Image Registration Between X-ray Fluoroscopy and Transesophageal Echocardiography: Novel Algorithms and Applications for Transcatheter Aortic Valve Replacement

By
Charles R Hatt

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Biomedical Engineering)

at the

UNIVERSITY OF WISCONSIN - MADISON

2015

Date of final oral examination: 6/10/2015

The dissertation is approved by the following members of the Final Oral Committee:

Walter F. Block, Professor, Biomedical Engineering
Amish N. Raval, Professor, Biomedical Engineering
Michael A. Speidel, Professor, Medical Physics
Vikas Singh, Professor, Computer Science
Charles A. Mistretta, Professor, Biomedical Engineering

Abstract

Since it was first performed in humans in 2002, transcatheter aortic valve replacement (TAVR) has emerged as a successful minimally invasive treatment for aortic valve disease. In TAVR, an artificial prosthetic valve is deployed by a catheter, typically under guidance of x-ray fluoroscopy (XRF). Because proper valve positioning is important for achieving optimal clinical outcomes, advanced image guidance systems and workflows that use all available imaging modalities may be able to further improve the success of this procedure.

Recently, image registration between XRF and transesophageal echocardiography (TEE) has been validated and clinically implemented (Philips EchoNavigator). This technology uses image processing to merge the XRF and TEE coordinate systems, allowing the information from both modalities to be fused into a single visualization framework. It is hypothesized that image guidance during TAVR can be improved using XRF/TEE registration by allowing anatomical information from TEE to be combined with device visualization from XRF.

In this thesis, technical contributions aimed at enhancing image guidance and clinical workflows for TAVR using XRF/TEE registration are presented. In the introductory chapter, the history of interventional cardiology is briefly discussed, followed by a description of the TAVR procedure. Novel clinical workflows aimed at improving procedural efficiency and prosthetic valve deployment accuracy are proposed, and specific technical problems involving the proposed workflows are identified and addressed in chapters 2-5.

In chapter 2, a novel implementation of the Hough forest algorithm for object detection is presented and applied to the problem of automatic TEE probe and prosthetic valve detection in XRF images. The purpose of this aim is to minimize the need for user interaction in the image registration process, enabling enhanced clinical workflows and image guidance. In clinical datasets from 48 patients, the TEE probe was successfully detected in 95.8% of images (n=1077) and the prosthetic valve was detected in 90.1% of images (n=388). These results indicate that the presented method is feasible and has potential

for clinical use.

Along with a summary of the technical background and prior work concerning XRF/TEE registration, chapter 3 presents two novel algorithms designed for improved registration accuracy and speed. In chapter 4, these algorithms were validated in simulated, phantom, and clinical datasets. It is shown that the first proposed algorithm was an order of magnitude faster and had a higher success rate than state-of-the-art methods, but was slightly less accurate. The second proposed algorithm was faster and more accurate than state-of-the-art methods, but had a lower success rate. When both algorithms were combined in a hybrid approach, state-of-the-art methods were greatly outperformed in all categories, leading to the first method for XRF/TEE registration that is purely image based (requires no extra hardware), accurate, and fast enough to operate at fluoroscopic frame rates.

In chapter 5, a clinical application of XRF/TEE registration is introduced. A method for contrast-free, intraprocedural optimization of TAVR projection angles using XRF/TEE registration was developed and tested in 10 patients. It was shown that the proposed method agreed with the standard aortographic method to within $3.46^\circ \pm 3.28^\circ$, while a previously introduced method using preoperative CT agreed to within $7.01^\circ \pm 2.78^\circ$. Furthermore, the proposed method can be performed intraoperatively with minimal disruption of clinical workflow and without the use of nephrotoxic x-ray contrast dose.

In the final chapter of this thesis, the potential impact of the presented algorithms are discussed in the context of future image guidance systems for TAVR. Limitations of the current methods and future work needed for clinical translation of the proposed technology are discussed.

Acknowledgments

I would like to thank my advisor, Dr. Amish Raval, for his support during my time in Wisconsin. I often found it remarkable that Dr. Raval was able to handle his very important clinical duties while simultaneously running a successful, multifaceted research program. On top of that, there was not a single moment during the last six years that I saw him without a cheery, relaxed, and positive attitude. I would also like to thank him for working so hard on my behalf to obtain funding each year, his unbelievably fast rough draft turn around times, and for all the coffees from EVP!

I would also like to thank Dr. Michael Speidel, who I worked with a lot during the final two years of my program. Dr. Speidel is a great example of a what a true scientist should be: patient, thorough, and integrous in their work. A lot of my productivity during the final year of my degree would not have been possible without his guidance and help.

As my academic advisor, I would like to thank Dr. Wally Block for helping me navigate the PhD process and for taking the time to talk with me whenever I was in need of advice. Another committee member, Dr. Vikas Singh, deserves special recognition for being enthusiastic in helping me with my research and providing valuable mentorship. I would also like to thank Dr. Charles Mistretta for agreeing to be on my committee at the last minute.

Special thanks goes out to Dr. Vijay Parthasarathy from Philips Research. Vijay was a great mentor during my first three years in Wisconsin, and I learned a great deal while working under him for a semester at Philips Research. Vijay is a person of great integrity and kindness, and I truly owe much of my success to his guidance and help.

I'd like to thank Dr. Ameet Jain, also from Philips Research, for his advice and friendship, as well as Dr. Terry Peters from Western Ontario for inviting me to visit his lab and for his friendliness whenever we ran into each other at conferences.

I'd like to single out Dr. Mike Tomkowiak for all of his help and advice throughout the whole grad school process. Dr. Eric Schmuck was also a

valuable mentor. I'd like to wish good luck to two great guys who are well on their way to finishing excellent PhDs, David Dunkerley and Jordan Slagowski, and thank them for being great lab mates. Also, I wish good luck to Nehal Shah, wherever he currently is, on whatever he is pursuing.

Finally, I would like to thank my family: my brothers for being great friends, my parents for their love and support, and my wife, Carolina, for all the love she gave me when I was struggling, extremely stressed out, or both. I couldn't have done it without you all.

Dedicated to:

My parents, for nurturing my desire to learn and achieve.

My wife, for comforting and supporting me through this challenging but rewarding process.

Contents

List of Figures

List of Abbreviations

AKI: Acute Kidney Injury
CAD: Coronary Artery Disease
Cath lab: Catheterization Laboratory
CBCT: Cone Beam Computed Tomography
CRA/CAU: Cranial-Caudal
CT: Computed Tomography
CUDA: Compute Unified Device Architecture
CVD: Cardiovascular disease
DOF: Degrees of Freedom
DRR: Digitally Reconstructed Radiograph
DSC: Direct Splat Correlation
EAM: Electroanatomic Mapping
EMT: Electromagnetic Tracking
fps: frames-per-second
GCC: Gradient Correlation
GPGPU: General Purpose Computing on Graphics Processing Units
GPU: Graphics Processor unit
HF: Hough Forest
ICE: Intracardiac Echo
IVUS: Intravascular Ultrasound
LAO/RAO: Left/Right Anterior Oblique
LCA: Left Coronary Artery
LV: Left Ventricle
MR: Magnetic Resonance
NCC: Normalized Cross-correlation
PGC: Patch Gradient Correlation
PHV: Prosthetic Heart Valve
PVL: Paravalvular Leak
RCA: Right Coronary Artery
rcGCC: GCC combined with raycasting
rcNCC: NCC combined with raycasting
SAVR: Surgical Aortic Valve Replacement
SCT: Stem-cell Therapy
TAVR: Transcatheter Aortic Valve Replacement
TEE : Transesophageal Echo
TTE : Transthoracic Echo
UW : University of Wisconsin - Madison

XRF : X-ray Fluoroscopy

1. Introduction

Cardiovascular disease (CVD) is the number one cause of death in the world. At 31% of all deaths, it outnumbered cancer, communicable diseases, and fatal injury in 2012 [1]. Due to this fact, investigation of causes and treatment of CVD will remain at the forefront of medical research efforts for the foreseeable future. Most forms of CVD can be treated pharmacologically, through lifestyle changes, or through the use of surgical or minimally invasive interventions. The field of interventional cardiology is concerned with the latter approach, where instruments known as catheters are introduced percutaneously ("through the skin") and manipulated within the cardiovascular system to correct structural defects or administer therapies. Medical imaging for guidance of catheters and visualization of anatomy is critical for successful performance of cardiac interventions. This is especially true for the transcatheter aortic valve replacement (TAVR) procedure.

Before addressing the main motivation for the work in this thesis, the biological and technical background will be introduced, including a brief discussion of cardiac anatomy followed by a history and summary of interventional cardiac procedures and the medical imaging modalities used to guide them.

1.1 The Cardiovascular System

The cardiovascular system is responsible for transporting nutrients and oxygen to the organs and tissues of the body via blood. The heart is the primary organ responsible for pumping blood throughout the cardiovascular system, which is done through a periodic process known as the cardiac cycle. The

heart's electrical system coordinates each phase of the cardiac cycle and helps it operate at maximum mechanical efficiency when healthy. The heart is self-perfused by the coronary arteries. The main branches are the right coronary artery (RCA) and left coronary artery (LCA), which both originate in the aortic root. Diagrams of the heart and the cardiac cycle are shown in Fig. 1.1 and Fig. 1.2, while a detailed anatomy of the aortic root is shown in Fig. 1.3.

Some common diseases of the heart requiring medical intervention include:

- Coronary artery disease (CAD): Narrowing of the coronary vasculature that can result in decreased blood flow to the heart, causing chest pain or myocardial infarction (heart attack). Ischemic heart disease, which covers all heart disease related to lack of blood supply, is the most common form of heart disease in the world.

- Congenital heart defects: Structural heart abnormalities that are present at birth.

- Heart rhythm disorders: Malfunction of the heart's electrical system resulting in irregular cardiac rhythm (arrhythmia), fast cardiac rhythm (tachycardia), or slow cardiac rhythm (bradycardia). Heart rhythm disorders can result in sudden death.

- Valvular disease: Malfunction of one of the four cardiac valves. An example is valvular stenosis, where stiffness and/or narrowing of the valve force the heart to work harder to pump blood, reducing cardiac output and leading to heart failure.

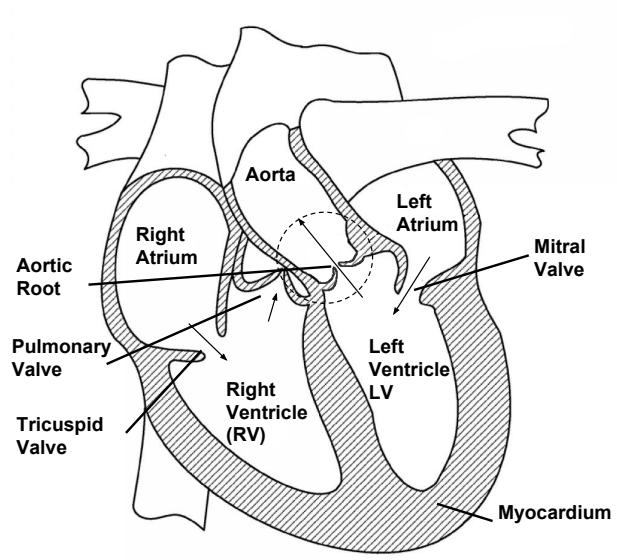

Figure 1.1: Cardiac anatomy.

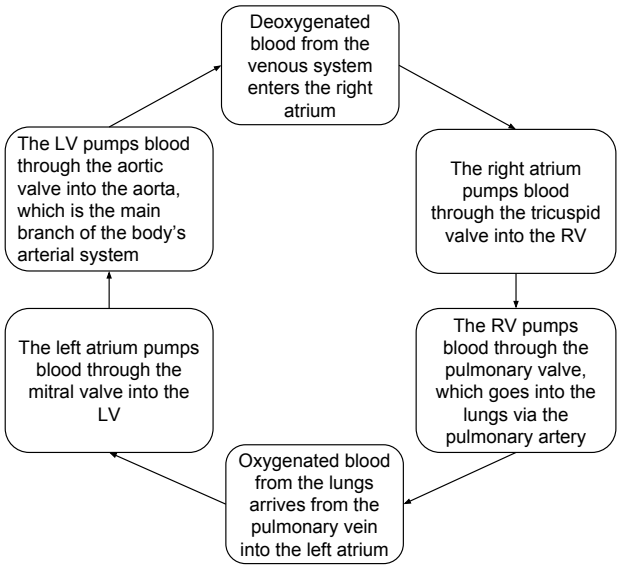

Figure 1.2: The cardiac cycle.

1.2 Interventional Cardiology

Interventional cardiology refers the subspecialty of cardiology that deals with catheter based treatment of structural heart disease. The history of cardiac

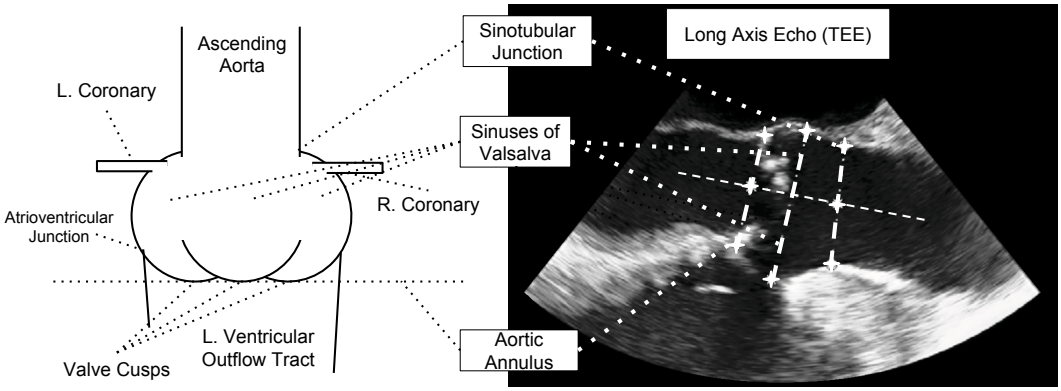

Figure 1.3: Aortic root anatomy, with an example echocardiographic image (from [2]).

catheterization dates back to the 1711 when Stephen Hales recorded pressure measurements from the heart of a living horse using a brass catheter tube. In 1929, Werner Forssmann tricked a nurse into aiding him with an auto-cardiac catheterization, where he advanced a ureteral catheter into his own right atrium and for the first time documented such a procedure by radiography. In 1977, Andreas Gruentzig performed the first balloon angioplasty, which along with coronary stenting is still the gold standard interventional treatment for occlusion of coronary arteries [3].

It should be noted that one of the most important aspects of Forssmann's experiment was that catheter placement was confirmed by x-ray imaging. Image guidance has always been an important aspect of cardiac catheterization.

1.3 Image Guidance for Cardiac Interventions

Interventional cardiology is considered minimally invasive because procedures are completed using only a small incision (normally in the groin) for catheter

introduction, as opposed to opening the chest and operating directly on the heart. Due to the fact that the anatomy cannot be visualized directly, image guidance is a key component of any interventional procedure. Below is a list of the primary imaging modalities utilized in interventional cardiology procedures:

1.3.1 X-ray Fluoroscopy

X-ray fluoroscopy (XRF) is considered the primary imaging modality in the cardiac catheterization lab (cath lab). In an XRF imaging system, a source emits x-rays, which pass through the patient and onto an x-ray detector. The detector in a modern flat panel system consists of a scintillator, which converts the incident x-rays to visible light, and an array of detector elements, which convert light produced by the scintillator to a digital signal representing the image. The source and detector pair are mounted to a movable gantry (aka C-arm, Fig. 1.4), which allows for rotations about the left-right axis (termed a cranial-caudal rotation, CRA/CAU) and inferior-superior axis (termed a left/right anterior oblique rotation, LAO/RAO) of the patient. An XRF system that utilizes two C-arms is called a "biplane" system. XRF produces 2D projection images, where (ignoring the impulse response of the system and propriety image processing pipelines) the image intensity at each pixel is a function of the attenuation line integral along a ray passing from the source to the pixel (see section 3.2.3 for more details). Soft-tissue visualization is generally poor, but x-ray attenuating contrast dyes can be injected into the vasculature to temporarily provide visualization of anatomy.

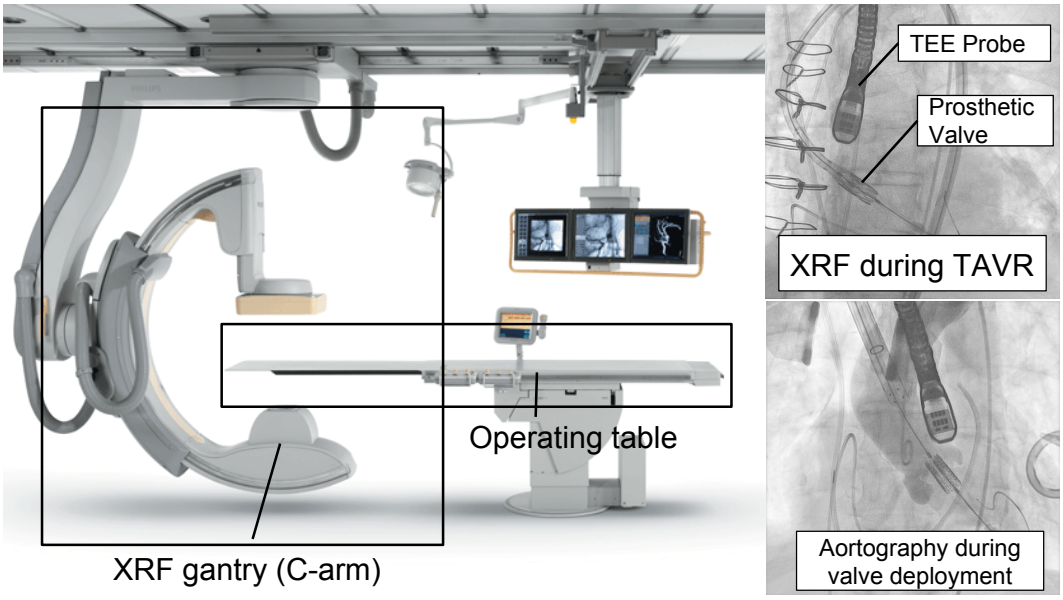

Figure 1.4: Left: An example of a cath lab with an XRF C-arm (Philips FD20) and operating table. Right: Example XRF images during TAVR, with and without contrast dye for aortography.

XRF is a critical imaging modality for guidance of cardiac interventions because of its excellent device visualization and fast imaging frame rate (15-30 fps). The main drawbacks of XRF are that it introduces ionizing radiation to the patient and clinical staff, which can result in skin burns and/or increased risk of cancer, and the fact that devices and anatomy are complicated 3D structures while XRF is an inherently 2D imaging modality.

1.3.2 Echocardiography

The term echocardiography (echo) refers to ultrasound imaging of the heart. Ultrasound imaging operates via the emission of ultrasonic waves from a transducer array, where images are formed based on the timing and intensity of the

received signals. Echo has historically been a high temporal frame rate two-dimensional (2D+t) imaging modality, but in the last decade, technological advances in hardware have allowed for the introduction of 3D+t echo imaging. 3D+t imaging can be volumetric or consist of intersecting 2D planes (biplane). Example of volumetric and biplane echo images are shown in Fig. 1.5.

There are four major types of echo transducers used in echocardiography:

- Transthoracic echo (TTE): External echo transducer that images the heart via the chest wall.
- Transesophageal echo (TEE): An echo transducer that is introduced into the esophagus, where it images from behind the heart.
- Intracardiac echo (ICE): An echo transducer that is embedded within a catheter and images the heart from within the cardiac chambers.
- Intravascular ultrasound (IVUS): Another catheter based imaging device, IVUS is designed for cross-sectional imaging of vessels.

The major advantages of echo are that it can image soft tissue structures in real time (generally 10-50 fps). Also, exposure to ultrasound carries minimal health risks. The main problems with echo are that, aside from a low signal to noise ratio, images often contain artifacts that make manual or automatic identification of anatomical features challenging or in some cases impossible.

Interventional MRI

Interventional magnetic resonance (MR) imaging guided procedures [4] are capable of providing the interventionalist with dynamic, high-resolution images

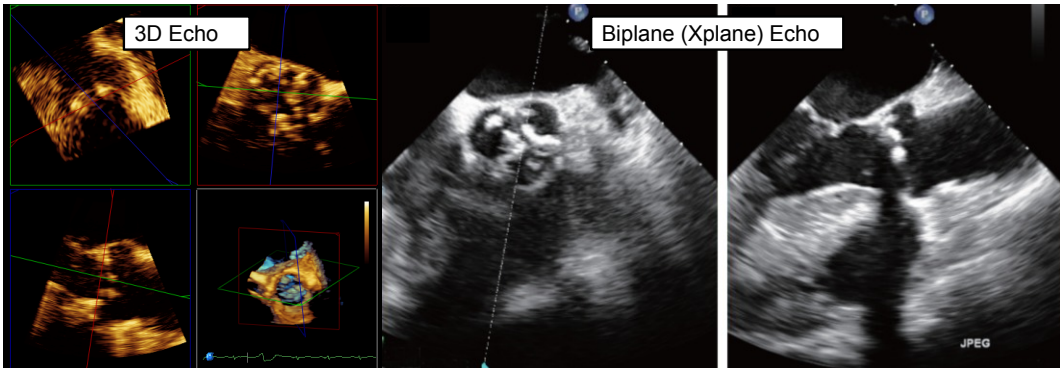

Figure 1.5: Left: 3D echo volume of the aortic valve. Right: Biplane echo of the aortic valve.

of the heart. A comprehensive review of MR guided procedures [5] noted severe limitations, including the ability to monitor patients, risk of device heating causing burns, and lack of commercial testing and development of compatible devices, all contributing to a lack of acceptance for clinical use.

Image Fusion with Preoperative Volumetric Imaging

Preoperative volumetric images are often acquired prior to interventional procedures and map the 3D anatomy and function of the heart. The most common modalities are x-ray computed tomography (CT) and MR. In the past, these modalities were mainly used for diagnosis and preoperative planning, but recent research has demonstrated their use for aiding procedural guidance by registration with intraoperative modalities such as XRF and echo [6, 7, 8, 9, 10, 11, 12].

Electroanatomical mapping

Electroanatomical mapping (EAM) refers to a technology that uses electromagnetic tracking (EMT) sensors to create anatomical roadmaps for guidance of interventional procedures. EMT sensors utilize currents induced in a small wire coil by an electromagnetic field generator for position and orientation localization. By placing these sensors in a catheter and moving the catheter to different positions within a cardiac chamber, a map of the anatomy can be reconstructed. An advantage of this technology is that the catheter is intrinsically registered to the roadmap. Disadvantages are that the mapping step can take a long time (depending on the type of procedure), patient motion can invalidate the roadmap, and the roadmap is often a crude, low resolution representation of the anatomy.

1.4 Complex Interventional Procedures

The most common cardiac interventions are balloon angioplasty and stent placement, as they have become the gold standard of care for CAD. These procedures are routinely guided by XRF alone. However, more complicated procedures have emerged that may require and/or benefit from multiple interventional imaging modalities. A few examples are:

Cardiac Stem-Cell Therapy (SCT) SCT involves the injection of stem-cells into the heart following myocardial infarction. The underlying technology is promising, but still has not been proven to result in improved outcomes for

patients. It is hypothesized that injections near infarcted tissue are most effective. For that reason, and for safety reasons, intraprocedural infarct imaging is required, which can be achieved using preoperative image overlays of delayed enhancement MRI [7, 6] or EAM [13].

