposure during catheter ablation of atrial fibrillation. *Circulation*, 110(19):3003–3010, 2004, http://circ.ahajournals.org/cgi/reprint/110/19/3003.pdf.

[21] Stijn De Buck, Frederik Maes, J. Ector, Jan Bogaert, Steven Dymarkowski, Hein Heidbuchel, and Paul Suetens. An augmented reality system for patient-specific guidance of cardiac catheter ablation procedures. *IEEE Trans. Med. Imaging*, 24(11):1512–1524, 2005.

[22] P. Fallavollita, P. Savard, and G. Sierra. Fluoroscopic navigation to guide rf catheter ablation of cardiac arrhythmias. *Engineering in Medicine and Biology Society, 2004. IEMBS 2004. 26th Annual International Conference of the IEEE*, 1:1929–1932 Vol.3, 1-5 Sept. 2004.

[23] Erik Franken, Peter Rongen, Markus van Almsick, and Bart M. ter Haar Romeny. Detection of electrophysiology catheters in noisy fluoroscopy images. In *MICCAI (2)*, pages 25–32, 2006.

[24] G. Medioni, M.S. Lee, and C.K. Tang. *A computational framework for segmentation and grouping*. Elsevier, 2000.

[25] Alexander Brost, Rui Liao, Joachim Hornegger, and Norbert Strobel. 3-D respiratory motion compensation during EP procedures by image-based 3-D lasso catheter model generation and tracking. In *MICCAI (1)*, pages 394–401, 2009.

[26] C. Molina, G. P. Prause, P. Radeva, and M. Sonka. 3-D catheter path reconstruction from biplane angiography using 3D snakes. In *SPIE - Medical Imaging*, San Diego, California, 1998.

[27] A. Brost, R. Liao, N. Strobel, and J. Hornegger. Respiratory motion compensation by model-based catheter tracking during ep procedures. *Medical Image Analysis*, 14(5):695 – 706, 2010. ¡ce:title¿Special Issue on the 12th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2009¡/ce:title¿.

[28] Peng Wang, T. Chen, Ying Zhu, Wei Zhang, S.K. Zhou, and D. Comaniciu. Robust guidewire tracking in fluoroscopy. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 0:691–698, 2009.

[29] Gregory G. Slabaugh, Koon Kong, Gozde B. Unal, and Tong Fang. Variational guidewire tracking using phase congruency. In *MICCAI (2)*, pages 612–619, 2007.

[30] S. A. M. Baert, W.J. Niessen, E. Meijering, A. F. Frangi, and M. A. Viergever. Guide wire tracking during endovascular interventions. In S. L. Delp, A. M. DiGioia, and B. Jaramaz, editors, *Medical image computing and computer-assisted intervention - MICCAI 2000*, volume 1935 of *Lecture Notes in Computer Science*, pages 727–734, Berlin, October 2000.

[31] Thomas Schoenemann and Daniel Cremers. Globally optimal shape-based tracking in real-time. In *CVPR*, 2008.

[32] I.H. Jermyn and H. Ishikawa. Globally optimal regions and boundaries as minimum ratio weight cycles. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(10):1075–1088, 2001.

[33] Thomas Deschamps and Laurent D. Cohen. Fast extraction of tubular and tree 3-D surfaces with front propagation methods. In *ICPR (1)*, pages 731–734, 2002.

[34] J. A. Sethian. A fast marching level set mmethod for monotonically advancing fronts. In *Proc. Nat. Acad. Sci*, pages 1591–1595, 1995.

[35] John N. Tsitsiklis. Efficient algorithms for globally optimal trajectories. *IEEE TRANSACTIONS ON AUTOMATIC CONTROL*, 40(9):1528–1538, 1995.

[36] M. Boue and P. Dupuis. Markov chain approximations for deterministic control problems with ane dynamics and quadratic cost in the control. *SIAM JOURNAL ON NUMERICAL ANALYSIS*, 36(3):667–695, 1999.

[37] Y.H.R. Tsai, L.T. Cheng, S. Osher, and H.K. Zhao. Fast sweeping algorithms for a class of Hamilton-Jacobi equations. *SIAM JOURNAL ON NUMERICAL ANALYSIS*, 41(2):673–694, 2003.

[38] H. Zhao. A fast sweeping method for Eikonal equations. *Mathematics of Computation*, 74(250):603–627, 2004.