Ablation Procedures In order to correct abnormal heart rhythms, tissue causing the abnormal rhythm is destroyed or isolated via ablation with a catheter. To do this, the anatomy needs to be visualized along with the catheter tip location, which is not easily done with XRF alone. Commercial EAM software (Carto 3, Biosense Webster and Ensite NavX, St. Jude Medical) has been used to decrease the procedural complexity and radiation dose associated with these procedures [14].

Transcatheter Aortic Valve Replacement (TAVR) TAVR is a procedure used to replace a malfunctioning aortic valve with a new prosthetic one. In the following section, the TAVR procedure is explained in detail.

1.5 TAVR: Background and Typical Workflow

TAVR is a minimally invasive approach for treatment of aortic valve stenosis that involves implanting a prosthetic aortic valve (aka prosthetic heart valve, PHV) in place of a malfunctioning native valve. An alternative to traditional surgical approaches, TAVR was first performed in a human patient by Cribier in 2002 [15].

The two main PHV device models on the market are the Edwards Sapien

and the Medtronic CoreValve (Fig. 1.6). The Sapien valve is deployed on top of a balloon, while the CoreValve is self-expanding.

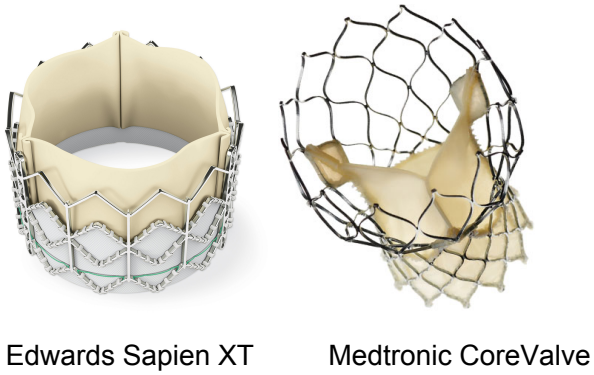

Figure 1.6: Left: Edwards Sapien XT. Right: Medtronic CoreValve

Although the specific clinical workflow for TAVR varies, the one used at the University of Wisconsin (UW) hospital is described below and contains many standard procedural elements.

1.5.1 TAVR Workflow at University of Wisconsin Hospital

1. Preoperative valve sizing and initial estimation of optimal XRF projection angle Prior to the procedure, the patient undergoes CT imaging of the aortic valve. From the CT image, the size of the native aortic valve is measured, which is critical for determining the proper PHV size. CT can also be used to estimate the 3D position of the aortic valve cusps when the patient arrives in the operating room, which in turn can be used to estimate the proper XRF projection views to use during the procedure. Finally, CT helps in determining which access route to use (anterograde or retrograde) via

analysis of the iliofemoral access route [16].

2. Cardiovascular Access In this step, guide wires and the PHV are introduced into the cardiovascular system. The most common techniques are via the aorta (transfemoral retrograde approach) or the apex of the heart (transapical anterograde approach) [17].

3. Determination of the optimal XRF projection A critical component of XRF guided TAVR is finding the optimal XRF projection angle for PHV deployment. Because XRF creates 2D projection images, the gantry angle should be chosen such that the relative positions of anatomy and devices in the XRF image correlate with their true three-dimensional positions. This is best achieved when the source-detector axis is perpendicular to the long axis of the aorta, and parallel to the annulus (Fig. 1.7). There are a number of ways to determine this projection angle, including estimation from preoperative CT, estimation from intraoperative cone beam CT (CBCT), and iterative aortography. At UW hospital, a combination of preoperative CT and iterative aortography is used, where an initial C-arm angulation is determined from preoperative CT and aortography is used to refine the angulation if deemed necessary.

4. PHV deployment Prior to deployment of the PHV, valvuloplasty is performed to dilate the orifice and reduce the mechanical resistance of the native valve. The PHV is then positioned within the aortic annulus, and the valve is deployed. Both valvuloplasty and valve deployment are performed

Figure 1.7: Left: CT volume renderings of an aorta with different virtual XRF projection views. The view in the middle aligns the inferior portions of all 3 coronary cusps, indicating that the annulus is parallel to the source-detector axis, as opposed to the view on the left (image modified from [18]). Right: An example XRF image during an aortogram, where all three cusps are aligned and therefore at the optimal XRF projection angle (image modified from [19]).

under rapid ventricular pacing and breath hold, minimizing cardiorespiratory motion and sudden movements of the PHV/balloon during LV contraction. Proper positioning is critical during deployment. As this is the most challenging step of the procedure, physician experience and imaging are important here.

5. Post-deployment Following deployment, the echocardiographer checks for paravalvular leak and/or regurgitation using Doppler TEE. If the PHV is not functioning properly, another PHV may be deployed within the first one.

1.5.2 TAVR Outcomes and Complications

The PARTNER (Placement of AoRtic TraNscathetER Valves) trial was the first prospective, randomized, and controlled trial for TAVR [20]. One co-

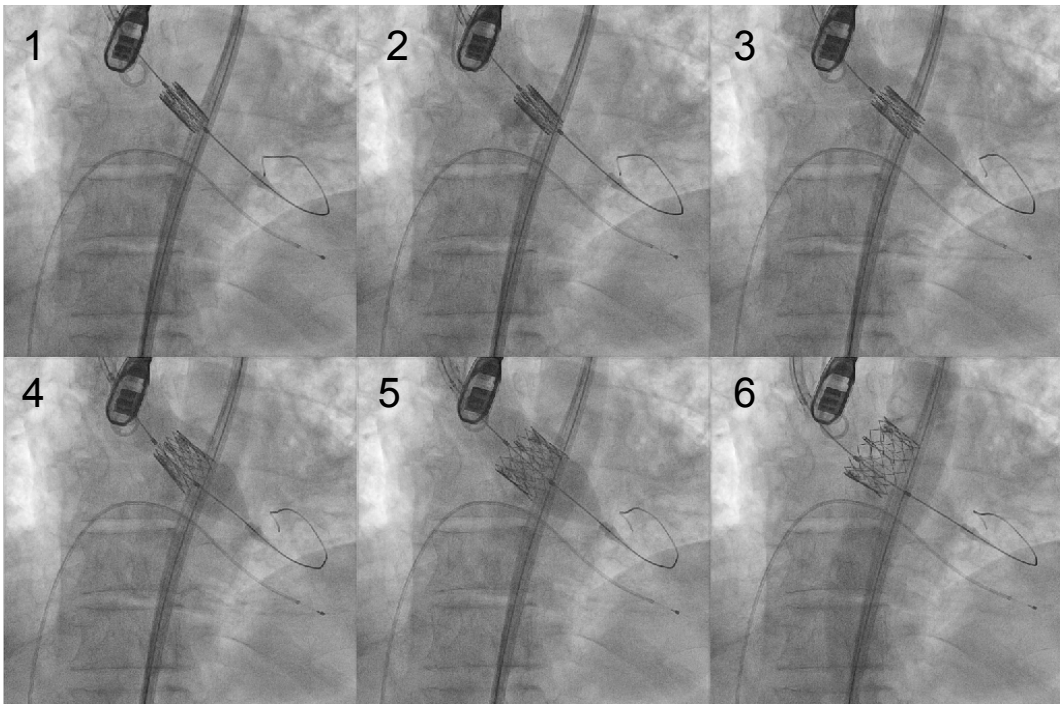

Figure 1.8: XRF guided PHV deployment. The PHV in this example is the Edwards Sapien, and can be seen expanding on top of a balloon catheter.

hort of patients (A) were those that were considered high-risk for surgery and were randomly assigned to either surgical aortic valve replacement (SAVR) or TAVR, while the second cohort (B) were those that were not approved for SAVR and were assigned to either TAVR or standard non-surgical treatment (balloon valvuloplasty).

Within cohort A, it was shown that TAVR and SAVR had similar outcomes in terms of 1-year mortality (TAVR: 24.2%, SAVR: 26.8%, $p = 0.44$) and that patients in the TAVR group were more likely to survive after 30 days (TAVR: 3.4%, SAVR: 6.5%, $p = 0.07$). Major vascular complications were significantly more common in the TAVR group [21] (TAVR: 11.0%, SAVR: 3.2%, $p < 0.001$).

For cohort B, TAVR greatly outperformed standard treatment in terms of

1-year mortality (TAVR: 30.7%, Standard: 50.7%, $p < 0.001$). A difference in mortality at 30 days was not statistically significant (TAVR: 5.0%, Standard: 2.8%, $p = 0.41$). A significant increase in strokes (TAVR: 5.0%, Standard: 1.1%, $p = 0.06$) and major vascular complications (TAVR: 16.2%, Standard: 1.1%, $p < 0.001$) were associated with TAVR at 30 days [20] (TAVR: 11.0%, SAVR: 3.2%, $p < 0.001$).

Many of the TAVR complications resulting in morbidity and mortality are possibly related to suboptimal PHV positioning. If the PHV is deployed too high in the aorta, it can block the coronary arteries or become displaced and migrate into the aorta. If it is deployed too low, it can become displaced and embolize into the LV, requiring surgery for removal. In addition, the PHV may interfere with mitral valve operation, or cause conduction disturbance [22].

Cardiogenic shock is a potential complication that can occur due to hypotension resulting from rapid pacing [23]. Rapid pacing causes the LV to beat rapidly using a pacing catheter, effectively "stunning" the LV, and is used to minimize cardiac motion while obtaining the optimal XRF projection and during PHV deployment. Because of potential cardiogenic shock, rapid pacing during TAVR should be minimized.

Impaired renal function from acute kidney injury (AKI) suffered during TAVR is associated with increased risk of mortality [24]. Although the cause of AKI is often multifactorial, it has been shown to be related to the amount of contrast dose used during the procedure [25, 26].

The PARTNER trial showed that paravalvular leak (PVL) is also an indicator of mortality post-TAVR [27]. PVL results from [28]:

- Incomplete attachment of the PHV to the aortic annulus due to calcifications and/or annular eccentricity.
- PHV undersizing.
- PHV malpositioning.

Due to the importance of proper valve positioning, image guidance is a key component of a successful TAVR outcomes. An image guidance system that can provide enhanced, simultaneous visualization of devices and the aortic root may be able to reduce intra- and post-procedural complications. Furthermore, imaging technologies and clinical workflows that minimize the need for x-ray contrast and LV rapid pacing may also improve outcomes.

1.6 Potential Clinical Workflows for TAVR

Use of echo during TAVR is mainly relegated to an ancillary role [29]. Valve deployment is still mainly guided by XRF, mostly due to ease of image interpretation and device visualization.

Recently, image registration between XRF and echo has been validated [30, 31, 32, 33, 34] and clinically implemented (Philips EchoNavigator). This technology uses image processing to merge the XRF and echo coordinate systems, allowing for image information from both modalities to be fused into a single visualization framework. It is hypothesized that image guidance during the TAVR procedure can be improved using XRF/TEE registration by allowing echo to assume a larger role during the positioning and deployment of the

PHV. In the next section, some example clinical workflows utilizing XRF/TEE registration are considered.

1.6.1 Real-time Anatomical Overlay

One of the main problems with XRF guidance during PHV deployment is the need for aortography to visualize the anatomy of the aortic root, which is critical for proper PHV positioning. Real-time, continuous visualization of anatomical structures without the use of x-ray contrast may allow for more optimal positioning of the PHV and reduce the risk of contrast induced AKI during TAVR.

The workflows in Fig. 1.9 show two possibilities for merging anatomical information from echo with XRF. Both rely on a real-time XRF/TEE registration. In one scenario, anatomical features (such as the aortic annulus) are segmented from the echo image in real-time, and then projected onto the XRF image. In the other, a preoperative CT-based roadmap of the aorta is continuously registered with the echo image, and the roadmap is then projected on the XRF image. By utilizing anatomical information from echo, cardiorespiratory motion of the aortic root is compensated for without the use of x-ray contrast.

1.6.2 Optimal XRF Projection via Echocardiography

As pointed out in sections 1.5.1 and 1.5.2, a critical component of successful XRF guided TAVR is selection of the optimal XRF projection, where the plane of aortic annulus is parallel to the source detector axis. Obtaining this projec-

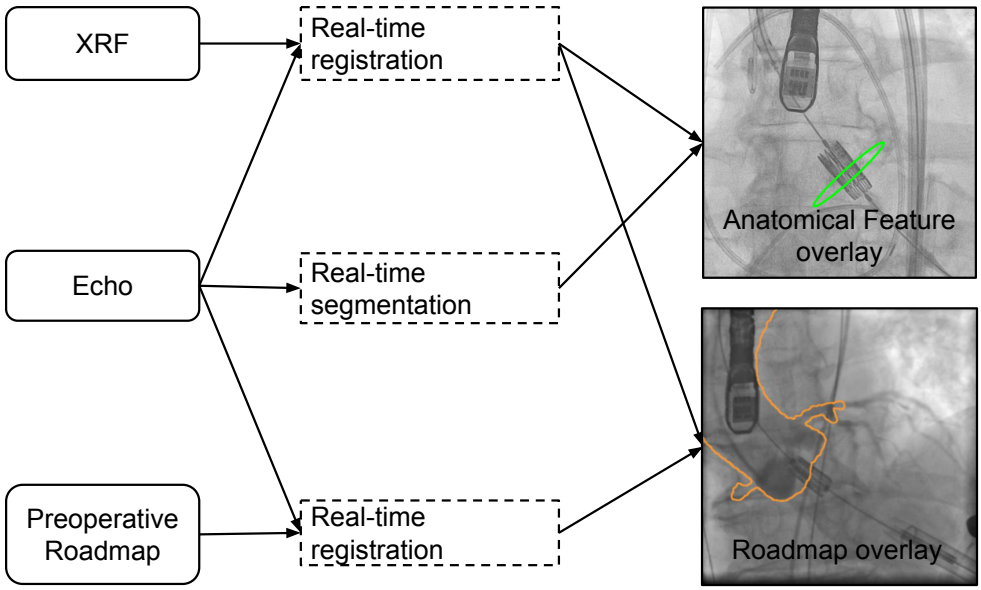

Figure 1.9: Two scenarios for echo-enhanced XRF guidance of TAVR using real-time anatomical overlays. The first scenario merges XRF with segmented anatomical features from echo, while the second scenario merges XRF with a preoperative roadmap via echo as an intermediary modality.

tion is typically achieved by performing repeated aortograms until the inferior tips of the aortic valve leaflets are co-linear in the XRF image (Fig. 1.7). This can often be accomplished using only a single aortogram by experienced physicians. However, due to the distorted and calcified aortic anatomy in some patients, more than one aortogram is often needed to obtain the optimal projection, resulting in a high contrast dose and increased use of rapid pacing. Due to the link between AKI and mortality following TAVR, as well as the risk of cardiogenic shock from rapid pacing, minimizing the use of aortography is important for optimizing clinical outcomes.

The workflow presented in this section is designed to minimize, or even

eliminate, the use of aortography during XRF projection angle optimization. The idea is to segment the aortic annulus in 3D echo, and use XRF/TEE registration to place the annulus in the XRF image space. By doing this, the optimal XRF projection can be obtained by finding the C-arm projection angle where the elliptical annulus projects as line segment in the XRF image. Fig. 1.10 demonstrates the proposed workflow.

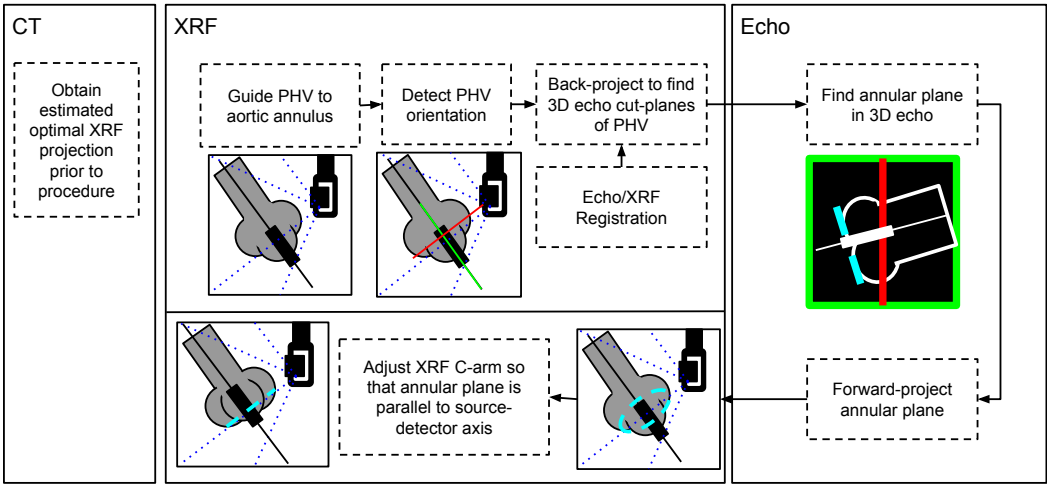

Figure 1.10: Workflow for intraoperative C-arm projection optimization using XRF/TEE registration. Note that the aortic root is visible in the XRF figures merely to demonstrate imaging concepts, but contrast is not necessary for this workflow. Note that the line representing the PHV in XRF can be back-projected to find the PHV centered cut-plane in 3D echo (both are colored green).

1.7 Thesis Contributions

Both of the workflows in Figs. 1.9 and 1.10 require the development and validation of novel algorithms. In particular, the first workflow requires a XRF/TEE registration algorithm that operates at fluoroscopic frame rates

(15 fps). For the second workflow, real-time performance is not necessary, but the method needs to be validated in-vivo against a clinical goal-standard. Both workflows can benefit from algorithms that minimize user interaction. This includes autonomous TEE probe and PHV localization in XRF images.

This thesis presents work aimed at enabling and validating these clinical workflows. The first three chapters outline technical contributions related to device detection and registration in XRF images, while the third clinical validates an application of XRF/TEE registration for TAVR.

Chapter 2: Fully Automatic, Real-Time Localization of Devices in XRF Images A novel Hough forest (HF) implementation for automatic detection of devices in XRF images is presented and validated in clinical datasets. The algorithm is applied to automatic TEE probe and PHV detection for TAVR. The purpose is to minimize user interaction during XRF/TEE enabled TAVR workflows. Validation is performed on XRF datasets from 48 patients.

Chapter 3: XRF/TEE Registration: Prior Work, Technical Background and Novel Algorithms First, a literature review on XRF/TEE registration is presented, followed by a technical summary of medical image registration, 2D/3D registration, and state-of-the-art XRF/TEE registration methods. Next, a `CUDA` implementation of state-of-the-art methods is described, followed by the introduction of three novel real-time methods.

Chapter 4: XRF/TEE Registration: Experimental Validation The novel algorithms presented in chapter 3 are validated *in-silico*, *in-vitro*, and *in-vivo* against state-of-the-art methods. Results indicated that the novel methods out-perform the state-of-the-art methods in terms of accuracy, speed, and success rate.

Chapter 5: A Novel Intraprocedural, Contrast-Free Method for Obtaining the Optimal Fluoroscopic Projection Angles for TAVR A pilot clinical study was designed and carried out that validated a novel method (Fig. 1.10) for determining the optimal XRF projection for TAVR using XRF/TEE registration in 10 patients. The proposed method was compared to the clinical gold-standard and a method that used preoperative CT. It was hypothesized that the novel method would agree with the clinical gold-standard to within a mean of 5°, and that the novel method would outperform the CT method in terms of comparison with aortography.

Chapter 6: Summary and Future Work A summary of this thesis is presented, along with a discussion of future work.

2. Fully Automatic, Real-Time Localization of Devices in Fluoroscopic Images

Detection and pose estimation of devices in XRF images is a challenging but important task for enabling multimodal image fusion for TAVR. This chapter addresses the problem of fully automatic, real-time detection of the Philips X7-2t TEE probe and the Edwards Sapien prosthetic valve in XRF images.

One example application is automatic determination of device-centered echo cut-planes via PHV detection (Fig. 2.1). By combining automatic PHV localization and XRF/TEE registration, the line representing the PHV can be back-projected into the echo volume, providing a device centered cut-plane that is difficult to achieve manually, even for experienced echocardiographers. Furthermore, detection of the PHV in biplane XRF and registration to echo enables the PHV to be virtually rendered within the 3D echo volume (Fig. 2.1), resulting in an alternate imaging tool for guidance of PHV deployment.

A key component of the clinical workflow is automatic localization of the devices at the beginning of an image sequence. In this chapter, a common framework for TEE and PHV localization in XRF images is described. A Hough forest (HF) detector is trained that can detect multiple parts of each device, allowing for estimation of in-plane pose parameters. The data is validated on 1077 clinical images for the TEE probe detector and 388 for the PHV detector.

The work from this chapter is accepted for publication as: Hatt, Charles R., Speidel, Michael A., and Raval, Amish N."Hough Forests for Real-time, Automatic Device Localization in Fluoroscopic Images: Application to TAVR" In Medical Image Computing and Computer–Assisted Intervention-MICCAI 2015. Springer International Publishing, 2015.

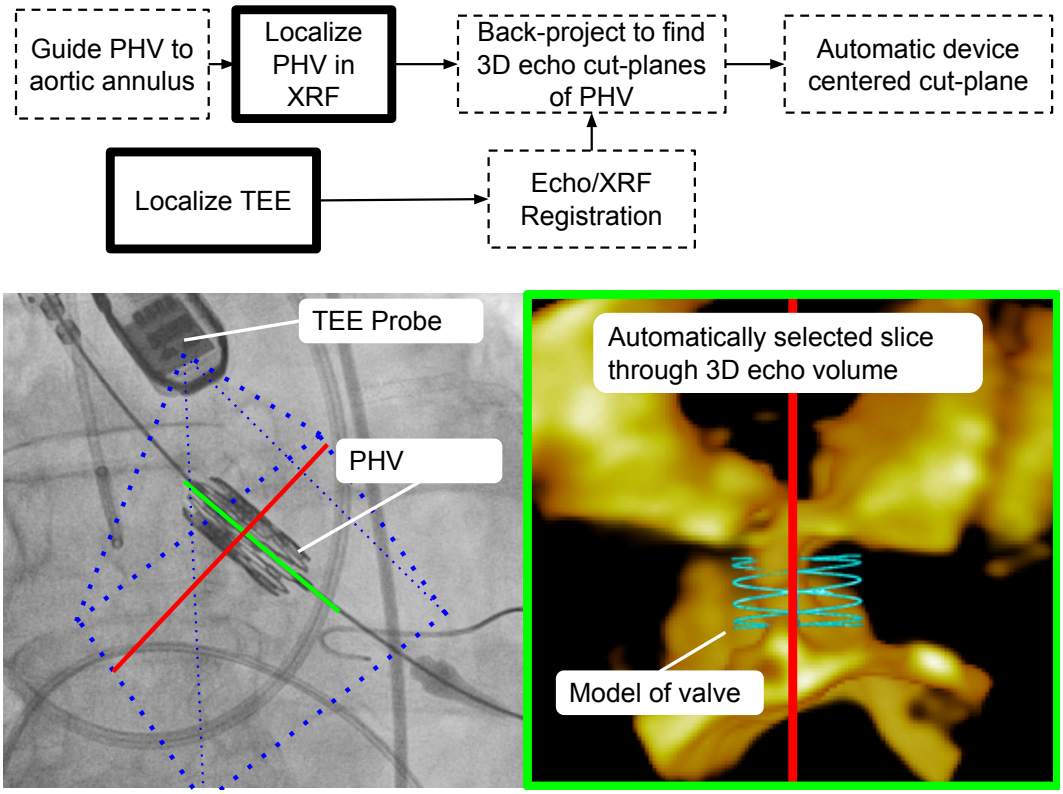

Figure 2.1: Top: Potential workflow enabled by TEE/XRF registration and PHV detection and tracking. Bottom: Rendered model of the PHV displayed within the echo volume following detection and XRF/TEE registration.