[39] L. Yatziv, A. Bartesaghi, and G. Sapiro. O (N) implementation of the fast marching algorithm. *Journal of Computational Physics*, 212(2):393–399, 2006.

[40] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. *Proc. CVPR*, 1:511–518, 2001.

[41] C. Pappone, H. Oral, and V. Santinelli. Atrio-esophageal fistula as a complication of percutaneous transcatheter ablation of atria fibrillation. *ACC Current Journal Review*, 13(9):37 – 38, 2004.

[42] Taishi Kuwahara, Atsushi Takahashi, Yasuhiro Yokoyama, Atsushi Kobori, Akira Sato, Yoshito Iesaka, Kenzo Hirao, Mitsuaki Isobe, and Kazutaka Aonuma. Importance of esophageal temperature monitoring for the avoidance of esophageal injury

during circumferential left atrial ablation. *Heart Rhythm*, 2(5, Supplement 1):S156 – S156, 2005.

[43] Eric Good, Hakan Oral, Kristina Lemola, Jihn Han, Kamala Tamirisa, Petar Igic, Darryl Elmouchi, David Tschopp, Scott Reich, Aman Chugh, Frank Bogun, Frank Pelosi Jr, and Fred Morady. Movement of the esophagus during left atrial catheter ablation for atrial fibrillation. *Journal of the American College of Cardiology*, 46(11):2107 – 2110, 2005.

[44] M. Galeazzi, S. Ficili, S. Dottori, M. A. Elian, V. Pasceri, F. Venditti, M. Russo, C. Lavalle, A. Pandozi, C. Pandozi, and M. Santini. Pain perception during esophageal warming due to radiofrequency catheter ablation in the left atrium. *J Interv Card Electrophysiol*, 27:109–115, Mar 2010.

[45] Emile G. Daoud, John D. Hummel, Mahmoud Houmsse, David T. Hart, Raul Weiss, Zhenguo Liu, Ralph Augostini, Steven Kalbfleisch, Macy C. Smith, Rohit Mehta, Ashish Gangasani, and Subha V. Raman. Comparison of computed tomography imaging with intraprocedural contrast esophagram: Implications for catheter ablation of atrial fibrillation. *Heart Rhythm*, 5(7):975 – 980, 2008.

[46] Mikael Rousson, Ying Bai, Chenyang Xu, and Frank Sauer. Probabilistic minimal path for automated esophagus segmentation. volume 6144, page 614449. SPIE, 2006.

[47] Johannes Feulner, S. Kevin Zhou, Alexander Cavallaro, Sascha Seifert, Joachim Hornegger, and Dorin Comaniciu. Fast automatic segmentation of the esophagus from 3d CT data using a probabilistic model. In *MICCAI '09: Proceedings of the 12th International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 255–262, Berlin, Heidelberg, 2009. Springer-Verlag.

[48] R. Kobza, A. W. Schoenenberger, and P. Erne. Esophagus imaging for catheter ablation of atrial fibrillation: comparison of two methods with showing of esophageal movement. *J Interv Card Electrophysiol*, 26:159–164, Dec 2009.

[49] Jean Serra. *Image Analysis and Mathematical Morphology*. Academic Press, Inc., Orlando, FL, USA, 1983.

[50] N. Otsu. A threshold selection method from graylevel histogram. *IEEE Trans. System, Man and Cybernetics*, 1(9):62–69, 1979.

[51] G. Sapiro. *Geometric Partial Differential Equations and Image Analysis*. Cambridge University Press, 1 edition, January 2001.

[52] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.

[53] A. Laurentini. The visual hull concept for silhouette-based image understanding. *IEEE Trans. Pattern Anal. Mach. Intell.*, 16(2):150–162, 1994.

[54] Roberto Cipolla and Peter Giblin. *Visual Motion of Curves and Surfaces*. Cambridge University Press, New York, NY, USA, 2000.

[55] Gill Barequet and Sariel Har-Peled. Efficiently approximating the minimum-volume bounding box of a point set in three dimensions. In *SODA '99: Proceedings of the tenth annual ACM-SIAM symposium on Discrete algorithms*, pages 82–91, Philadelphia, PA, USA, 1999. Society for Industrial and Applied Mathematics.