2.1 Previous work

2.1.1 TEE probe detection

Two papers concerning fully automatic detection of a TEE probe have been presented in the literature. In [33], the TEE probe was detected using the probabilistic boosting-tree (PBT) approach with Haar wavelets and steerable features [35]. The PBT object detector has been used extensively by Siemens

Corporate Research on many different problems related to automatic anatomy and device detection in medical images [36, 37, 38, 39]. Like most object detectors, the PBT detector can only estimate in-plane parameters - the location and orientation of the object in the image plane. In [33], out-of-plane parameters were estimated using an oriented gradient binary template library. Average TEE pose-estimation time was 0.53 seconds, and a true-positive rate of 0.88 and a false-positive rate of 0.22 were reported for in-vivo experiments. Mean position measurements of 1.10 $\pm$ 0.8 mm and 0.70 $\pm$ 0.6 mm for t_x and t_y were reported for phantoms. Relatively high errors of 11.5° $\pm$ 12.0° and 11.8° $\pm$ 9.8° were reported for rotations about the image axes (θ_x and θ_y, respectively).

In [40] and [41], the work from [33] was extended by focusing on a framework for adapting a classifier generated with *in silico* training data to perform more accurately on *in-vivo* test data. Only in-plane parameters were estimated. Impressive localization results were achieved (position: 0.8 mm $\pm$ 0.5 mm, orientation: 1.4° $\pm$ 1.1°), as well a low false positive rate (0%) and fast detection speed (25 ms).

2.1.2 PHV detection

In the literature, only one paper could be found that addressed the topic of PHV tracking. It should be noted that "tracking" is actually an entirely different problem than "detection," as "tracking "refers to frame-to-frame motion estimation of an already located object, while "detection" refers to automatic or semi-automatic recognition and/or localization of an object. In [42], the

authors introduced a method to track the PHV in fluoroscopic images using a normalized cross-correlation based template tracking method. The purpose was simply to aid in visualization of the PHV prior to deployment, as the method did not track the PHV as it was being expanded. The template tracking method required manual initialization. However, a key component of making computed-aided image guided interventions more usable is the minimization or elimination of user interaction.

PHV detection is similar to catheter detection, for which there are quite a few methods reported in the literature. Most methods deal with detection and tracking of electrophysiology catheters [43, 44, 45], which can be used to compensate for cardiorespiratory motion during interventional procedures. Despite many publications on the topic, there are only a few examples of fully automatic catheter detection. In [46], a method that used filtering combined with sparse coding was employed to automatically detect catheters in real-time. In [47], the authors fused hypotheses from a number of machine-learning based methods to detect coronary sinus catheters.

2.2 Methods

In this chapter, the HF framework for object localization [48] was employed for device detection in XRF images. A key component of the implementation was the simultaneous detection of multiple object parts, which allowed for estimation of both location and orientation of the device. In the following section, the HF object detector is described.

2.2.1 Algorithm

HF detector

A HF is a specific type of random forest that is designed for object detection. A random forest is a collection of decision trees that perform classification and/or regression [49]. HFs take image patches as input, and simultaneously perform both classification (is it part of an object?) and regression (where is the object?). The term "Hough" comes from the idea that each input image patch classified as a part of the object votes for the position of the object center, as is done in the original work on line detection by Hough [50] and later generic shape detection [51]. Votes are added in an accumulator image ("Hough" image, Fig. 2.3), and peaks are considered to be object detection hypotheses. In the original work on HFs [48], the object detector was designed to localize object centers, with the main example being human pedestrians. In the implementation presented in this chapter, the HF was designed to locate two ends of a device, referred to as the "tip" and "tail" (Fig. 2.2).

The HF framework was chosen for the following reasons:

- Because HFs operate by voting from multiple parts, they tend to be robust to object occlusions, which is a common scenario in XRF images due to the presence of wires, bones, and contrast agent injections.
- HFs naturally handle object appearance variations through their training mechanism.
- The data structure used by HFs (trees) are highly amendable to fast,

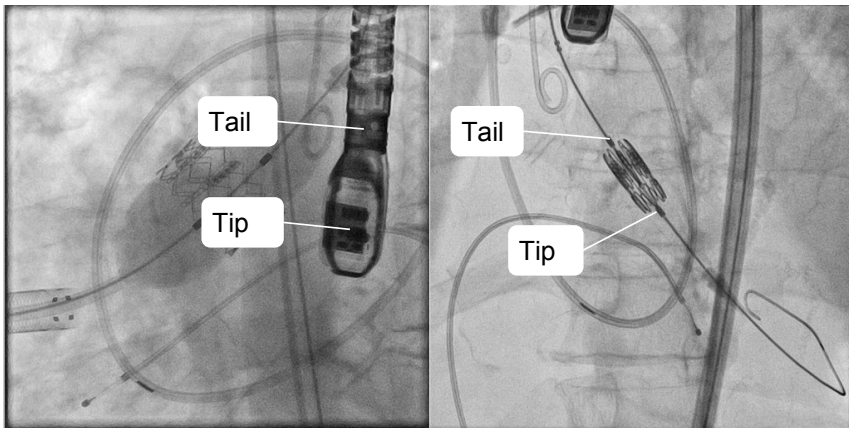

Figure 2.2: Left: TEE probe, with tip and tail labeled. Right: PHV, with tip and tail labeled.

parallel implementation.

Decision Trees A decision tree is an acyclic directed graph where each node contains a single input edge (except at the root node) and multiple output edges (except at the terminal nodes). In random forests, each node only contains two output edges. During testing, data is input into the root node, and rules based on binary tests (aka features) determine which edge to travel down. For image patches, these binary tests typically encode patch appearance. Eventually the data will arrive at a terminal "leaf" node. The leaf node contains data, learned during training, about how to classify (or regress) the input data.

Each tree is trained by computing a set of binary tests on labeled training data, which are used to establish splitting rules. The splitting rules are chosen to maximize class discrimination at each node. In this work, binary pixel comparison tests are used due to their computational efficiency. Binary pixel

comparisons are very simple tests that simply compare two pixels in an image patch to determine if the difference between the two is greater than a certain threshold. Multi-channel image patches are used as input data. This means that, aside from just the raw patch intensities, comparisons can be made between pixels in the x-gradient of the image, y-gradient of the image, gradient magnitude of the image, wavelet filtered image, etc... For each multi-channel input training patch I_n, a set of binary tests are computed as follows:

$$F_{k,n}(p_k, q_k, r_k, s_k, \tau_k, z_k) = I_n(p_k, q_k, z_k) - I_n(r_k, s_k, z_k) < \tau_k \tag{2.1}$$

Where (p, q) and (r, s) are patch pixel coordinates, τ is a threshold used for detecting varying contrast, and z is the channel index. Image channels used in this work were image intensity, the x-gradient and the y-gradient. In order to make sure that the threshold parameter τ can be meaningfully applied to any patch, each channel of each patch is normalized to have a range of 1:

$$I_z(u, v) = \frac{I_z(u, v)}{max(I_z) - min(I_z)}) \tag{2.2}$$

$$\text{Where } I_z \text{ is the patch for channel } z. \tag{2.3}$$

Training Training begins by inputing a $K \times N$ training matrix with N training patches and K tests into a root node (Fig. 2.9). The orientation, offset vectors, and class attributes associated with each patch are known. A metric is calculated for each test (each row). The metric is computed from all N samples. Different metrics are used for classification and regression. In this

work, the metric used for classification is the information gain [52]:

$$G_k^c = H(S) - \frac{|S_1|}{|S|}H(S_1) - \frac{|S_0|}{|S|}H(S_0) \tag{2.4}$$

$$H(S) = -\sum_{c \in C} p(c)log(p(c)) \tag{2.5}$$

Where S is the entire set of training data, S_0 is the set of training data where F_k is false and S_1 is the set of training data where F_k is true, and $H(S)$ is the Shannon entropy over all classes (device or background) in the set S. The purpose of this metric is to find the splitting rule, out of all K splitting rules, that achieves the highest amount of information gain by partitioning the input data into different classes (device or background).

Alternatively, for regression of continuous variables, the metric is:

$$G_k^r = |S|var(S) - |S_1|var(S_1) - |S_0|var(S_0) \tag{2.6}$$

Where $var(S)$ is the variance of continuous data describing the device orientation or offset vectors within each set (non-device patches are ignored for this calculation).

A random decision is made at each node on which attribute to base the splitting rule on: class, offsets, or device orientation. If the offsets are chosen, a random choice about which offsets to regress ("tip" or "tail") is made. The test that gives the maximum value of G_k^c or G_k^r is stored as the splitting rule for that node, and the training data is passed onto the left or right child node according to the splitting rule. The same process is completed until a

maximum tree depth D is reached or all of the samples in a node belong to the background class. The terminal node is termed a "leaf" node, and it stores the classes labels and offsets associated with all of the training data that arrived at that node. In order to speed up run-time, offsets in each leaf node are partitioned into 16 clusters using k-means and the cluster means replace the original offsets. A small tree trained on 128 image patches is shown in Fig. 2.9.

A key feature of HFs is the use of randomness during training, which helps prevent over-fitting the classifier to the training data. This is accomplished by only generating a small random subset of binary pixel tests for each tree, as well as randomizing whether each node will build a splitting rule based on class, offset vector, or device orientation. For example, in the implementation for the TEE probe, only 8192 out of over 1 million binary tests are available to each tree.

During testing, a new image patch centered on (u_p, v_p) is fed into the root node of each tree and traverses the tree according to the splitting rules established during training. When it arrives at a leaf node, each offset (u_o, v_o) in the leaf node votes for the device parts in the Hough image, incrementing the value at pixel (u_H, v_H) accordingly:

$$I_H(u_H, v_H) \rightarrow I_H(u_H, v_H) + \frac{C_L}{|D_L|} \tag{2.7}$$

Where $(u_H, v_H) = (u_p, v_p) + (u_o, v_o)$, C_L is the proportion of device samples in the leaf node, and $|D_L|$ is the number of offsets in the leaf node.

This process is then repeated at every patch and for every tree in the HF. The final Hough image is blurred with a Gaussian kernel and peaks are classified as tip and tail detection hypotheses (Fig. 2.3).

HF input patches can be sampled densely at random locations or sparsely at salient key-points. For our application, it was found that device detection was faster and more reliable using densely sampled patches at random locations.

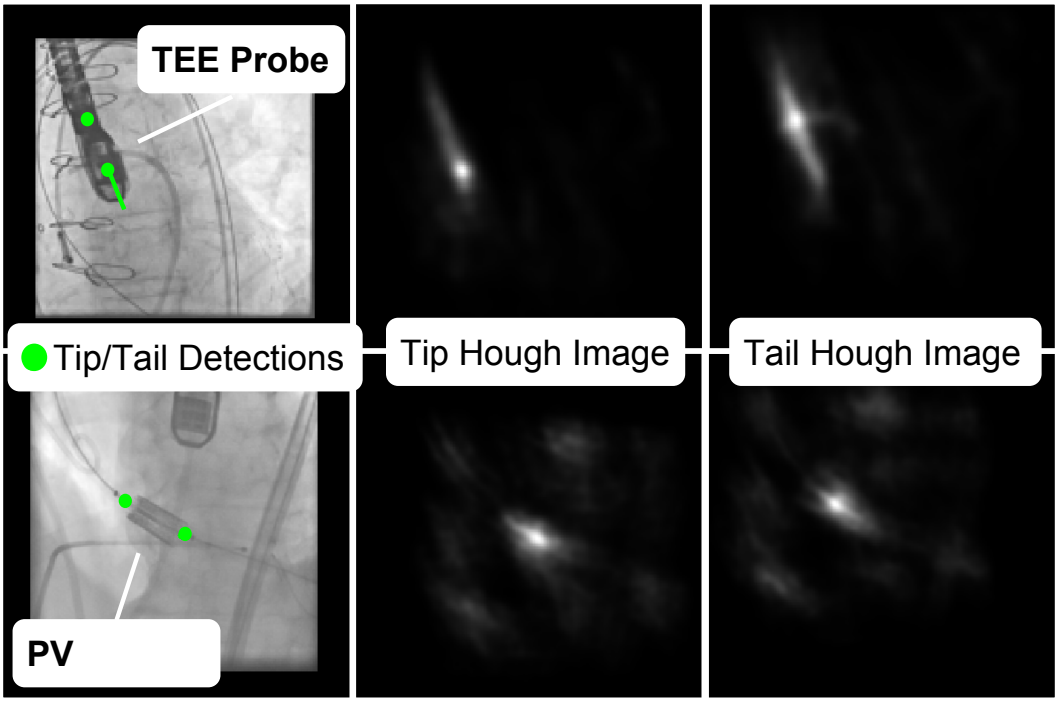

Figure 2.3: TEE probe and valve detection hypotheses with corresponding Hough images showing clearly defined peaks at the tip and tail of the devices.

Hypothesis scoring A Hough image peak was considered a valid hypothesis if it was greater than $0.8 * max(I_H)$ following non-maximum suppression. At most, the top 10 peaks were retained as part hypotheses, but in practice usually only a few peaks survived the first criteria. All L tail and M tip hypotheses

are combined to form $L \times M$ tip-tail pair hypotheses.

Next, unfeasible tail-tip pair hypotheses were removed. This was done by creating tail-tip pair distance and orientation matrices, and removing pair hypotheses that fell outside of the ranges of distance and orientation seen in the training datasets. Remaining tip-tail hypotheses are then given a score:

$$S_{lm} = I_{H_{tip}}(u_l, v_l) \cdot I_{H_{tail}}(u_m, v_m) \tag{2.8}$$

The tip-tail pair with the highest score is selected as the detected device.

2.2.2 `CUDA C` Implementation

The TEE/PHV detection code was implemented in `C` and `CUDA` according to the following procedure:

1. Code initialization: The application starts by reading in a HF data structure from a text file, as well as image patches from a raw binary file. The HF data structure is stored as a `cudaArray` and bound to a 3D `texture`. Patch data was transfered to `global` device memory.

2. HF traversal: The HF was implemented as multichannel 3D `cudaArray` in order to allow for fast `texture` reads from multiple threads (Fig. 2.4). The use of multiple channels allowed for all five node parameters, $((p, q), (r, s), z, \tau)$ and the binary status of whether or not the node was a leaf node, to be stored. The size of the array is $2^{D-1} \times D \times T$, where D equals the maximum tree depth and T is the number of trees in the forest. This limits the maximum depth of the tree to be 12, as the maximum

dimension of a 3D `cudaArray` is currently 2048 elements. Note that this data structure only needs to store information about non-leaf nodes.

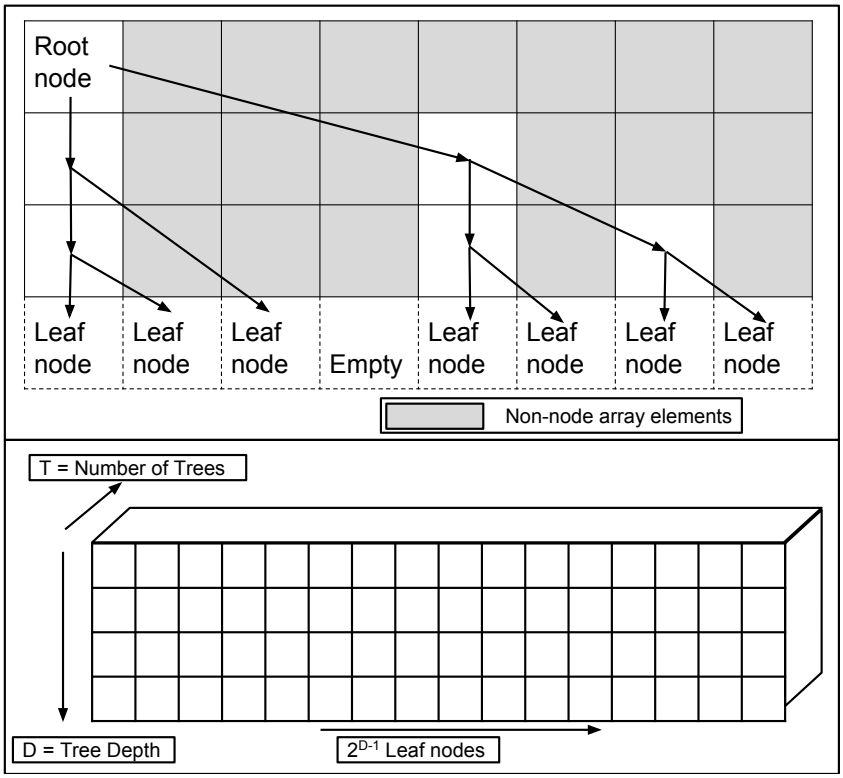

Figure 2.4: Structure of the `cudaArray` used to store the HF. White table entries indicate active nodes. Leaf nodes are shown but it should be noted that only non-leaf nodes are included in the array.

During run-time, a `CUDA kernel` passes each of the input patches though every tree in the HF. The `CUDA kernel` is invoked with N `blocks` and T `threads` (N = number of patches). Each block processes a patch and each thread processes a tree, resulting in $N \times T$ total threads. Before processing the tree, each block uses multiple threads to load the patch data from `global` memory to `shared` memory. This improved memory access performance for asynchronous patch intensity

lookups. The following pseudocode showing how a thread passes through a tree is shown in Fig. 2.5.

```
//Passing a patch through a tree
goleft=true;
col=0;
row=0;
tree=threadIdx.x;
patch=blockIdx.x;
while(d<MXDEPTH)
{
        row = row++;
        col = GetCurColumn(col,goleft,depth)
        node    = tex3D(hfTex,col,row,tree);
        thresh  = node.threshold;
        isleaf  = node.isleaf;
        channel = node.channel;
        goleft = shrPatch[channel][node.q]
        - shrPatch[channel][node.r] < thread;
        d = row*(!isleaf) + MXDEPTH*isleaf;
}
leafIndices[tree][patch]=col;
```

Figure 2.5: Pseudocode for HF patch traversal

3. Hough Image Creation: Once each patch has been processed by the `kernel`, the Hough image is generated. Each leaf votes for an essentially random set of pixels, which is known as a "scatter" operation. Scatter operations are known to be inefficient in a GPU programing environment, because GPUs are not optimized for random write operations to memory, especially when many threads are trying to write to the same pixel. For this reason, the list of HF leaf indices for each patch and tree were

transfered back to the CPU and Hough image generation was completed serially. The tip and tail Hough images were then written to file.

4. Device Detection: Peak detection and hypothesis scoring were performed in `MATLAB`.

2.2.3 Experimental Validation

Computer Hardware and Software

All experiments were run on a Dell Precision T7500 work station running Ubuntu Linux with a 3.47 GHz Intel Xeon processor and a NVIDIA Tesla K20 GPU. The Philips X2-7t probe and the Edwards Sapien valve were used in this study.

Training datasets

Image datasets from real clinical TAVR procedures were used for this study. XRF image sequences were downloaded and processed following approval by the local institutional review board. For the TEE probe, the classifier was trained on simulated XRF images (see chapter 4.3 for details). For the PHV, 389 clinical images from TAVR cases were manually annotated and used for training. In order to increase the size of the training dataset for the PHV detector, each training image was randomly rotated and re-used as if it were a new image. The PHV was only trained and detected in the pre-deployment state.

Table 2.1: HF parameters for the TEE and PHV. N = number of training samples. K = number of tests per tree. T = number of trees. D = tree depth. P = Number of patches at run-time

	Patch size	N	K	T	D	Image res.	P
TEE	17	65536	8192	32	10	1.0 mm	16384
PHV	25	16384	8192	64	8	0.5 mm	16384

Validation The TEE and PHV detector were tested on 1077 and 388 clinical XRF images, respectively. Ground truth data for the TEE images was obtained by manually registering a model of the TEE probe to the image (see Appendix, section B.2). The PHV ground truth was obtained by manual annotation of the tip and tail of the PHV in the test images.

For validation, true positive rate (TPR), mean localization error for true positives, and orientation error for true positives were measured. HF runtime was also reported, which was the amount of time it took for the HF to process all patches for each tree and create the Hough images. A detection was considered a true positive if the localization error was less than 5 mm and the orientation error was $< 10°$. Localization error was the Euclidean distance between the true and measured device centers computed at the detector (i.e. projection magnification was not considered.)

2.2.4 Results

Results are summarized in Table 2.2. True positive detection rate was 95.8% for the TEE probe and 90.1% for the PHV. In addition to summary statistics, error histograms for both devices are shown in Fig. 2.10. Although a true comparison with previous methods from the literature was impossible due to

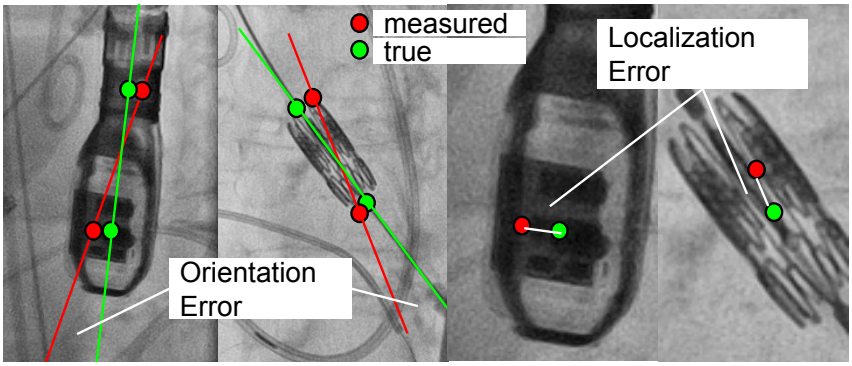

Figure 2.6: Validation error metrics for device detection.

lack of code availability, the reported TPR and localization errors for the TEE probe are competitive with results from the literature [33, 40], especially when considering that the HF was trained on only simulated images. For successful detections, both devices resulted in localization errors less than 1.5 mm on average, and orientation errors less than 3.0°. The histograms indicated that, for the large majority of cases, position errors were less than 5 mm and orientation errors were less than 5°.

Table 2.2: Detection results for the HF device detector

	# Test images	TPR	Localization Error (mm)	Orientation Error (°)	Run-time(ms)
TEE	1077	95.8	1.42 ± 0.79	2.59 ± 1.87	38.8 ± 5.00
PHV	388	90.1	1.04 ± 0.77	2.90 ± 2.37	37.0 ± 2.29

2.2.5 Discussion

The presented method was able to accurately detect both the TEE probe and the PHV in over 90% of images. Most of the false positive detections were due to occlusion from x-ray contrast during aortography. The true positive

rate was high for the PHV, despite the fact that a large percentage of the PHV test images were recorded during contrast infusion. Furthermore, the PHV in the training and testing images varied greatly in size and appearance due to different patient sizes and valve models. This indicates that the Hough forest classifier is robust to appearance variation and that greater detection performance may be possible using a classifier trained on specific valve sizes and models. Examples of challenging images where the device was obscured by contrast agent and/or catheters are shown in Figs. 2.7 and 2.8.

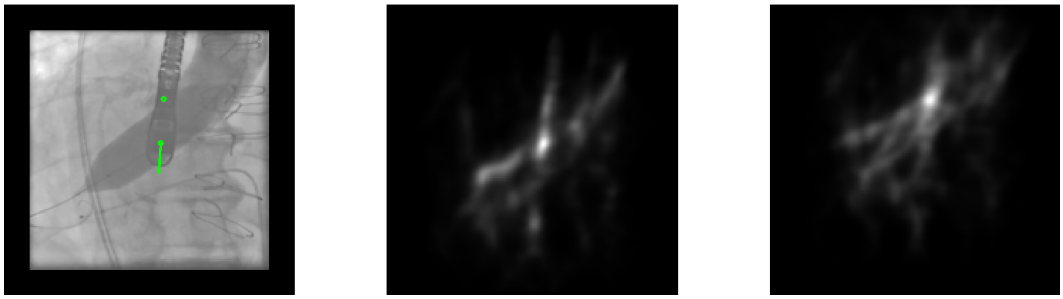

Figure 2.7: Example of a challenging image for TEE probe detection. Left: Image. Middle: Tip Hough image. Right: Tail Hough image.

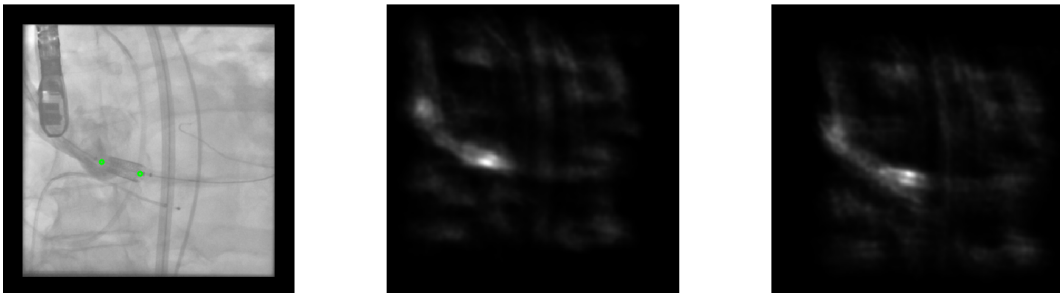

Figure 2.8: Example of a challenging image for PHV detection. Left: Image. Middle: Tail Hough image. Right: Tip Hough image.

The real-time performance of the method is contingent on the full image processing workflow. However, we expect that the bulk of processing is re-

quired by the Hough forest, which was shown to have a maximum run-time less than 50 ms. The other steps, which comprise random patch location generation and extraction, can be implemented very efficiently in `CUDA` using texture reads. It is expected that the full image processing workflow can be completed in less than 60 ms, which is sufficient for typical fluoroscopic imaging frame rates (15 fps)

The main application of these methods is to enable XRF/Echo image fusion, where the PHV will either be rendered in the echo image, or soft-tissue information from echo will be projected onto the XRF image. It is expected that these image processing tools will minimize the need for use of x-ray contrast, which is not only healthier for the patient, but also decreases the risk of device detection failure. For the TEE probe, future work will focus on detection of the out-of-plane pose parameters, which is often a necessary step for fully automatic initialization of 2D/3D registration. For the PHV, future work will focus not only on detecting the PHV location and orientation, but also on the deployment state during implantation. This will allow a dynamic model of the PHV to be rendered in echo images, resulting in new image guidance paradigms for TAVR.

2.2.6 Conclusion

A method for real-time, automatic detection of devices in fluoroscopic images is presented. Based on the Hough forest object detection framework, the method is fully automatic, and has the potential to operate at fluoroscopic frame rates. The percentage of successful device detections was 95.8% for

the TEE probe and 90.1% for the prosthetic valve, despite the presence of x-ray contrast in many of the image frames. Future work will focus on detecting PHV deformation during and after valve deployment for enhanced multi-modal guidance of TAVR.

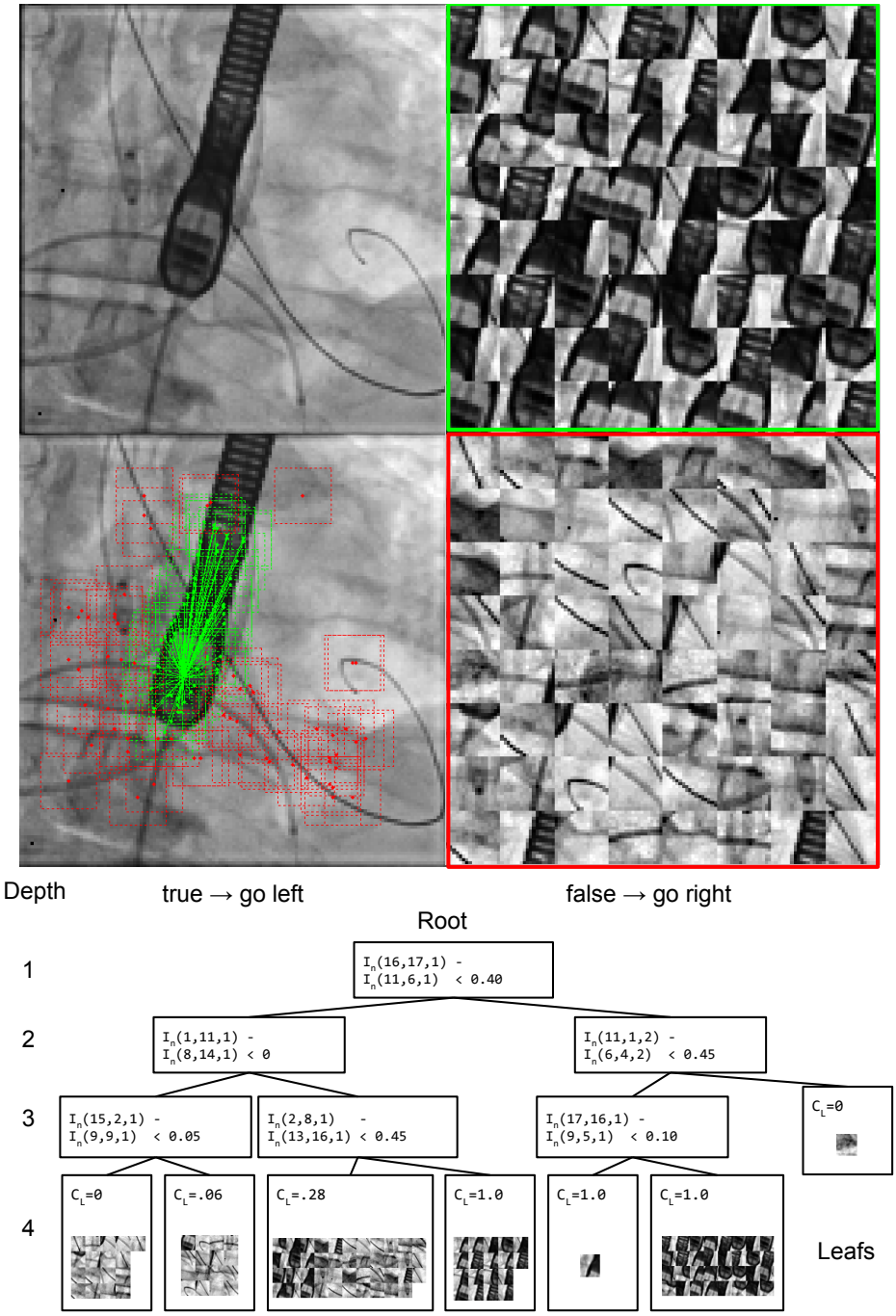

Figure 2.9: A simple example of a decision tree trained on a single image of the TEE probe. Top: Example TEE probe image, with locations of background (red) and device (green) training patches. Bottom: Example of a simple decision tree. Input data traverses the nodes based on binary test results and arrives at leaf nodes. In this example, all of the patches from the training image are shown in their destination leaf nodes.

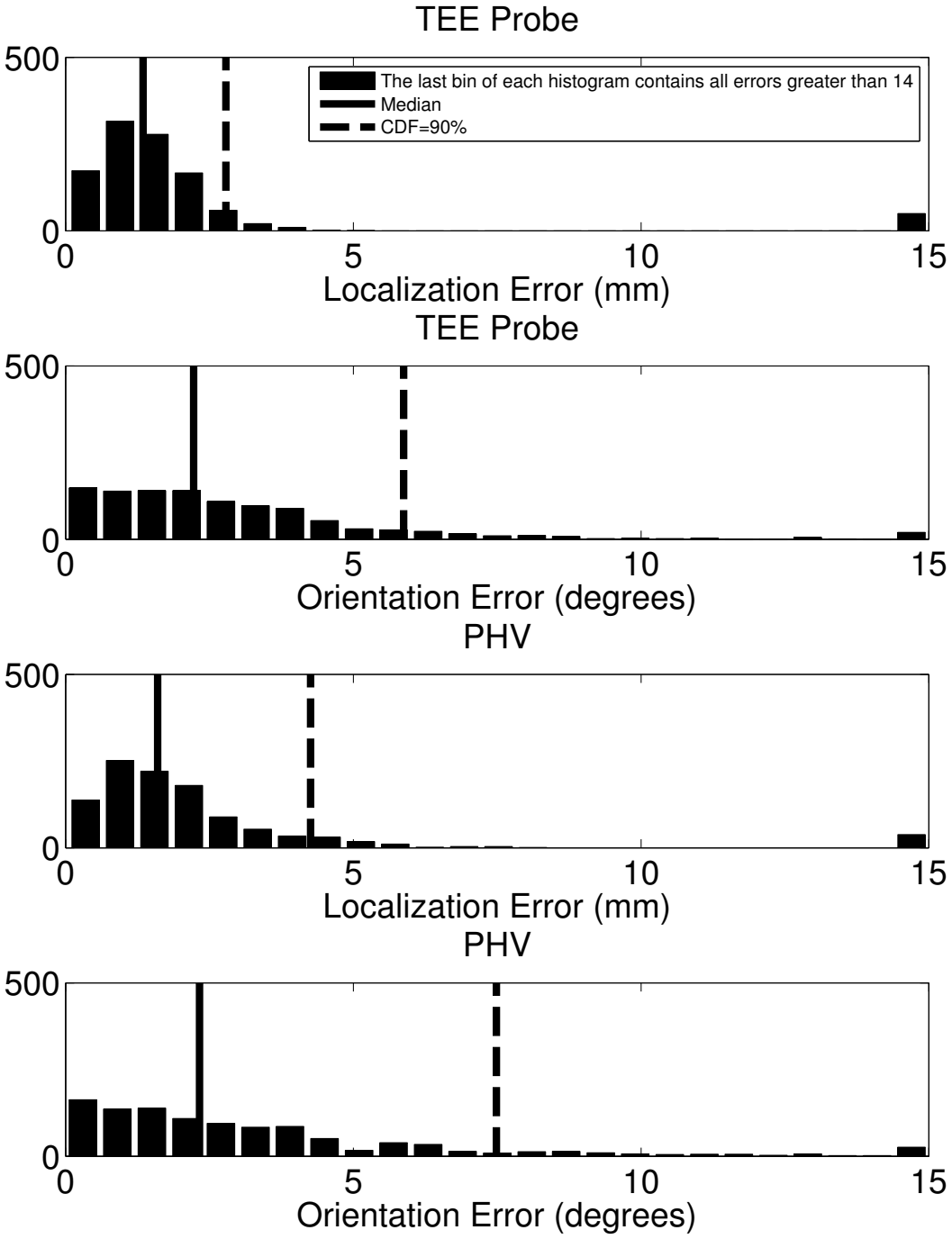

Figure 2.10: Histograms for the device detection error metrics. CDF=90% refers to the error value corresponding to the cumulative error distribution function (CDF) equaling 0.9.

3. XRF/TEE Registration: Prior Work, Technical Background and Novel Algorithms

The focus of this chapter is XRF/TEE registration via image-based pose estimation of the TEE probe. Specifically, we discuss recent work on the topic, provide a technical background on the subject of image registration in general, discuss prior methods of XRF/TEE registration and our optimized implementation of those methods, and finally introduce novel methods aimed at increasing the speed of XRF/TEE registration while maintaining clinically acceptable accuracy.

3.1 Literature Review

3.1.1 XRF/TEE Registration

The first paper to propose a method for XRF/TEE image registration was [53]. In that work, registration was accomplished by:

1. Calibrating the spatial relationship between the EMT sensor on the TEE probe and the echo volume.

2. Attaching EMT sensors to the TEE probe.

3. Registering the EMT and XRF coordinate systems using a specially designed phantom.

Portions of the work from this chapter are accepted for publication as: Hatt, Charles R., Speidel, Michael A., and Raval, Amish N."Robust 5DOF Transesophageal Echo Probe Tracking at Fluoroscopic Frame Rates" In Medical Image Computing and Computer–Assisted Intervention-MICCAI 2015. Springer International Publishing, 2015.

By doing this, the spatial relationship between echo image volume and the XRF coordinate system was known, and the two images could be registered. The main problems with this method were that extra hardware and modifications to the TEE probe were required, and the system was prone to inaccuracies due to electromagnetic fields generated by the XRF system.

The first image based TEE/XRF registration algorithm was presented by Gao et. al. [30, 54]. In this work, the feasibility of using GPU accelerated raycasting techniques to quickly generate DRRs of the TEE probe for 2D/3D registration was demonstrated. The initial results indicated that the method was accurate (mean projection distance = 1.8±1.13 mm). Registration time was 8.0±2.0 s for 1024 × 1024 images and 1.8±0.6 s for 512 × 512 images. The authors found that cardiorespiratory motion caused projection errors up to 4 mm, and noted that a real-time algorithm was necessary to overcome these errors. In follow up work [31], the same group implemented their algorithm on a system that was able to stream echo and XRF data and display both in a common coordinate system. They reported a registration speed-up (0.5-1.0 seconds per registration), a mean projection error of 2.9 mm, and a maximal projection error of 10.9 mm in 13 clinical data sequences.

In [32, 55, 56], radio-opaque fiducials were attached to the TEE probe to perform feature based registration, rather than image based registration as in [30]. The use of fiducials made the registration problem more accurate and faster. A registration time of 50 ms (20 fps) was reported, as well as a point projection accuracy of 1.20±0.91 mm and a 3D point registration accuracy of 2.23 ± 0.72 mm. The main issue with this approach, however, was that

the excellent speed and accuracy results could only be achieved through the addition of the external fiducials, which increased the size of the probe. Aside from the fact that this solution required a custom modification that could not easily be reproduced, it was also undesirable due to increased risk of esophageal injury.

The image-based 2D/3D registration approach was re-visited in [57]. A new method for fast DRR generation was presented, which accelerated the process by modeling the TEE probe as a mesh. This enabled the use of very-fast `OpenGL` rendering primitives for DRR creation. DRR generation was indeed significantly accelerated (0.05 ms), but the overall timing still depended on the computation of the similarity between the DRR and the XRF image. The combined process was reported as 1 ms per iteration. This was much faster than the method reported in [30], but likely not fast enough for registration rates faster than 4 fps, based on the assumption that the mean number of similarity function evaluations is roughly 250. Furthermore, the accuracy of the method was not reported.

It was shown in [30] that using two XRF views can result in more accurate TEE pose estimation. This can be done using a biplane system or imaging the probe from multiple views with a single-plane system. In [58], the problem of accurate multi-view TEE pose-estimation was addressed. The authors developed a method for oblique bi-plane registration that was more robust than previous methods in terms of capture range.

Machine learning based techniques have also been used for XRF/TEE registration, and were discussed in chapter 2.

Based on a review of the literature, there still appeared to be a need for an algorithm that could perform accurate image-based registration at fluoroscopic image frame rates (15 fps) without the use of external hardware or fiducials.

3.1.2 XRF/TEE Registration Applications

Most of the prior work on XRF/TEE registration has focused on methods demonstrating feasibility or presenting incremental improvements. Some applications, however, have been presented.

XRF/TEE registration was used to improve the accuracy of 3D echo mosaicking in [59]. Since 3D echo has as relatively small field-of-view (FOV), previous research efforts have focused on fusing multiple acquisitions to create a larger image. This can be done using echo-to-echo image registration, but the initialization is important for getting this technique to work well. Tracking of the TEE probe in XRF images allowed for accurate initialization, which improved robustness of the multi-view mosaicking approach.

Catheter visualization and segmentation in 3D echo is challenging due to appearance and intensity variations of the catheter and echo imaging artifacts. In [60], a method was presented for improved catheter segmentation in 3D echo using XRF/TEE registration. Catheter segmentation was first performed in XRF, which is a much easier problem. The segmented catheter from XRF was then back-projected into the echo volume using the XRF/TEE registration transformation, which helped initialize and improve the echo based catheter segmentation.

3.2 Technical Background

3.2.1 Medical Image Registration

Medical image registration is the process of aligning images so that there is a maximal spatial correlation between corresponding anatomical features. Typically, one image (the "moving" image) undergoes a spatial transformation, while the other image remains static (the "fixed" image). The standard process begins with estimation of an initial spatial transformation, defined by parameters ϕ. A medical image registration process consists of the following components:

- Spatial transformation (aka warping): The spatial transformation defines the ways in which the moving image can be transformed so that it matches the fixed image. For rigid transformations, the moving image can only undergo translations and rotations, while for affine transformations, it can undergo translations, rotations, anisotropic scaling, and shearing. Deformable transformations define more complicated image warps that are characterized by many local spatial transformations, and tend to have many parameters and are therefore usually slower and less accurate.

- Similarity metric (F_S, aka cost, objective, or energy function): The similarity metric is a scalar or vector value that determines how well the fixed and moving images are aligned. If the images are well-aligned, then the similarity (cost) metric will have a high (low) value. Some common

metrics are sum of squared differences, normalized cross-correlation, and mutual information.

- Optimizer: The optimizer is the specific method used to maximize the similarity between the fixed and moving image. Typically, some initial spatial transformation parameters ϕ_0 are provided, from which an initial similarity (F_{S_0}) can be computed. It is the job of the optimizer to iteratively compute a new set of transformation parameters, evaluate the similarity metric for those new parameters, and from that information, determine how to keep searching the parameter space until the similarity (cost) is at a maximal (minimal) value.

The iterative process of medical image registration is shown in Fig. 3.1, and proceeds as follows:

1. Initialization: An initial set of transform parameters ϕ_0 are chosen.

2. Moving image transformation: The moving image is transformed according to ϕ_0 ($\hat{I}_M = T_{\phi_0}(I_M)$).

3. Similarity function computation: The similarity between the transformed moving image and the fixed image is computed ($F_S(\hat{I}_M, I_F)$).

4. Update parameters: The parameters ϕ are updated by the optimizer and the process repeats (starting at to step 2). The process terminates when a maximum number of iterations have occurred or when changes in F_S and ϕ are smaller than a predefined threshold.

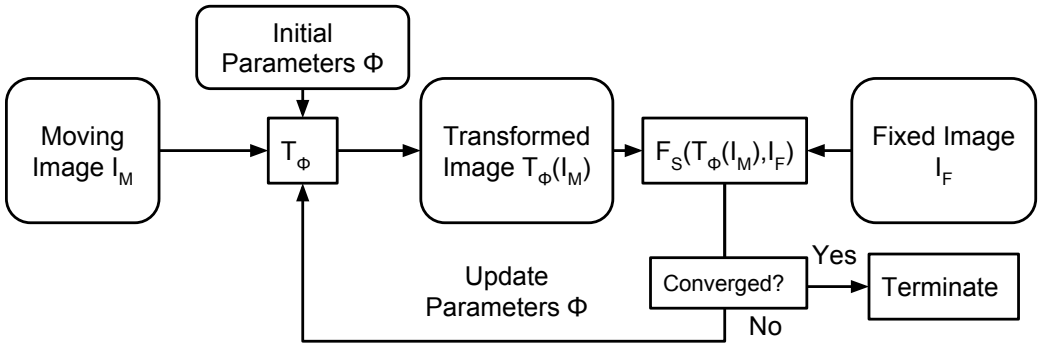

Figure 3.1: Image registration process.

When evaluating an image registration algorithm, three things need to be considered: accuracy, speed, and capture range. These attributes are used to answer the following questions:

- Accuracy: How accurate is the registration? Specifically, what are the alignment errors, usually in mm, between homologous points in each modality after registration? How does this relate to the needs of the clinical application?

- Speed: How fast can the registration problem be solved on a computer? Does it fit the clinical requirements?

- Capture range: How accurately does the registration need to be initialized in order to consistently converge to an acceptably accurate solution? How much user interaction is needed to ensure that the registration will properly converge, and how can failure be detected?

3.2.2 2D/3D Registration

Medical image registration is mostly used to align images that have the same dimensionality. For example, when registering two brain MRI scans, both volumes are three-dimensional. However, there is a class of image registration problems that attempts to align 3D volumes with 2D projection images. This is referred to as "2D/3D registration", also known as "pose estimation" as it essentially attempts to estimate the 3D pose of an object from a set of projection images. The standard process is shown in Fig. 3.2 and explained below:

- Initialization: An initial set of transform parameters ϕ are chosen.
- Moving volume transformation: The moving volume is transformed according to the provided parameters ($\hat{V}_M = T_\phi(V_M)$).
- DRR generation: A DRR (D) is generated given the pose of transformed moving volume. This is done by modeling the projection imaging process using a projection matrix (P, see section 3.2.3).
- Similarity function computation: The similarity between the DRR and the fixed image is computed ($F_S(D, I_F)$).
- Update parameters: The parameters ϕ are updated by the optimizer and the process repeats. The process terminates when a maximum number of iterations have occurred or when changes in F_S and ϕ are smaller than a predefined threshold.

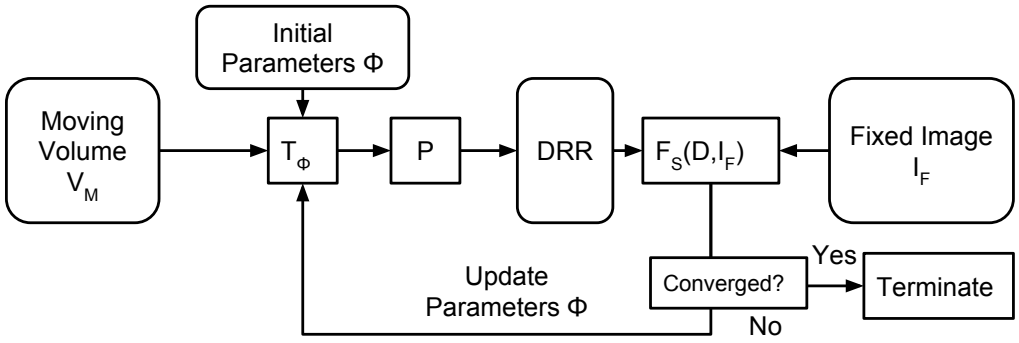

Figure 3.2: 2D/3D image registration process

An exhaustive survey of the topic is found in [61]. The main application of 2D/3D registration is alignment of preoperative data, usually from CT, to intraoperative x-ray images. The moving image in 2D/3D registration is a simulated x-ray image, called a digitally reconstruction radiograph (DRR), and the fixed image is the clinical x-ray image. In the next section, the x-ray image formation process is briefly discussed along with methods for generating DRRs via raycasting.

3.2.3 XRF Projection Geometry

The XRF image formation process is often modeled assuming a point x-ray source and a detector grid. In this thesis, the convention is that the XRF coordinate system is centered on the detector (Fig. 3.3). The following parameters define the projection geometry:

- x_s, y_s, z_s: The coordinates of the x-ray source relative to the origin of the detector. z_s is often referred to as the source-to-image distance (SID).
- p_x, p_y: The detector element spacing of the detector, also known as

"pitch". Expressed in units of $\frac{del}{mm}$ or $\frac{px}{mm}$.