[56] Edith M. T. Dieleman, Suresh Senan, Andrew Vincent, Frank J. Lagerwaard, Ben J. Slotman, and John R. van Soernsen de Koste. *International journal of radiation oncology, biology, physics*, volume 67, chapter Four-dimensional computed tomographic analysis of esophageal mobility during normal respiration, pages 775–780. Elsevier Science Inc., Mar 2007.

[57] A. Khamene, J.K. Warzelhan, S. Vogt, D. Elgort, C. Chefd Hotel, J.L. Duerk, J. Lewin, F.K. Wacker, and F. Sauer. Characterization of internal organ motion using skin marker positions. *Lecture Notes in Computer Science*, pages 526–533, 2004.

[58] M. R. Paling and J. R. Brookeman. Respiration artifacts in mr imaging: reduction by breath holding. *J Comput Assist Tomogr*, 10(6):1080–1082, 1986.

[59] Gikas S Mageras and Ellen Yorke. Deep inspiration breath hold and respiratory gating strategies for reducing organ motion in radiation treatment. *Semin Radiat Oncol*, 14(1):65–75, Jan 2004.

[60] D. Manke, P. Rosch, K. Nehrke, P. Bornert, and O. Dossel. Model evaluation and calibration for prospective respiratory motion correction in coronary mr angiography based on 3-d image registration. *Medical Imaging, IEEE Trans. on*, 21(9):1132–1141, Sept. 2002.

[61] A. P. King, R. Boubertakh, K. S. Rhode, Y. L. Ma, P. Chinchapatnam, G. Gao, T. Tangcharoen, M. Ginks, M. Cooklin, J. S. Gill, D. J. Hawkes, R. S. Razavi, and T. Schaeffter. A subject-specific technique for respiratory motion correction in image-guided cardiac catheterisation procedures. *Med Image Anal*, 13(3):419–431, Jun 2009.

[62] S.A. de Winter, R. Hamers, M. Degertekin, K. Tanabe, P.A. Lemos, P.W. Serruys, J.R.T.C. Roelandt, and N. Bruining. A novel retrospective gating method for intracoronary ultrasound images based on image properties. *Computers in Cardiology, 2003*, pages 13–16, Sept. 2003.

[63] H. Zhu, K.D. Oakeson, and M.H. Friedman. Retrieval of cardiac phase from IVUS sequences. In *Proceedings of SPIE*, volume 5035, pages 135–146, 2003.

[64] B. Martin-Leung, K. Eck, I. Stuke, J. Bredno, and T. Aach. Mutual information based respiration detection. In *International Congress Series*, volume 1256, pages 1085–1092. Elsevier, 2003.

[65] T Moser, J Biederer, S Nill, G Remmert, and R Bendl. Detection of respiratory motion in fluoroscopic images for adaptive radiotherapy. *Physics in Medicine and Biology*, 53(12):3129–3145, 2008.

[66] S. Webb. Motion effects in (intensity modulated) radiation therapy: a review. *Physics in Medicine and Biology*, 51(13):403, 2006.

[67] R.I. Berbeco, H. Mostafavi, G.C. Sharp, and S.B. Jiang. Towards fluoroscopic respiratory gating for lung tumours without radiopaque markers. *Physics in Medicine and Biology*, 50(19):4481, 2005.

[68] CD Kuglin and DC Hines. The phase correlation image alignment method. In *Proc. Int. Conf. on Cybernetics and Society*, volume 4, pages 163–165, 1975.

[69] QX Wu, PJ Bones, and RHT Bates. Translational motion compensation for coronary angiogram sequences. *Medical Imaging, IEEE Transactions on*, 8(3):276–282, 1989.

[70] W Kaplan. *Advanced Calculus*. Addison-Wesley, 4 edition, 1992.

[71] W. Wein, J.-Z. Cheng, and A. Khamene. Ultrasound based respiratory motion compensation in the abdomen. In *MICCAI 2008 Workshop on Image Guidance and Computer Assistance for Soft-Tissue Interventions*, September 2008.

[72] Wolfgang Wein, Estelle Camus, Matthias John, Mamadou Diallo, Christophe Duong, Amin Al-Ahmad, Rebecca Fahrig, Ali Khamene, and Chenyang Xu. Towards guidance of electrophysiological procedures with real-time 3d intracardiac echocardiography fusion to c-arm ct. In *MICCAI 2009*. Springer, September 2009.