These parameters are used to compute the projection matrix P:

$$P = \begin{bmatrix} \frac{1}{p_x} & 0 & 0 & x_s \\ 0 & \frac{1}{p_y} & 0 & y_s \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{z_s} & 1 \end{bmatrix} \tag{3.1}$$

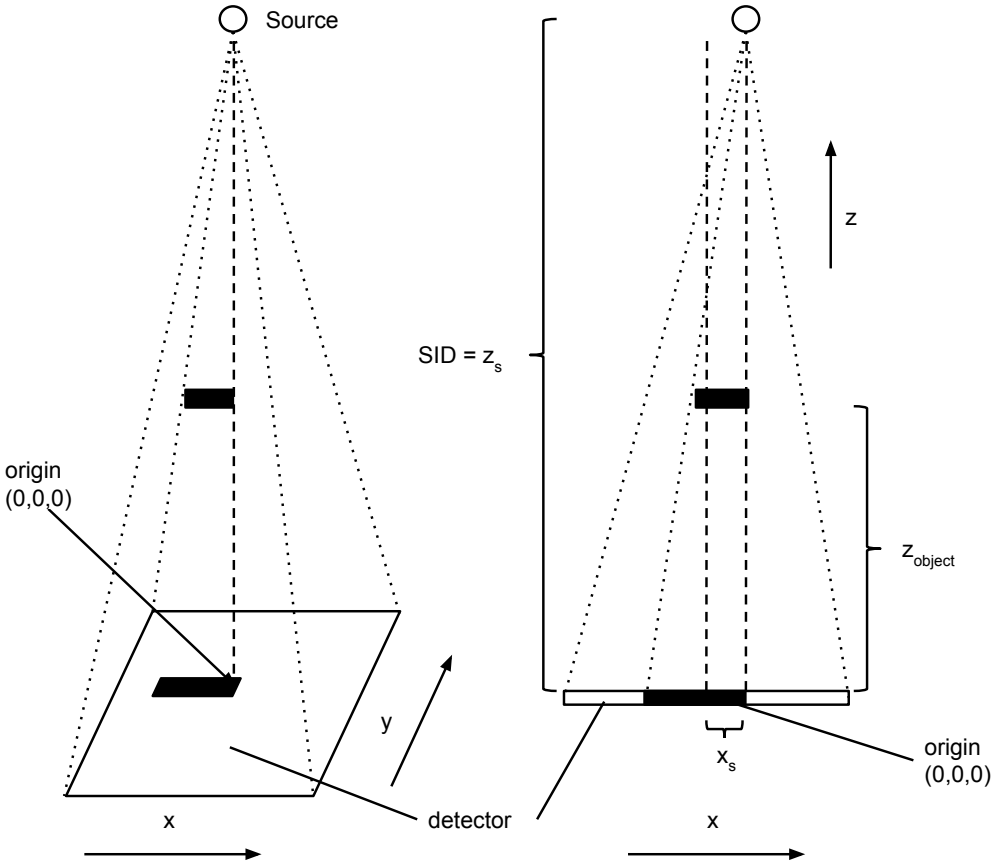

Figure 3.3: XRF coordinate system and projection geometry. The black rectangle illustrates object magnification.

The projection of a point onto the detector ($p_{3D} = [x, y, z, 1]^\top \rightarrow p_{detector} = [u, v]^\top$) is calculated using the following equations:

$$p_{detector} = [u, v]^\top = [\frac{\hat{x}}{w}, \frac{\hat{y}}{w}]^\top \tag{3.2}$$

$$[\hat{x}, \hat{y}, 0, w]^\top = P \cdot [x, y, z, 1]^\top = P \cdot p_{3D} \tag{3.3}$$

From Eqs. 3.2 and 3.3 and Fig. 3.3, note that the magnification of the projected object depends on its z-position.

3.2.4 XRF Image Formation and DRR Generation

During XRF imaging, x-rays are emitted from the source and pass through the patient before arriving at the detector. The image intensity at each pixel ($I(u, v)$) is proportional to the source intensity times the exponential function of the attenuation (μ) line integral along a ray (l) passing from the source to the detector:

$$I(u, v) = I_0 e^{-\int_l \mu(l) dl} \tag{3.4}$$

Clinical XRF images are therefore of function of the attenuation of organs and/or devices such as catheters and TEE probes. To imitate real XRF images, simulated XRF images (DRRs) attempt to recreate this process using volumetric data (usually a CT image) in place of a real object (patient) and discretely sampled virtual x-rays instead of real x-rays. The most natural method of simulation is known as raycasting, shown in Fig. 3.4 and Fig. 3.6.

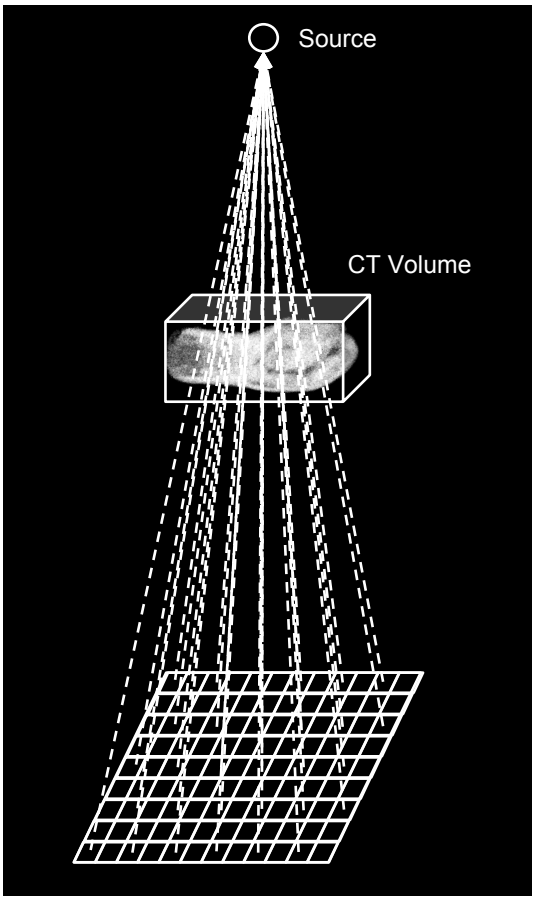

Figure 3.4: DRR generation via raycasting. Line integrals through a CT volume are computed at each pixel.

The equation for generating a DRR (D) via raycasting is:

$$D(u,v) = e^{-\sum_{n \in S} V({}^{V}T_{xrf} \cdot \vec{x_n})\Delta l},\ \vec{x_n} = [x_n, y_n, z_n, 1]^\top \tag{3.5}$$

$$S = \{n | P \cdot \vec{x_n} = [u,v]\} \tag{3.6}$$

where V is the data volume, $\vec{x}$ is a set of discrete points passing along a ray from pixel (u,v) to the source (x_s, y_s, z_s), Δl is the path length between

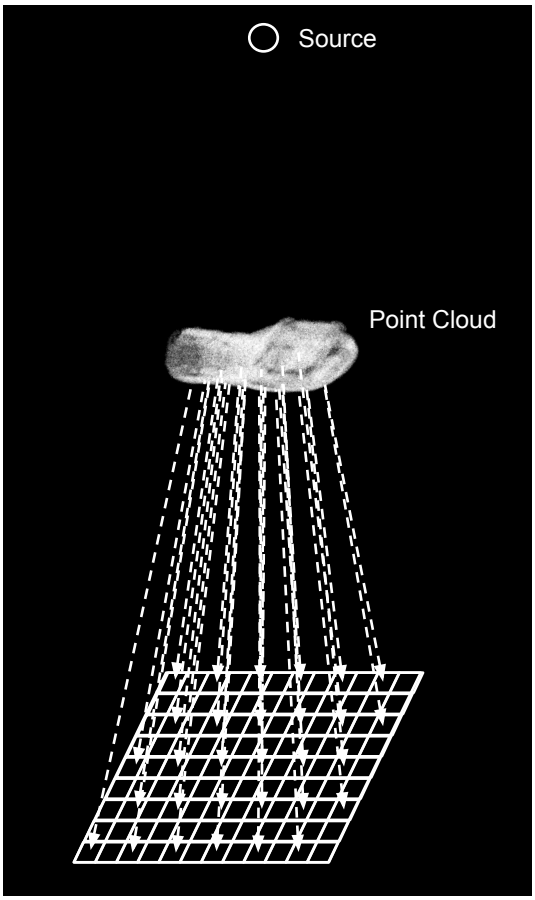

Figure 3.5: DRR generation via splatting. Voxels (points) are projected onto discrete pixel locations and summed.

each discrete point in the ray, and ${}^{V}T_{xrf}$ is the spatial transformation matrix registering the XRF coordinate system to the coordinate system of the data volume.

An alternative DRR generation method is known as splatting (Fig. 3.5 and Fig. 3.6). In splatting, rather than estimating line integrals through a data volume, the voxels from the data volume are considered points and are directly projected onto the image. Wherever a point lands on the image, that pixel is incremented by the intensity value associated with that point. The

equation for generating splat rendered DRRs is:

$$D(u_i, v_i) = \sum_{j \in S_i} -V(\vec{x_j}),\ S_i = \{\ j \mid \lfloor P \cdot^{xrf} T_V \cdot \vec{x_j} \rceil = [u_i, v_i]\ \} \tag{3.7}$$

Here, S_i refers to the set of points that project onto pixel i after the model has been spatially transformed, $\vec{x_j} = [x_j, y_j, z_j, 1]^\top$ is the 3D coordinate of the jth model point, and $V(x_j)$ is the intensity value associated with the jth model point.

One advantage of splatting is that voxels with low intensities do not contribute much information to the final DRR and therefore can be discarded. In some cases, this can lead to a dramatic reduction in the size of the volumetric data needed to generate a DRR. Discarding even the meaningful voxels can also be used to further reduce the volumetric data size, resulting in faster DRR rendering. Fig. 3.6 shows a comparison of TEE probe DRRs generated using raycasting and splatting with multiple levels of point (voxel) downsampling.

3.3 Implementation of XRF/TEE Registration Using Raycasting

3.3.1 Fast Implementation of Raycasting in CUDA

In order to compare the novel algorithms presented in this chapter with prior methods of TEE/XRF registration, it was necessary to implement the al-

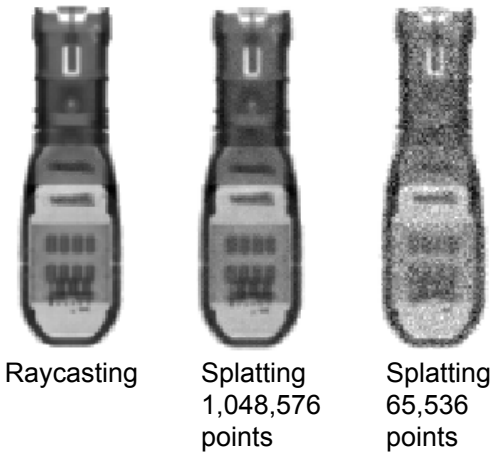

Figure 3.6: Methods for generating probe DRRs.

gorithm from [30]. This required a GPU implementation of raycasting for DRR generation, which was implemented using CUDA. The volumetric image ($CT_{probe}(x, y, z)$) used for generating DRRs of the TEE probe was a 432×150×146 CBCT image with 0.1037 mm^3 pixel resolution acquired with a Philips FD20 system. The raycasting implementation was designed with some special features that improved computational efficiency compared to the method presented in [30] by precalculating a DRR region-of-interest (ROI) and only performing computations at the pixels within the ROI.

The CUDA accelerated DRR generation algorithm is presented below:

1. CPU Program:

   - Transform the corners (Fig. 3.7) of the CT volume to the XRF image space using the spatial transformation parameters $\phi = (t_x, t_y, t_z, \theta_x, \theta_y, \theta z)$:

$$p_{corners}^{(XRF)} = {}^{XRF}T_{CT} \cdot p_{corners}^{(CT)} \tag{3.8}$$

$$T_\phi = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_z & -s_z & 0 & 0 \\ s_z & c_z & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & c_x & -s_x & 0 \\ 0 & s_x & c_x & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_y & 0 & -s_y & 0 \\ 0 & 1 & 0 & 0 \\ s_y & 0 & c_y & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.9}$$

Where $c_\alpha = \cos(\theta_\alpha), s_\alpha = \sin(\theta_\alpha)$.

Record the minimum and maximum z-coordinates of CT bounding box after transformation to the XRF coordinate system:

$$z_{start} = min(z_{corners}^{(XRF)}) \tag{3.10}$$

$$z_{stop} = max(z_{corners}^{(XRF)}) \tag{3.11}$$

- Project the transformed corners of the CT volume onto the image:

$$p_{corners}^{(image)} = (u_{corners}, v_{corners}) = P \cdot p_{corners}^{(XRF)} \tag{3.12}$$

- Record the minimum and maximum projected corners coordinates defining bounding box ROI:

$$B_{ROI} = (min(u_{corners}), max(u_{corners}), min(v_{corners}), max(v_{corners})) \tag{3.13}$$

- Compute the inverse spatial transform ${}^{CT}T_{XRF} = {}^{XRF}\ T_{CT}^{-1}$

- Transfer all spatial transform and bounding box information $({}^{CT}T_{XRF}, B_{ROI}, z_{start}, z_{stop})$ to CUDA `constant` memory.
- Transfer all image information (CT_{probe}, I_{xrf}) to CUDA `texture` and `global` memory.

2. GPU Program:

- Launch 1 thread for every pixel within the bounding box B_{ROI}.
- Set $DRR(u, v)$=0.
- Trace a ray r through the volume according to the following procedure:
  - Compute the current position of the ray: $r_{current} = (x_{current}, y_{current}, z_{current})$. Set $z_{current} = z_{start}$. Compute $x_{current}$ and $y_{current}$ from the pixel position and the XRF system geometry P.
  - Spatially transform $r_{current}$: $\hat{r}_{current} = {}^{CT}T_{XRF} \cdot r_{current}$.
  - Sample and record the CT value: $DRR(u, v) += CT(\hat{r}_{current})$
  - Set $z_{current} = z_{current} + \Delta z$
  - Repeat until $z_{current} > z_{stop}$
- Set $DRR(u, v) = exp(-DRR(u, v))$

Note in Fig. 3.7 that the projected bounding box is much smaller than the entire image, which results in a much faster DRR generation step compared to when rays are cast at every pixel.

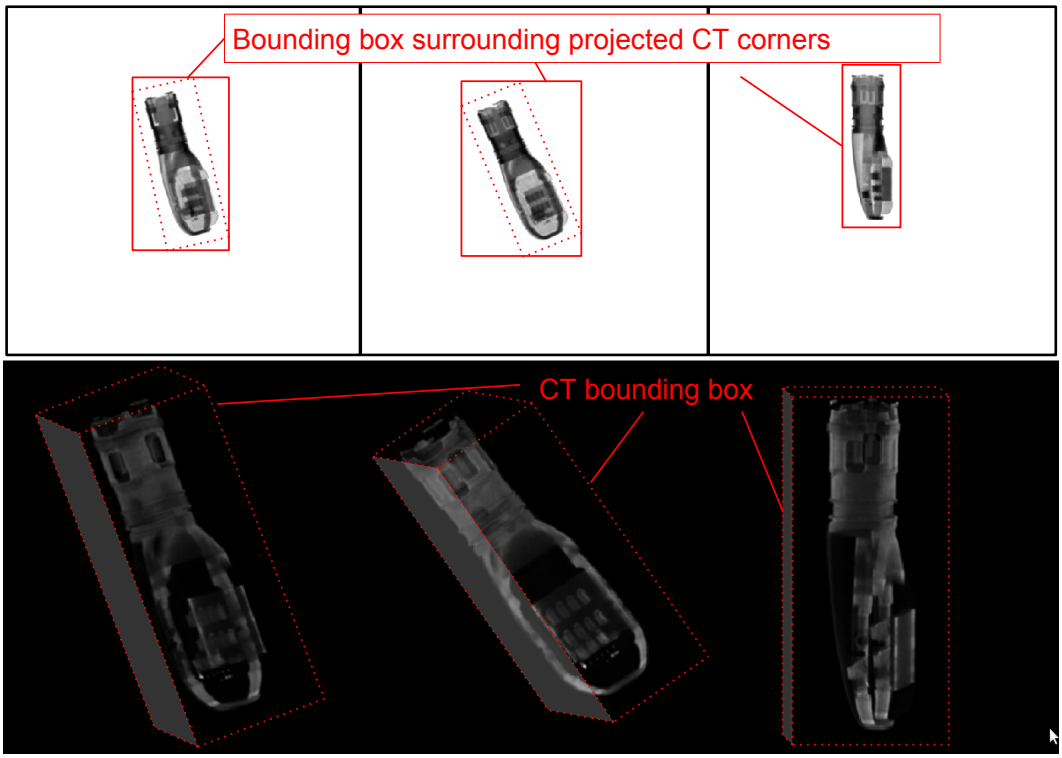

Figure 3.7: Illustration of the probe CT image used to generate the DRRs via raycasting, as well as the bounding box scheme for acceleration of the procedure.

3.3.2 Similarity Metrics

In this work, the normalized cross-correlation (NCC) and gradient correlation (GCC) similarity metrics were used due to their relative computational simplicity, as well as the fact that these metrics were used in prior work on XRF/TEE registration [30, 55]. They are defined with the following equations:

$$F_{NCC}(I, D) = \sum_i \sum_j \hat{I}(u_i, v_j)\, \hat{D}(u_i, v_j) \tag{3.14}$$

$$\hat{I}(u_i, v_j) = \frac{I(u_i, v_j) - \mu_I}{\sigma_I} \tag{3.15}$$

$$F_{GCC}(I, D) = F_{NCC}(\frac{dI}{du}, \frac{dD}{du}) + F_{NCC}(\frac{dI}{dv}, \frac{dD}{dv}) \tag{3.16}$$

These functions were implemented in CUDA using the parallel reduction method presented in [62] for mean and standard deviation calculation and finite differences for gradient calculation.

To show that these standard optimization methods combined raycasting with the NCC and GCC metrics, they are termed "rcNCC" and "rcGCC".

3.3.3 Optimization Framework

The final requirement for implementation of XRF/TEE registration using raycasting was the optimization framework. Open-source optimization libraries (VNL) were used, from which two methods were chosen for experiments: the Nelder-Mead method and the Powell method. These methods were chosen because:

- They work well for low-dimensional optimization problems.
- They can perform optimization without needing analytical similarity function derivatives, which are unavailable in this case.

In [63], a tutorial on combining the VNL non-linear optimization libraries with CUDA was presented.

3.4 Novel Algorithms

In this section, two novel 2D/3D registration algorithms are introduced. The main design goal was to achieve large speed-ups in registration time while maintaining reasonable accuracy compared to the standard algorithms presented in section 3.3.

3.4.1 Direct Splat Correlation

Since 2D/3D registration is an iterative process, a key determinant of overall registration speed is the time needed to generate the DRR at each step of the optimization. For our application, optimization typically requires 150-300 similarity computations, depending on how close the initial pose is to the final solution. This means that both DRR generation and similarity computation need to be completed in roughly 300 μs on average for 15 fps registration.

In section 3.2.4, the concept of splatting for DRR generation was introduced. When using a CPU for computation, splatting can be more computationally efficient than raycasting, especially when the size of the point cloud is small compared to the size of the data volume it was derived from. However, splatting does not necessarily translate well to the GPU, because every projected point must perform an *indirect write* operation at a random pixel (see Fig. 3.8, left, line 3). However, it can be shown that the correlation similarity metric can be reformulated so that all of the write operations are replaced with extremely fast `texture` reads on the GPU.

Consider the correlation (CC) between an XRF image $I(u)$ and a DRR

$D(u)$:

$$CC = \sum_i I(u_i) D(u_i) \tag{3.17}$$

For a splat-generated DRR, we can substitute the $D(u)$ term to obtain:

$$CC = \sum_i I(u_i) \sum_{j \in S_i} -V(x_j) = -\sum_i \sum_{j \in S_i} I(P \cdot T \cdot \ x_j) V(x_j) \tag{3.18}$$

Finally, this simplifies to a sum over all 3D points:

$$CC = -\sum_j I(P \cdot T \cdot \ x_j) V(x_j) \tag{3.19}$$

This expression simply states that the correlation is equal to the sum, over all 3D points, of the point intensity times the value of the pixel that it projects onto. Equation 3.19 shows that the DRR does not need to be explicitly generated to compute the similarity, enabling more efficient computation (Fig. 3.8).

```
//Splat correlation w/
   DRRs
1  for each point x
2    u = P*T*x
3    DRR[u]=DRR[u]+V[x]
4  for each pixel u
5    cc+=DRR[u]*I[u]
```

```
//Direct method w/o DRRs
1  for each point x
2    u = P*T*x
3    cc+=I[u]*V[x]
4
5
```

Figure 3.8: Pseudo-code for computation of correlation for explicit DRR generation (left) and the proposed method (right).

Because this similarity is inspired by splatting and is directly computed

from the image, without the explicit generation of a DRR, we refer to it as "direct splat correlation" (DSC).

Point Cloud Representation

A point cloud representation of the probe (PC_{probe}) was required for the implementation of the DSC algorithm, and was created using the following procedure:

1. High-intensity features in the TEE probe CT (CT_{probe}) were manually segmented.

2. Segmented voxels were organized into 4×4×4 cubes, which was done to make sure that voxels located nearby each other in space were also located nearby each other in memory. This was done to improve the efficiency of the 2D `texture` read operations, which are faster when threads within the same warp access nearby spatial locations in an image.

3. 2^N points were randomly generated within the segmented region. The number of points affects the algorithm's run-time and accuracy, and so varying values of N (10-20) were tested.

4. Linear interpolation was used to assign an intensity value to each point.

3.4.2 Patch Gradient Correlation

The DSC method was designed to perform similarity computations without the need for generating DRRs, potentially resulting in fast registration speeds.

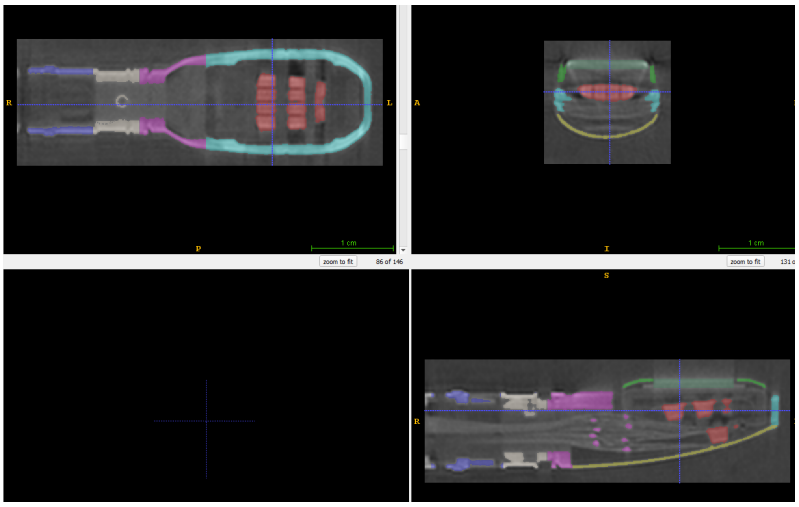

Figure 3.9: Manual segmentation of high-intensity voxels from the TEE probe CT.

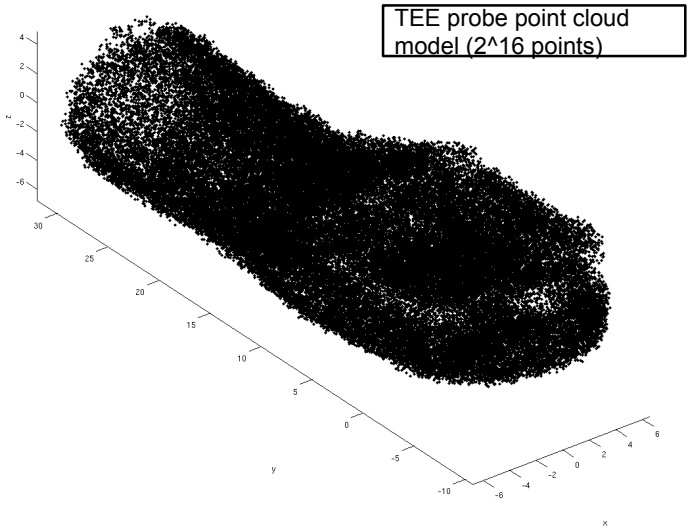

Figure 3.10: Point cloud model of the probe. The dimensions are in mm.

This was achieved by reformulating the cross-correlation similarity metric. However, other, more complicated similarity metrics have been introduced that, although much more computationally expensive, perform better in terms

of registration accuracy. Some examples include GCC and pattern intensity [64].

As an alternative to DSC, an algorithm was designed with a goal of achieving accuracy equal to or better than the GCC method, while achieving faster registration speeds. One of the key features of device pose estimation in XRF images is that the device typically has a high degree of *saliency* compared to other components of the image. This means that the device produces many high-contrast, high-entropy patches in the image that can be used as registration features. It was hypothesized that focusing on these salient features, rather than the entire image, would allow for accurate registration with a much lower computational burden. To that end, the new algorithm computed the GCC metric on a subset of patches rather than the whole image, and was therefore called "patch gradient correlation" (PGC).

In the PGC algorithm, a set of K key-points on the TEE probe are manually defined. The key-points are chosen based on their proximity to high contrast features on the probe (Fig. 3.11). During each optimization iteration, the key-points are projected on the image. At all K projected key-points, 16×16 DRR patches ($D_1 ... D_K$) are generating by raycasting. Patches are also extracted from the XRF image at same locations ($I_1 ... I_K$). The similarity metric is the sum of the gradient correlation over all patches:

$$F_{PGC} = \sum_{k}^{K} F_{GCC}(D_k, I_k) \tag{3.20}$$

The location of key-points and patch extraction process is illustrated in Fig. 3.11.

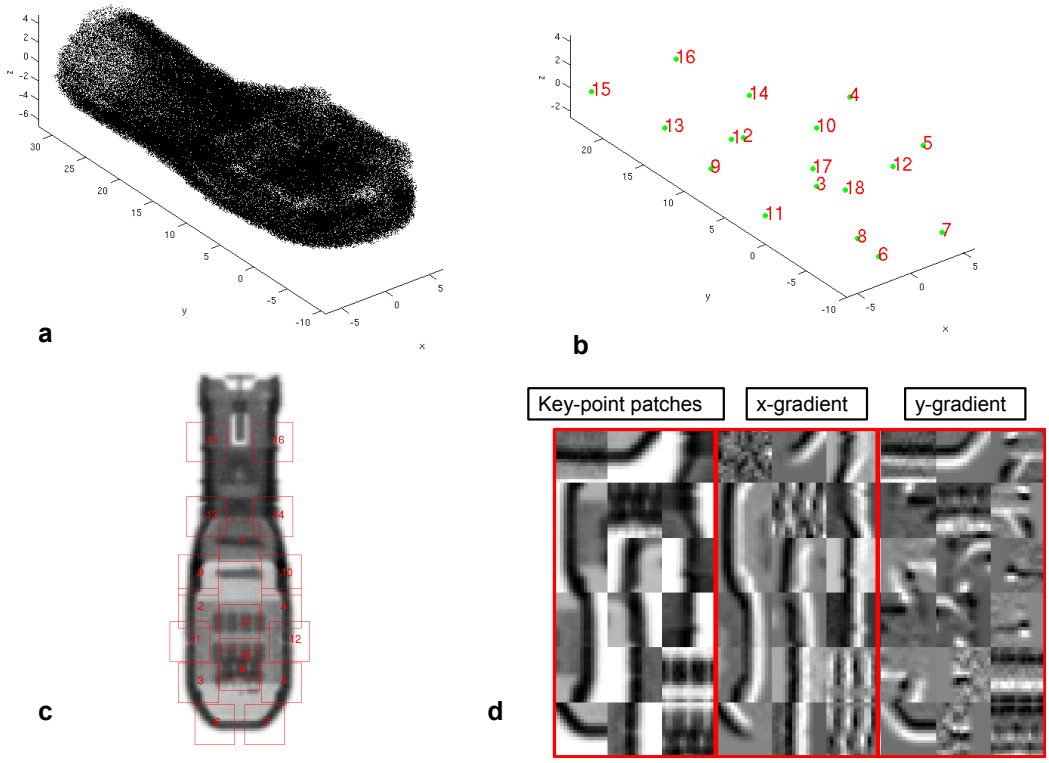

Figure 3.11: The PGC method. a) 3D point cloud model of the probe. b) Location of 18 key-points. c) An example of probe key-points projected onto an image of the probe, as well as the spatial footprint of the corresponding patches d) Patches extracted from the locations of the projected key-points, with x and y gradients

The following procedure is used to compute the PGC metric:

1. CPU:

   - Transform the key-points to the XRF image space using the spatial transform ($^{XRF}T_{CT}$, see Eq. 3.9) generated from the parameters $\phi = (t_x, t_y, t_z, \theta_x, \theta_y, \theta_z)$:

$$p^{(XRF)}_{key-points} = {}^{XRF}\,T_{CT}\, p^{(CT)}_{key-points} \tag{3.21}$$

- Project the key-points onto the XRF image.

$$p_{key-points}^{(image)} = P\,{}^{XRF}T_{CT}\,p_{key-points}^{(CT)} \tag{3.22}$$

- As specified in section 3.3.1, transform the CT bounding box coordinates to the XRF image space using ${}^{XRF}T_{CT}$ and record the minimum and maximum z-coordinates of CT bounding box after transformation to the XRF coordinate system (z_{start}, z_{stop}).
- Compute the inverse spatial transform ${}^{CT}T_{XRF} = {}^{XRF}\,T_{CT}^{-1}$
- Transfer all spatial transform and bounding box information to `CUDA constant` memory $({}^{CT}T_{XRF}, p_{key-points}^{(image)}, z_{start}, z_{stop})$.

2. GPU:

- Launch 1 block for every key-point, and launch 256 (16×16) threads per block.
- Using the same procedure from section 3.3.1, raycast through the CT volume to produce a 16×16 DRR patch D_k centered on a projected key-point.
- Using finite differences, compute the x-gradient and y-gradient of the DRR patch.
- Using a 2D `texture` read, extract a (16×16) patch I_k from the image centered on the projected key-point.
- Using finite differences, compute the x-gradient and y-gradient of the image patch.

- Compute the gradient correlation between the DRR patch and the image patch (Eq. 3.16).
- Transfer the gradient correlation for each patch back to the CPU.

3. Sum all the gradient correlations into a single scalar value. Report to optimizer.

3.4.3 Hybrid Method

In the previous sections, two novel 2D/3D registration algorithms were presented: DSC and PGC. It was hypothesized that the DSC method would be faster, while the PGC method would be more accurate. With this in mind, a hybrid (HYB) method was designed that combined both methods. In the hybrid method, DSC was first used to arrive at an initial solution, while PGC was used to refine the solution. The idea was that this strategy would decrease the total number of iterations needed for the PGC method to converge by using the DSC method to provide a fast and accurate initialization. It was hypothesized that the hybrid method could achieve the same accuracy as PGC, but with a faster registration speed and higher capture range.

3.5 Discussion and Conclusion

In this chapter, the technical principles and prior art concerning XRF/TEE registration were presented. A modified implementation of TEE pose estimation via standard 2D/3D registration methods (rcNCC,rcGCC), inspired by the work in [30], was then described. Finally, two novel algorithms aimed at

faster and more accurate registration were introduced (DSC, PGC, and HYB), as well as details about implementation.

In the next chapter, the DSC, PGC and HYB methods are validated against the standard XRF/TEE registration methods. The results of experiments in simulated, phantom, and clinical datasets are presented.

4. XRF/TEE Registration: Experimental Validation

In this chapter, experiments are presented characterizing the performance of the standard (rcNCC,rcGCC) and novel (DSC,PGC,HYB) XRF/TEE registration methods presented in the previous chapter. Prior to presenting experimental results, cost function curves were generated for each method in order to qualitatively characterize the capture range and accuracy. Next, three different experimental scenarios were tested: simulated, phantom, and clinical datasets.

4.1 Common Materials and Methods

In the following sections, experimental materials and methods common to all experiments are described.

4.1.1 Computer Hardware and Software

A Philips X7-2t TEE probe was used in this study. All experiments were run on a Dell Precision T7500 work station running Ubuntu Linux with a 3.47 GHz Intel Xeon processor and a NVIDIA Tesla K20 GPU. `VNL` libraries were used for optimization. Retrospective clinical dataset processing was approved by the local institutional review board.

Portions of the work from this chapter are accepted for publication as: Hatt, Charles R., Speidel, Michael A., and Raval, Amish N."Robust 5DOF Transesophageal Echo Probe Tracking at Fluoroscopic Frame Rates" In Medical Image Computing and Computer–Assisted Intervention-MICCAI 2015. Springer International Publishing, 2015.

4.1.2 Optimization Strategy

All experiments assumed a single x-ray projection, which typically results in inaccurate estimates of t_z. Therefore, we focused on optimization of the other five parameters, which resulted in a smaller optimization problem. Optimization consisted of two stages, first dealing with in-plane parameters (t_x,t_y,θ_z), and then all of the parameters except t_z (t_x,t_y,θ_x,θ_y,θ_z, Fig. 4.1). For the hybrid method, optimization consisted of three stages, shown in Fig. 4.1.

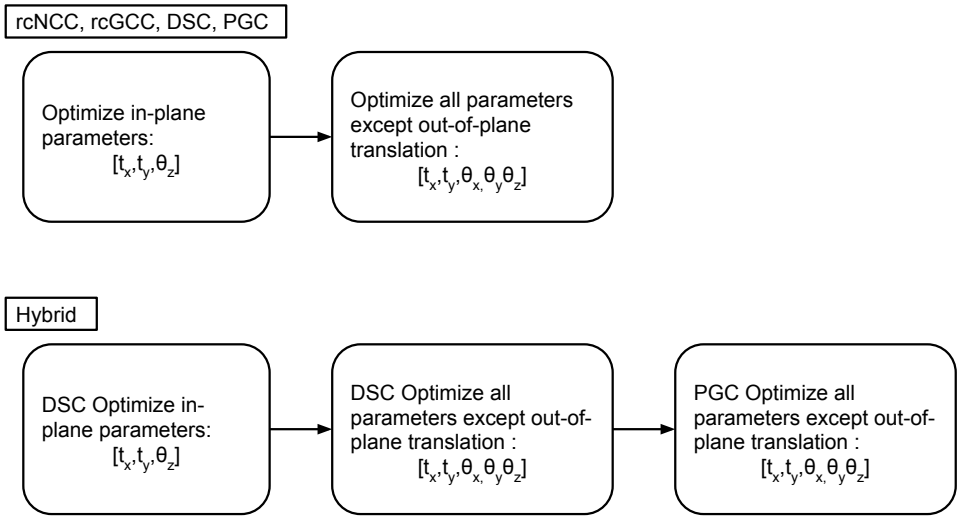

Figure 4.1: Optimization strategy for the simulation, phantom, and clinical experiments.

4.2 Cost function curves

In order to characterize the topology of each method, cost function curves were generated for a few representative TEE probe orientations. These curves were

generated by setting all of the registration parameters except one (ϕ_α) to the ground truth, varying ϕ_α over a range of values, and plotting the cost function at each value. The purpose was to investigate how the cost function behaves for different initial position and orientation errors, and how accurately the correct pose can be recovered. In particular, these curves can be used to examine how convex the cost functions are, which in turn can be used to estimate the capture range. Furthermore, the position of the cost function minimum relative to the origin indicates how accurately the parameter in question can be recovered. An example plot is shown in Fig. 4.2. These plots were generated from simulated images, which are described in section 4.3. Plots for different probe orientations are included in Appendix C.

By examining the cost function plots, a few interesting properties of the algorithms can be seen:

- The rcNCC and DSC methods have very similar topologies, with a wide, convex curve indicating a large capture range for in-plane position and all rotations.

- DSC consistently performs the worst for the θ_x parameter (TEE probe pitch), although it does converge to within less than 2.0° in all plots.

- DSC doesn't accurately estimate t_z (translation along the source-detector axis) indicating that it is only accurate for 5DOF 2D/3D registration.

- The rcGCC and PGC methods have similar topology, with a small translational capture range.. However, they are both more accurate than rcNCC and DSC for all parameters.

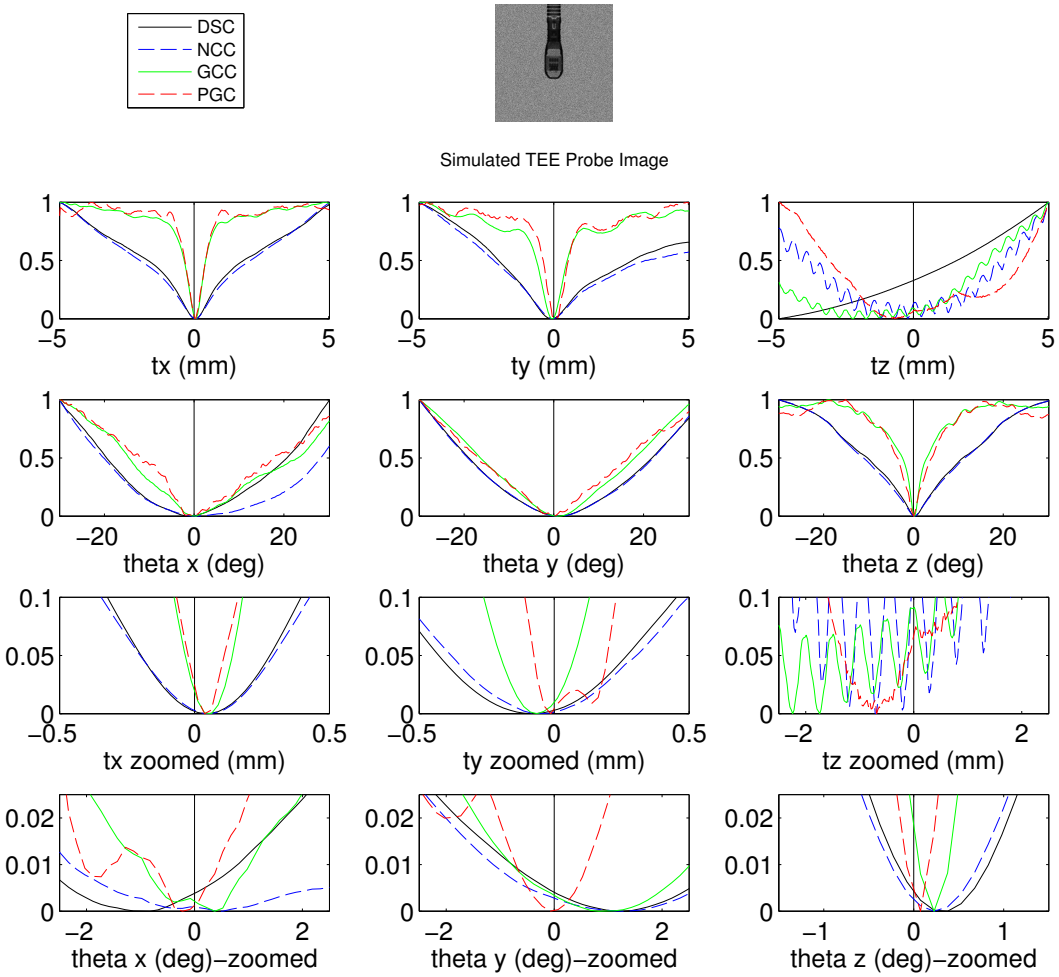

Figure 4.2: Cost function curves for a simulated image ($\theta_x = 0°$, $\theta_y = 0°$)

- All algorithms perform very well for estimation of the θ_z parameter (in-plane rotation).

4.3 Simulations

4.3.1 Experimental Design

The simulation experiments tested the accuracy and speed of the DSC, PGC, rcNCC, rcGCC, and HYB methods *in silico*. Simulation images (I_{sim}) were a hybrid of real background anatomy (I_{xrf}) and synthetic DRRs (I_{drr}).

A DRR of the probe was rendered using the splatting method with a point cloud large enough to generate a high quality DRR (2^{21} points). The background anatomy was obtained using images from TAVR procedures and the hybrid image was formulated as:

$$I_{sim} = I_{xrf} \cdot e^{-\alpha I_{drr}} \tag{4.1}$$

The parameter α controlled the probe to background contrast and was randomly varied to generate a contrast ratio ranging from 0.45 to 0.85 for each experiment. Fig. 4.3 shows a few examples of the simulated images.

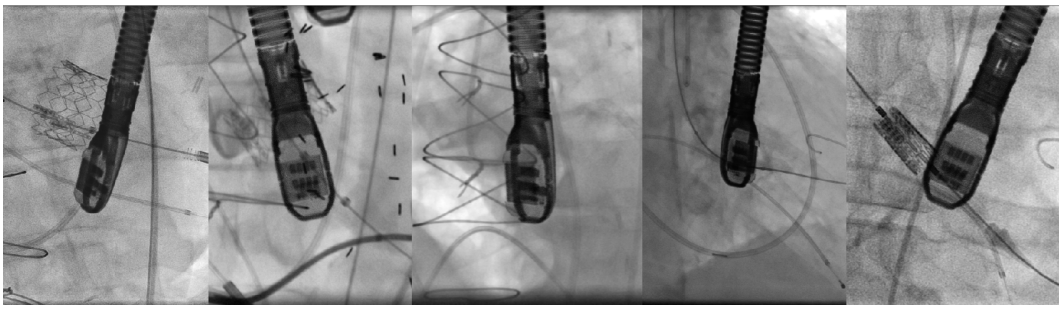

Figure 4.3: Examples of simulated images.

For each experiment, the TEE probe was placed at a random location and

orientation within the XRF C-arm image space. Based on our observations from an image database of TAVR cases, the TEE probe rarely has Euler angle rotations outside the range of -75° to 75° about its primary axis (y-axis), -30° to 30° about the x-axis of the image detector, and -45° to 45° about the source-detector axis (z-axis). Therefore, the initial pose was randomly generated from a uniform distribution within those ranges. Once an initial pose was created, a random mis-registration was applied, which the experiments attempted to recover. The random mis-registration was chosen from a zero mean normal distribution with standard deviations of 1.5 mm, 1.5 mm, 2.5 mm, 10°, 10°, 5° for the parameters $t_x, t_y, t_z, \theta_x, \theta_y, \theta_z$, respectively.

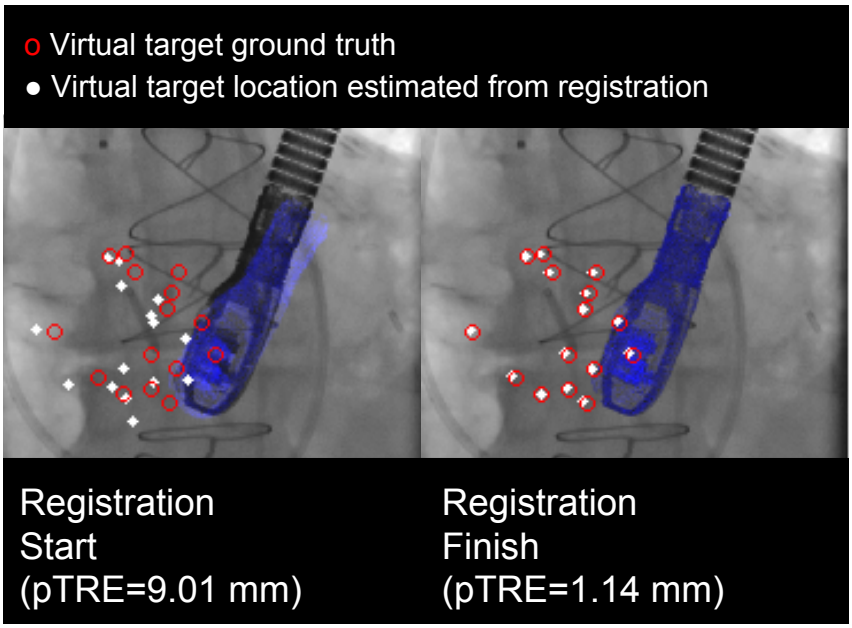

Figure 4.4: Virtual targets used to compute pTRE before and after registration.

When projecting echo data onto the XRF image, the errors in the estimation of t_z have little effect on target projection errors (for the same reason that

t_z is difficult to estimate in the first place). Therefore, the accuracy metric used was projection target registration error (pTRE), which was the root-mean-square error between known target points in XRF and estimated target points from echo following registration and projection to the XRF image:

$$\text{pTRE} = \sqrt{\frac{1}{N}\sum_{n}\left(\left\|\frac{1}{m_n}(p_n^{(xrf)} - P \cdot T \cdot p_n^{(echo)})\right\|_2\right)^2} \tag{4.2}$$

where m_n is the projective magnification of point $T \cdot p_n^{(echo)}$.

pTRE, % of successful registrations, and frame rate were reported for each experiment. We chose to define a successful registration as a pTRE $< 5.0\ mm$ based on results from [31] where pTRE of 2.9 mm was the mean error, but in reality this measure is application dependent. pTRE is only computed for successful registrations, to avoid large registration errors skewing the statistics.

For all experiments, virtual target points were used to compute pTRE. The virtual target points were randomly generated from within the center of a virtual ultrasound volume emanating from the TEE probe, at a mean distance of 50 mm from the probe face (Fig. 4.4). Errors from the probe model to echo volume calibration (${}^{echo}T_{probe}$) were not considered in the analysis.

The first set of experiments were conducted to examine the effect of point cloud size on the accuracy, speed, and success rate of the DSC algorithm. This was done by generating point cloud models of the TEE probe (see section 3.4.1) with varying numbers of points (2^{10}-2^{20}). The Nelder-Mead optimizer was used, and a total of 1000 simulated experiments were performed.

For the second set of experiments, all algorithms (DSC, PGC, rcNCC,

rcGCC, HYB) were tested with the Nelder-Mead and Powell methods. A point cloud size of 2^{16} was used for the DSC method. 5000 simulated experiments were performed.

4.3.2 Results

DSC Timing and Accuracy

For the experiments examining DSC accuracy, speed, and success rate, results are shown in Fig. 4.6. A surprising aspect of these results was that mean registration errors and success rates for small point clouds were comparable to those from large point clouds. Statistical tests were therefore performed to determine if the error distributions showed meaningful differences with respect to pTRE. A paired t-test was used to determine the p-values between each point cloud size, which are shown in Table 4.1, where the null hypothesis was that different point cloud sizes produced the same values of pTRE. Finally, median pTRE and the pTRE value corresponding to the cumulative distribution function (CDF) equaling 90% were calculated and reported in Table 4.2.

Based on these statistics, it was determined that a point cloud size of 2^{16} provided the best trade-off between accuracy and registration speed, as increasing the number of points did not increase the median pTRE. Furthermore, there was no statistically significant difference between the 2^{16} point cloud and larger point clouds in terms of pTRE. In general, increasing the point cloud size by a magnitude of 4 roughly doubled the mean registration time, and 2^{16}

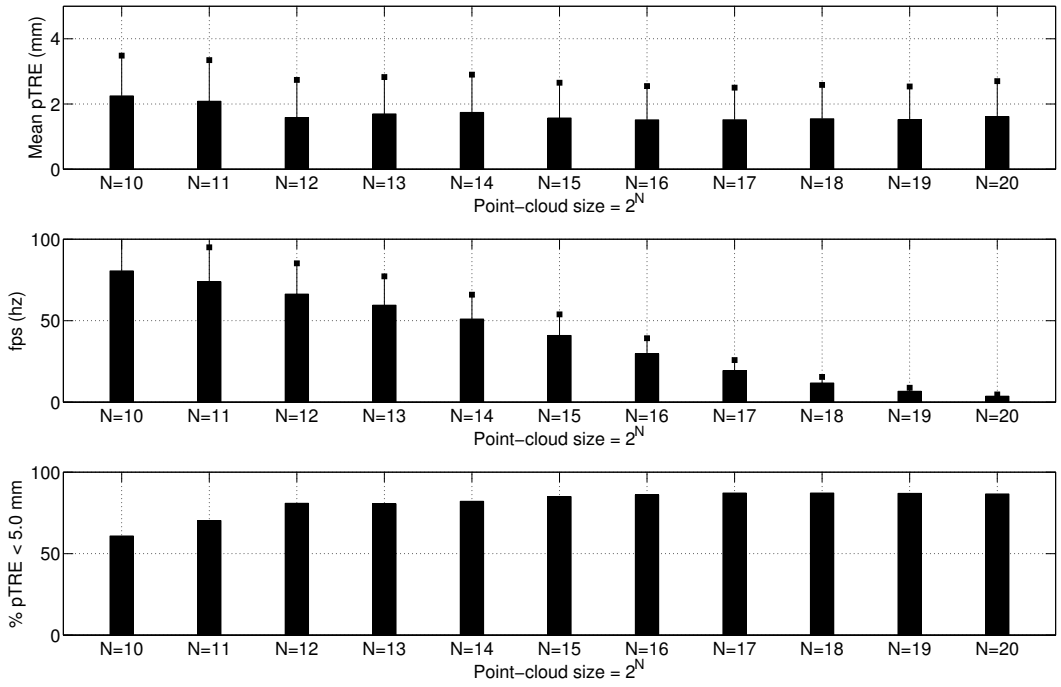

Figure 4.5: Timing and accuracy of different sized point clouds used for the DSC method.

points resulted in registration times well above fluoroscopic frame rates.

Algorithm Comparison

Results for the algorithm comparison studies are shown in Fig. 4.6. In addition, pTRE histograms are shown in Fig. 4.7 and Fig. 4.8. The following observations can be drawn from these results:

- The DSC method is an order of magnitude faster than the rcGCC and rcNCC methods, averaging 30±10 fps. The PGC and HYB methods are faster as well, both averaging faster than 10 fps registration speeds.

- Mean registration accuracy was different for the Nelder-Mead and Powell

Table 4.1: p-values between pTRE distributions for different point cloud sizes. A value of 0 indicates a p-value of less than 0.005. $p > 0.05$ in **bold**. The p-value represents the probability that differences in error distributions are due to chance.

Point cloud size	2^{10}	2^{11}	2^{12}	2^{13}	2^{14}	2^{15}	2^{16}	2^{17}	2^{18}	2^{19}	2^{20}
2^{10}	-	0	0	0	0	0	0	0	0	0	0
2^{11}		-	0	0	0	0	0	0	0	0	0
2^{12}			-	**.542**	**.634**	**.015**	0	0	0	0	0
2^{13}				-	**.940**	.011	0	0	0	0	0
2^{14}					-	.006	0	0	0	0	0
2^{15}						-	.006	.022	**.078**	.023	**.069**
2^{16}							-	**.997**	**.461**	**.729**	**.354**
2^{17}								-	**.477**	**.756**	**.412**
2^{18}									-	**.670**	**.874**
2^{19}										-	**.579**
2^{20}											-

methods. For the Nelder-Mead method, pTRE was similar for all methods except HYB, which had the lowest errors. However, for the Powell method, mean pTRE is highest for DSC, similar for PGC and rcNCC, 2nd lowest for rcGCC, and finally lowest for HYB.

- DSC and HYB clearly outperformed the other methods in terms of registration success rate, which is a proxy measure for capture range. PGC appeared to have the smallest capture range.

- The distributions in Fig. 4.7 and Fig. 4.8 show that the error modes were more distinct than the error means for the DSC method. For the PGC, rcNCC, rcGCC, and HYB methods, the error modes were all less than 0.5 mm, while for the DSC method, the error mode was slightly

Table 4.2: Median pTRE and threshold T such that the CDF = 90%.

	Median pTRE (mm)	T | CDF=0.9
10	3.66	12.9
11	2.67	10.6
12	1.52	8.38
13	1.75	8.10
14	1.74	7.68
15	1.45	7.41
16	1.38	6.64
17	1.43	6.28
18	1.39	6.01
19	1.40	6.49
20	1.50	6.69

less than 1 mm. However, the DSC method failed less than the other methods.

- The HYB method outperformed all methods in terms of success rate and registration accuracy, and was also the second fastest method. While the DSC method greatly outperformed even the HYB method in terms of registration speed, the HYB method was almost fast enough to operate at fluoroscopic frame rates.

- All methods resulted in mean registration errors lower than the results reported in [31] for clinical validation. However, errors resulting from echo to probe calibration were not considered in this study.

- Differences in speed and accuracy between the Nelder-Mead and Powell optimizers were most likely related to optimization settings controlling the magnitude of changes in the pose parameters at each iteration, which generally resulted in a larger number of function evaluations for the

Powell optimizer. Due to fundamental differences in the optimizers, it was not possible to keep these settings equal between experiments.

- It is also important to note that the registration speeds obtained for the rcNCC and rcGCC method were faster than reported in the literature [30, 31], indicating that the differences in registration speeds were not artificially inflated.

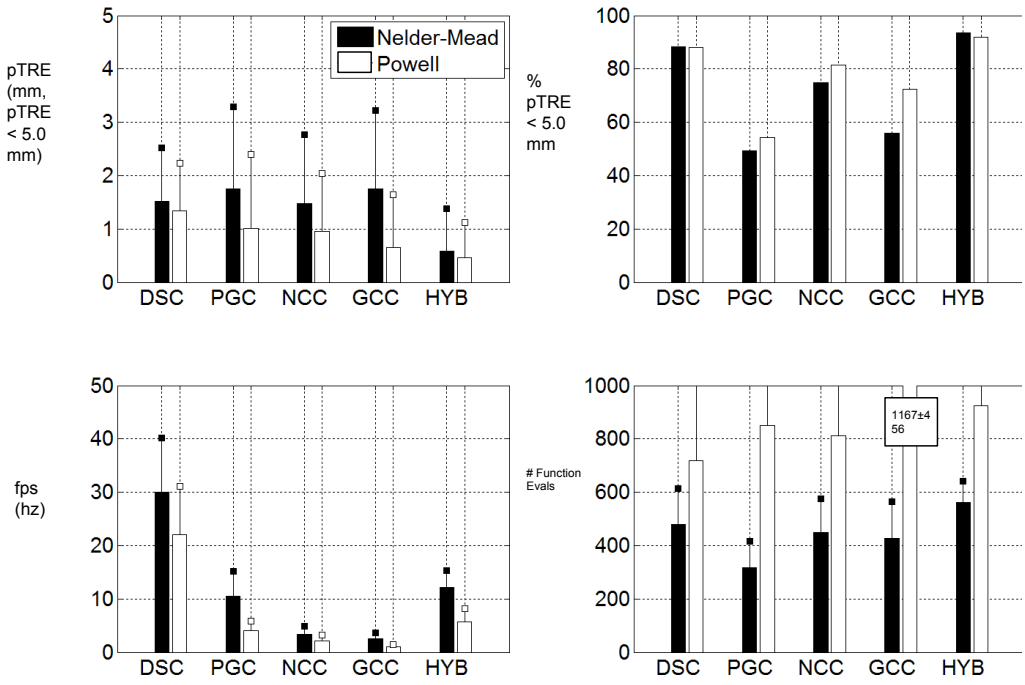

Figure 4.6: Summary statistics of pTRE, success rate, frame-rate, and number of function evaluations for each method in simulated datasets (mean ± std).

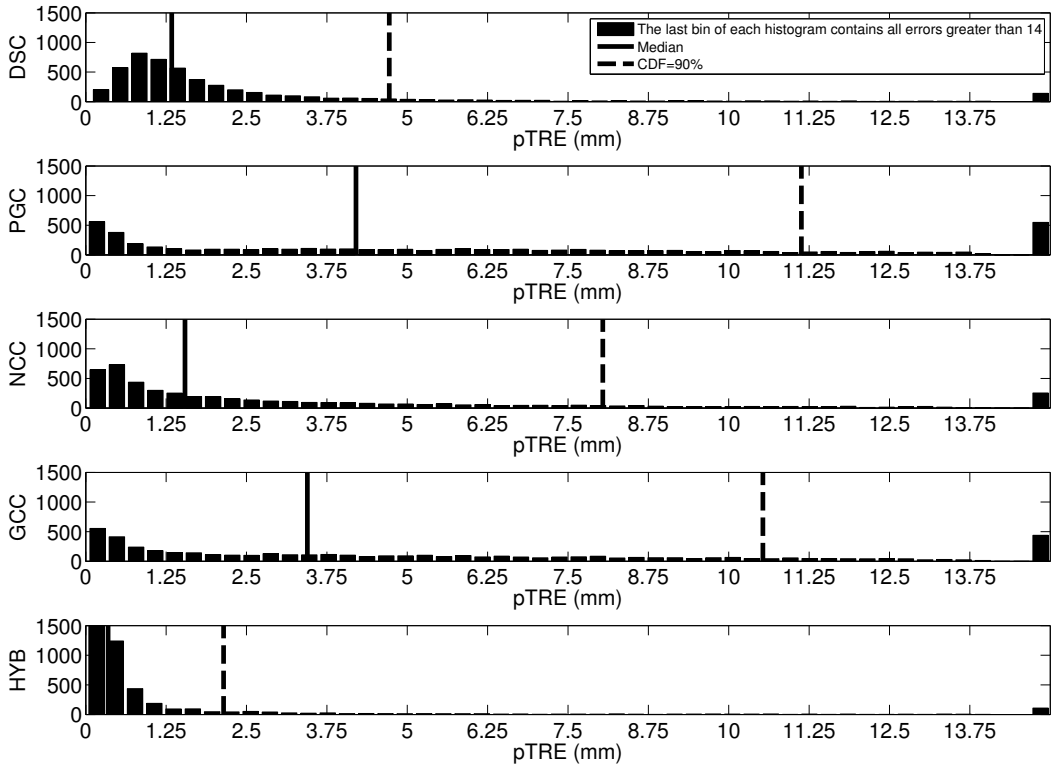

Figure 4.7: pTRE histograms for the simulated datasets (Nelder-Mead optimizer).

4.4 Phantom Experiments

4.4.1 Experimental Design

The purpose of this experiment was to measure 5DOF registration accuracy against ground truth, which was determined using a set of stainless steel markers attached to the cylinder containing the probe (Fig. 4.9). This was done to establish a rigid spatial relationship between the TEE probe and the fiducial markers. Because the markers were discrete, high contrast objects that are spread out over a larger area than the TEE probe, the marker based regis-

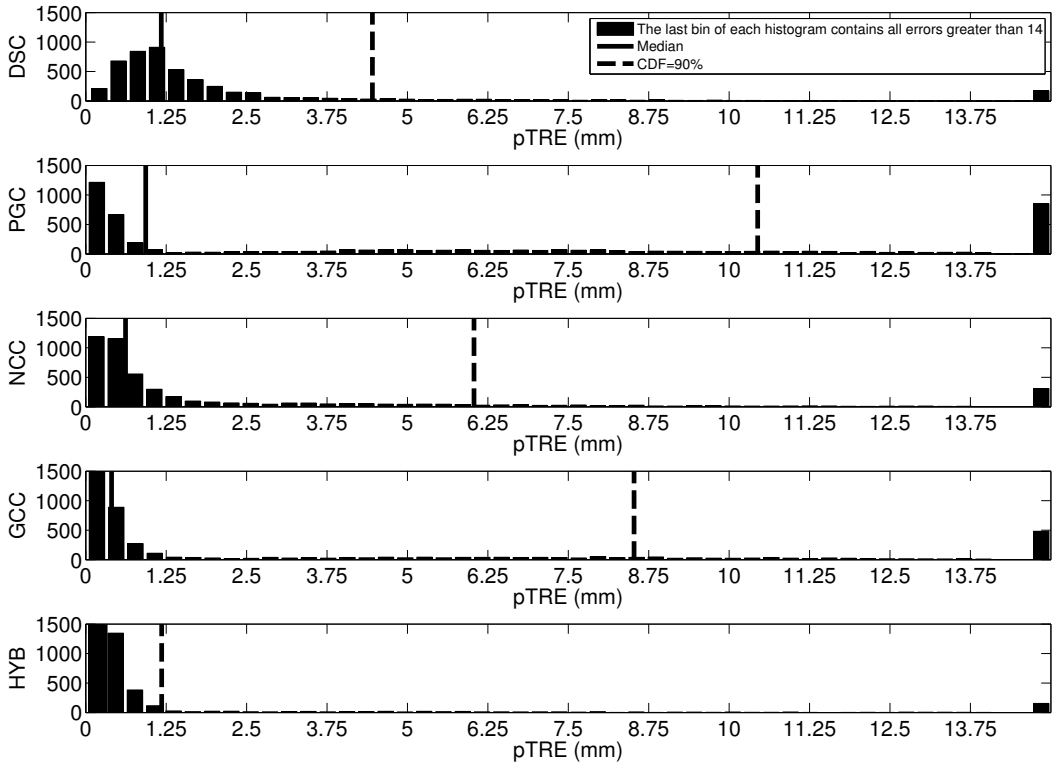

Figure 4.8: pTRE histograms for the simulated datasets (Powell optimizer).

tration accuracy was assumed to be accurate enough to be used as a ground truth measure (based on prior studies, see [65] and [66].

Following experimental setup, a CBCT ($CT_{phantom}$) scan of the probe and affixed cylinder was acquired. The XRF projections used to create the CBCT image were also recorded and used for the analysis.

Data Processing

In order to use the stainless steel fiducials as a ground truth reference, it was necessary to register the fiducials to the TEE probe. The was done by segmenting the fiducials from the $CT_{phantom}$ volume, and registering the TEE

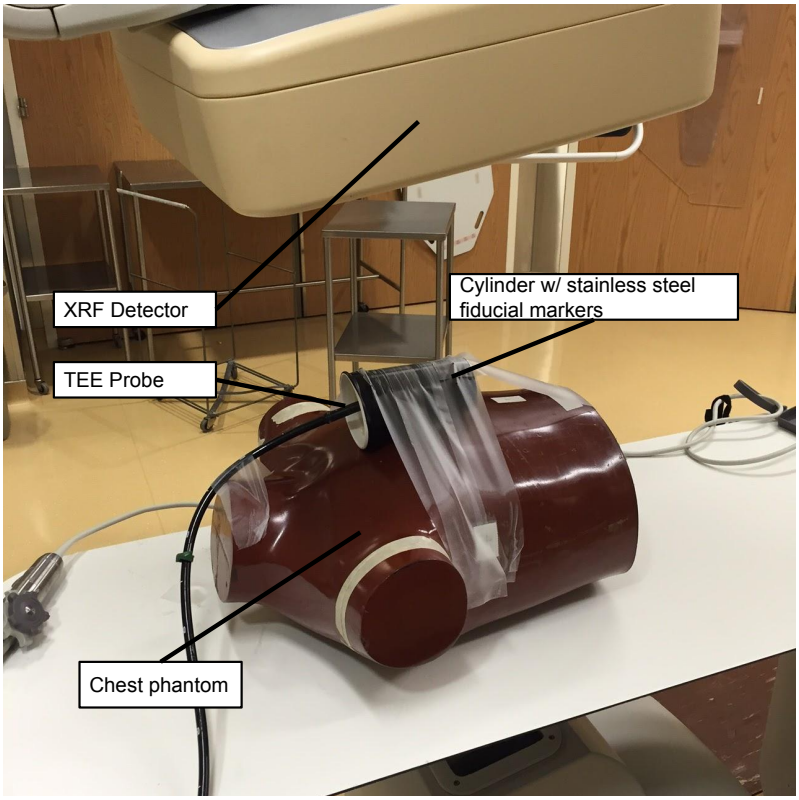

Figure 4.9: Setup for the phantom experiment.

probe point cloud (PC_{probe}) to the $CT_{phantom}$.

The fiducials (S_{CT}) were segmented from the $CT_{phantom}$ by manually setting an intensity threshold that segmented each fiducial into a single connected component. For each of the N connected components, a segmentation mask M_n was created, and the centroid of the fiducial was found using a center of mass calculation.

$$x_{n,com} = A \sum_{z=1}^{H} \sum_{y=1}^{L} \sum_{x=1}^{W} x \cdot CT(x,y,z) \cdot M_n(x,y,z) \tag{4.3}$$

$$y_{n,com} = A \sum_{z=1}^{H} \sum_{y=1}^{L} \sum_{x=1}^{W} y \cdot CT(x,y,z) \cdot M_n(x,y,z) \tag{4.4}$$

$$z_{n,com} = A \sum_{z=1}^{H} \sum_{y=1}^{L} \sum_{x=1}^{W} z \cdot CT(x,y,z) \cdot M_n(x,y,z) \tag{4.5}$$

$$A = \frac{1}{\sum_{z=1}^{H} \sum_{y=1}^{L} \sum_{x=1}^{W} CT(x,y,z) \cdot M_n(x,y,z)} \tag{4.6}$$

In order to register PC_{probe} to the $CT_{phantom}$, the following correlation cost function was used:

$$F_{CC} = \sum_{n}^{N} CT_{phantom}({}^{CT}T_{probe} \cdot [x_n, y_n, z_n, 1]^{\top}) \cdot V_n \tag{4.7}$$

where $[x_n, y_n, z_n]$ was the n^{th} point in PC_{probe} and V_n was the image intensity associated with that point. Following initialization, the Nelder-Mead optimizer was used for registration, providing ${}^{CT}T_{probe}$. The fiducials were then transformed to the TEE probe space:

$$S_{probe} = ({}^{CT}T_{probe})^{-1} \cdot S_{CT} \tag{4.8}$$

For pose-estimation during the experiment, the ground truth was established by finding the pose parameters that aligned the fiducials with their projected locations in each XRF image frame (I_t). For this optimization, the following cost function was used:

$$F_{fiducials} = \sum_{n}^{N} I_{xrf}(P \cdot^{xrf} T_{probe}^{(fiducials)} \cdot S_{probe}) \tag{4.9}$$

This was also solved using the Nelder-Mead optimizer following manual registration for each frame, providing a ground truth measure for ${}^{xrf}T_{probe}^{(fiducials)}$.

Alternatively, ${}^{xrf}T_{probe}^{(tee)}$ was estimated using the rcNCC, rcGCC, DSC, PGC, and HYB methods. The pTRE was computed using the following equation:

$$\text{pTRE} = \sqrt{\frac{1}{N}\sum_{n}\left(\left\|\frac{1}{m_n}(P \cdot^{xrf} T_{probe}^{(fiducials)} \cdot S_n) - (P \cdot^{xrf} T_{probe}^{(tee)} \cdot S_n)\right\|_2\right)^2} \tag{4.10}$$

Where m_n is the projective magnification of point ${}^{xrf}T_{probe}^{(fiducials)} \cdot S_n$.

The experiment consisted of tracking the fiducials and TEE probe throughout the rotational scan, and comparing the pose estimation accuracy. For TEE tracking, the first frame was initialized to the ground truth result, and then the pose was estimated using 2D/3D registration. For the following frame, the previous pose estimation result was used for initialization. In this way, the experiment tested the ability of the algorithms to accurately track the pose of the TEE probe during motion. Although in reality the XRF C-arm was the object moving, we analyzed the results as if it were the probe undergoing motion relative to a static C-arm. Fig. 4.10 illustrates the probe motion and example XRF image frames throughout the sequence.

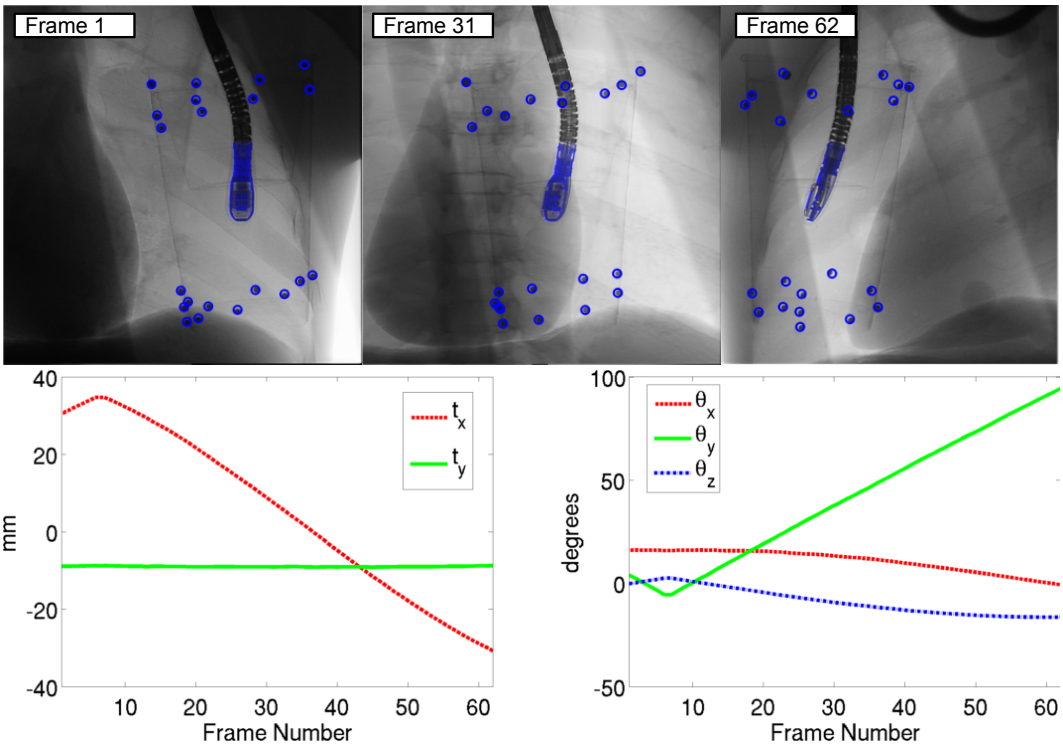

Figure 4.10: Phantom experiment image sequence, showing the probe DRR overlaid on the image in blue, as well as the fiducial marker locations as estimated by tracking the probe ($P \cdot^{xrf} T_{probe}^{(tee)} \cdot S_n$).

4.4.2 Results

The pTRE for each frame in the image sequence is shown in Fig. 4.11 for the Nelder-Mead optimizer and Fig. 4.12 for the Powell optimizer. Summary statistics are shown in Fig. 4.13. The phantom experimental results reinforce the conclusions from the simulation experiments in the following ways:

- Again, DSC is less accurate than the other methods (in terms of the error mode) but an order of magnitude faster.
- The hybrid method is the second fastest method, but is also very ac-

curate. In this case, the rcGCC method was marginally more accurate than the hybrid method, but the hybrid method appeared to provide the greatest trade-off between speed and accuracy, especially for the Nelder-Mead optimizer.

For the Powell optimizer, neither the DSC, rcNCC or PGC method were robust enough to track the TEE probe throughout the entire image sequence. The PGC method failed beyond recovery, rcNCC failed for a large number of frames, and DSC failed for one frame. This reinforces the conclusion that, although the PGC method is fast and accurate, it has a small capture range and may not be able to track large inter-frame pose differences. On the other hand, the Nelder-Mead optimizer resulted in zero failed frames for every method, including the PGC method, at slightly faster registration frame rates (except for DSC, which was faster using the Powell optimizer).

4.5 Clinical Datasets

4.5.1 Experimental Design

In a final set of experiments, validation was performed on images from TAVR procedures. 91 image sequences from 81 cases (7994 frames) were identified as containing significant probe movement due to physician manipulation or cardiorespiratory motion. Due to a lack of ground truth, a surrogate ground truth was created. It was assumed that the gold standard method used in prior work, rcGCC [30], was the best alternative in the absence of a true ground

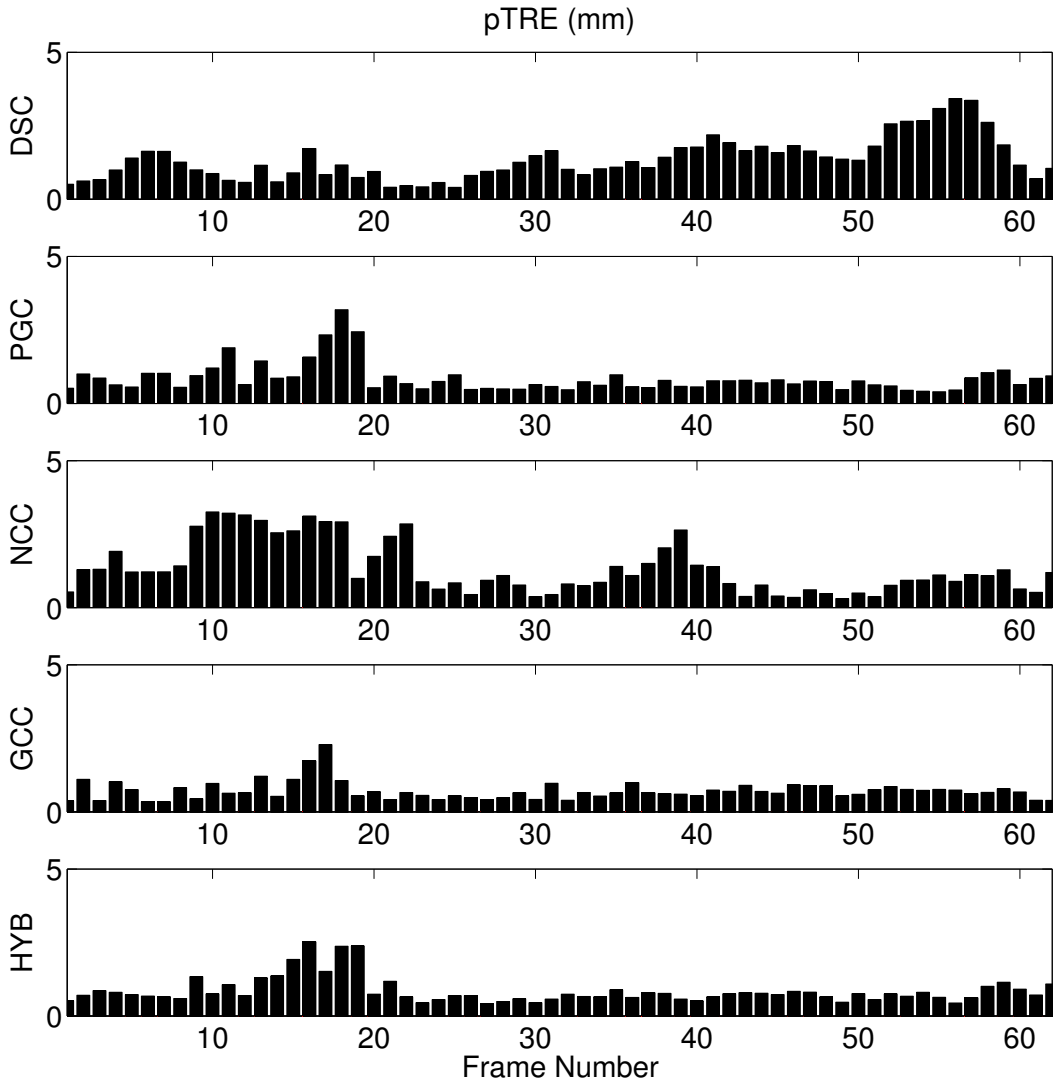

Figure 4.11: Phantom experiment, frame-by-frame pTRE for each method using the Nelder-Mead optimizer.

truth, and the following procedure was used to process all 7994 image frames:

1. Manual registration was performed for the first frame of each sequence.

2. The initial manual registration was refined using the rcGCC method with the Powell optimizer (rcGCC-pwl).

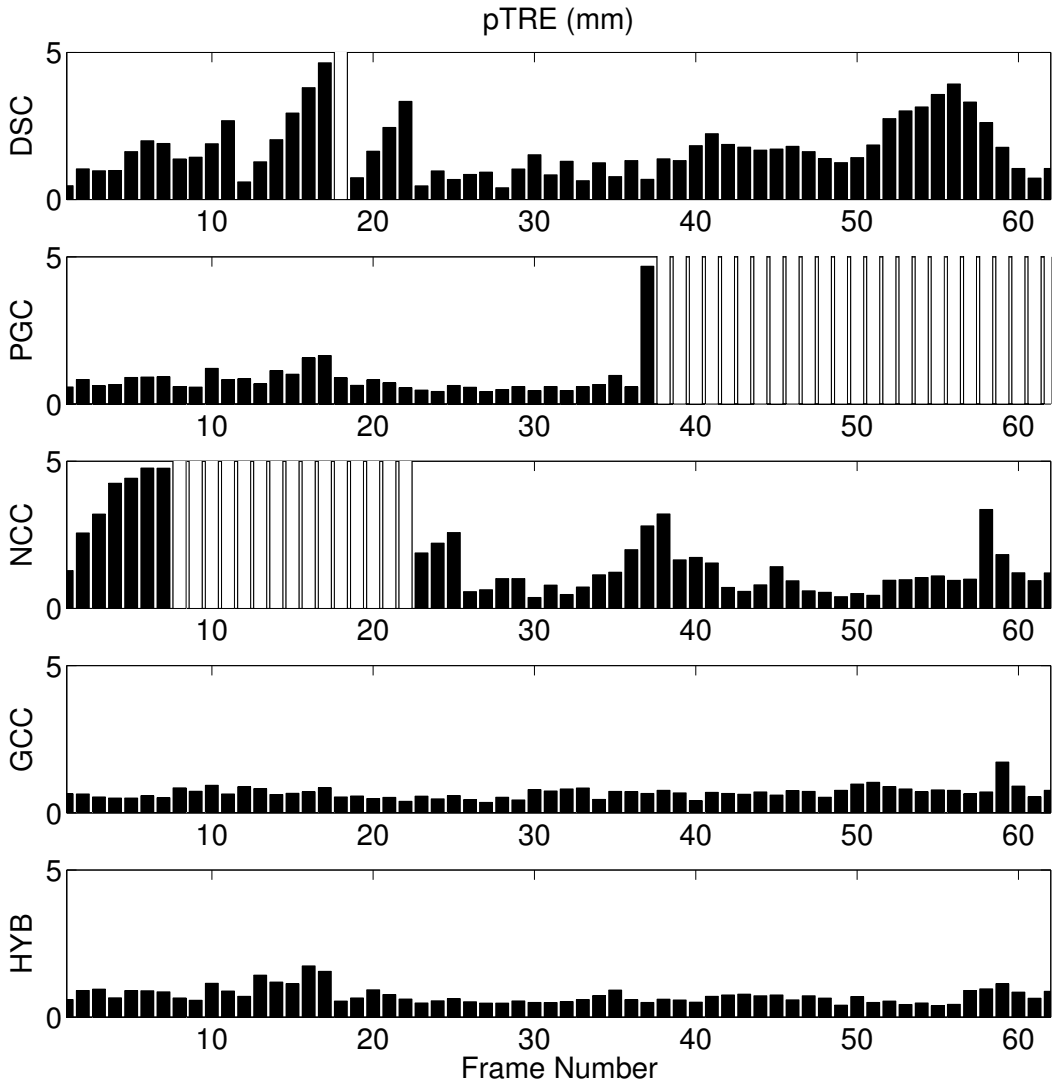

Figure 4.12: Phantom experiment, frame-by-frame pTRE for each method using the Powell optimizer. White bars indicate errors over 5.0 mm.

3. rcGCC-pwl was used to define the ground truth registration at each consecutive frame. Each sequence was then visually checked, frame-by-frame, for errors. The probe could not be tracked in 2 sequences and they were removed from the analysis.

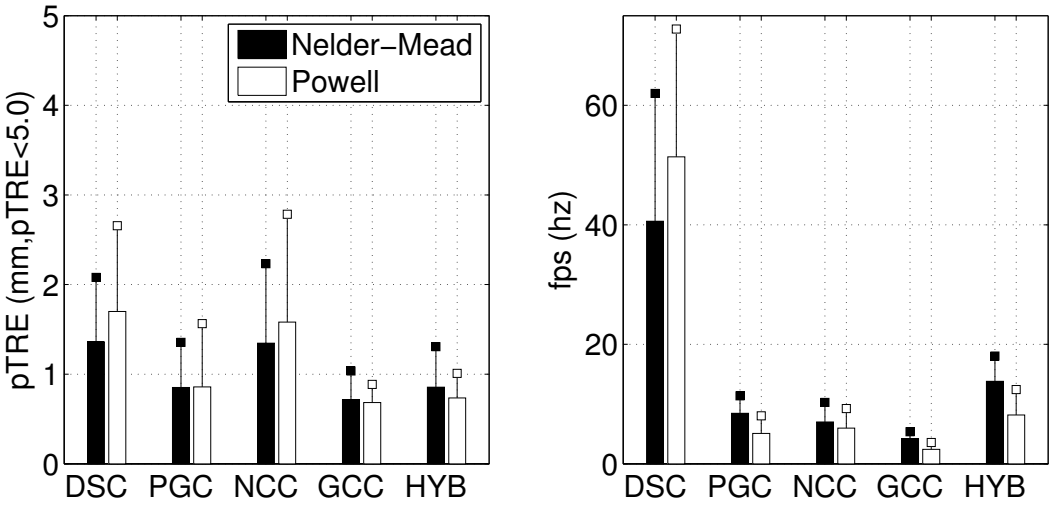

Figure 4.13: Summary statistics for the phantom experiment (mean ± std).

For the clinical datasets, we examined the tracking accuracy and robustness under real clinical conditions where both image streaming and registration delays must be considered. Assuming 15 fps image streaming, the registration lag was calculated as $n_{lag} = ceil(15 \times t_{registration})$ frames. Each skipped frame was only allowed to use the most recently finished registration result for its pTRE calculation, and every registration (T_n) was initialized with the most recently processed registration result ($T_{n-n_{lag}}$). Note that slower registration times resulted in increased pTRE due to not only more skipped frames but also less accurate frame-to-frame initialization.

4.5.2 Results

Results from the clinical datasets are summarized in Fig. 4.14. In this case, DSC and HYB outperformed all other methods in terms of registration accuracy, with HYB having the lowest errors for both optimizers. In terms of

pTRE, success rate, and registration speed, the Nelder-Mead optimizer outperformed the Powell method, with the HYB method achieving a nearly 100 % rate of success, the lowest registration errors, and mean pTRE less than 1.5 mm.

In experiments that did not consider x-ray video and image streaming lags, the DSC method was the least accurate and the HYB method showed comparable accuracy to rcGCC. The fact that both the DSC and HYB methods showed the lowest errors in the clinical experiments, where image streaming was considered, indicated that target registration error is reduced when tracking at imaging frame rates, justifying the need for a real-time algorithm.

4.6 Discussion

In this chapter, the DSC, PGC and HYB methods presented in chapter 3 were validated against the standard rcNCC and rcGCC methods in simulated, phantom, and clinical datasets. The main conclusions from these experiments were the following:

- The DSC method was much faster than all of the other methods, and was capable of achieving registration frame rates that greatly exceeded clinical requirements ($> 15\ fps$). The DSC method was also more robust than the other methods in the sense that it converged to less than 5.0 mm pTRE more often.

- The PGC method was similar in terms of accuracy, slightly less robust, and 2-4 times faster than rcGCC method.

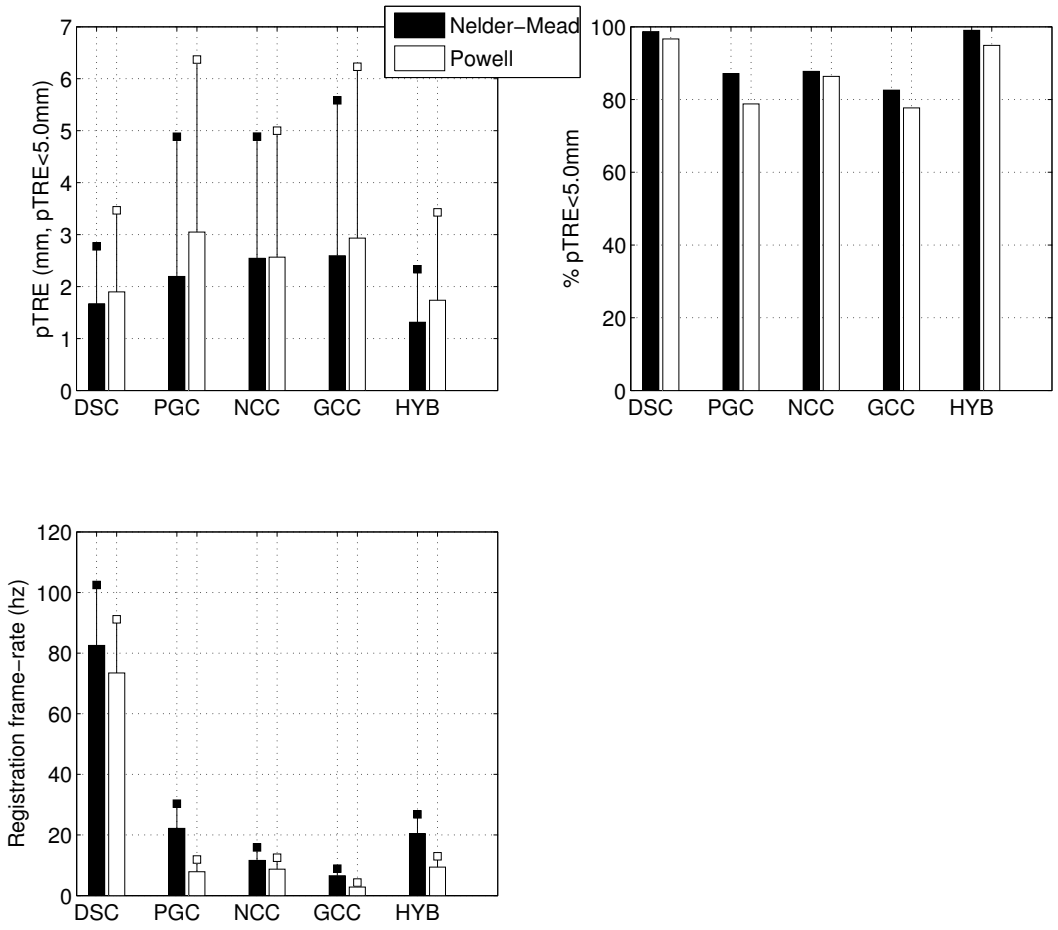

Figure 4.14: Summary statistics for clinical datasets (mean ± std).

- Results from all experiments indicated that the HYB method was able to successfully combine the speed and high capture range of DSC with the accuracy of PGC, resulting in robust and accurate 5DOF registration at or near fluoroscopic frame rates.

By having a method for performing real-time, accurate, and robust XRF/-TEE registration, new clinical workflows for TAVR are now possible. Examples of this are image guidance paradigms envisioned in the introductory chapter

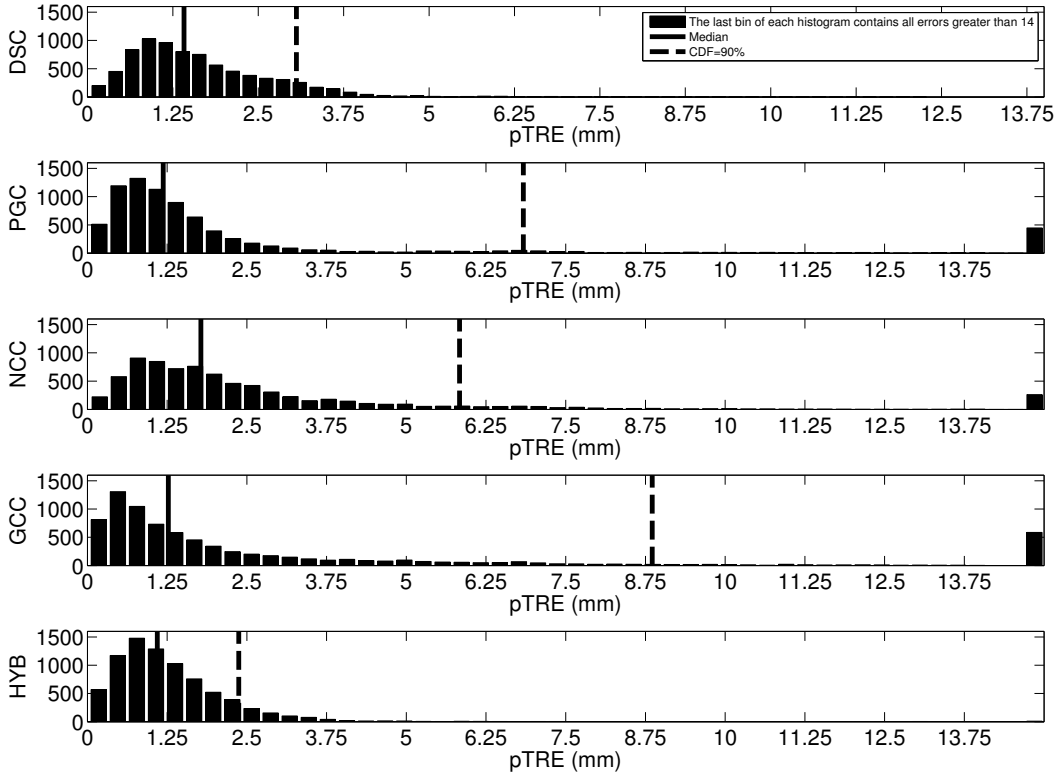

Figure 4.15: pTRE histogram for the clinical datasets (Nelder-Mead optimizer).

(Fig. 1.9). Under this scenario, image information provided by TEE is used to visualize the aortic root, removing the need for x-ray contrast throughout the procedure. In this paradigm, real-time XRF/TEE registration is needed to minimize errors resulting from cardiorespiratory motion of the probe, which was shown to be as high as 4.1 *mm* in [30].

Aside from the problem of tracking the pose of the TEE probe, the DSC and HYB algorithms may also be applicable to other 2D/3D registration problems. For example, lung tumor tracking by 2D/3D registration for radiotheraphy was proposed in [67], where the authors showed an update rate of 2

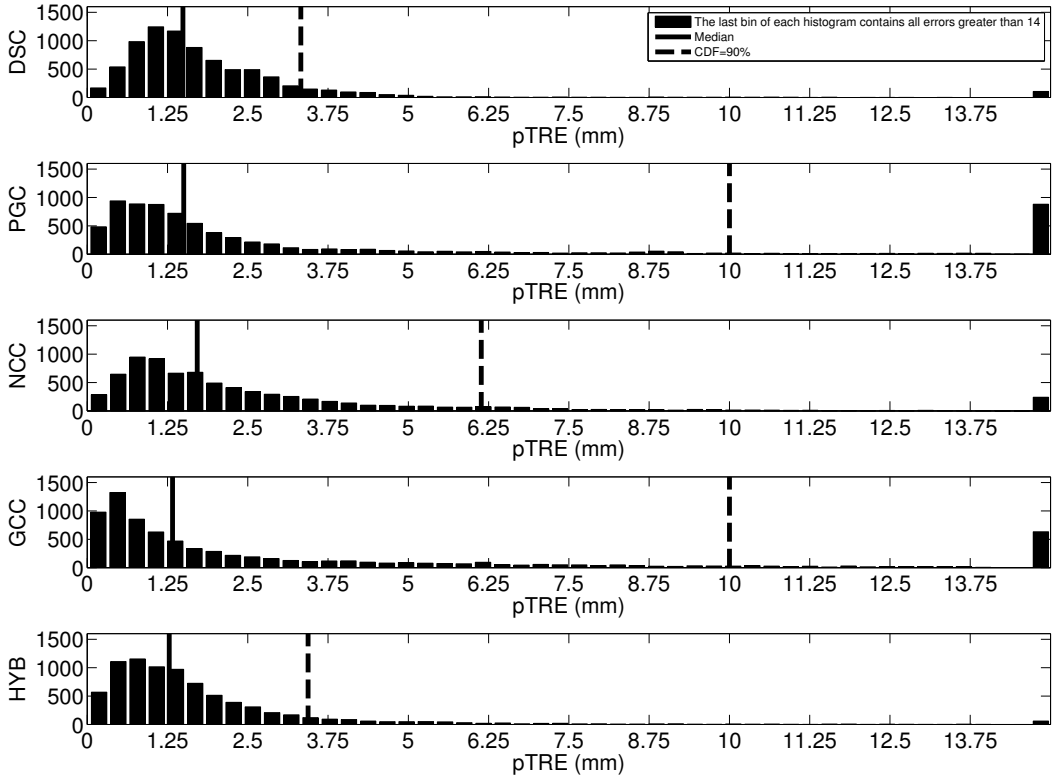

Figure 4.16: pTRE histogram for the clinical datasets (Powell optimizer).

fps. Using the DSC or HYB methods presented in this thesis could result in improved therapeutic targeting accuracy by improving the update speed. As another example, an ICE catheter with embedded fiducial markers for XR-F/ICE registration was recently proposed [68], and the methods presented in this chapter may be applied to that application as well.

5. A Novel Intraprocedural, Contrast-Free Method for Obtaining the Optimal Fluoroscopic Projection Angles for TAVR

A critical step in the TAVR clinical workflow is deployment of the prosthetic valve. In order to minimize the chance of paravalvular leak, device embolization, coronary obstruction, and/or conduction block, the PHV needs to be positioned and expanded at the correct location within aortic annulus. This step is typically guided by XRF, but XRF is a 2D imaging modality, and in order to successfully position the PHV within the 3D anatomical structure of the aortic root, the use of an optimal XRF projection angle is required. The optimal projection angle is defined as the gantry angulation (pair of LAO/RAO and CRA/CAU angles) that produces a 2D XRF image which is perpendicular to the aortic annulus. When the aortic annulus is perpendicular to the XRF image (and therefore parallel to the source-detector axis), the positioning of the PHV as seen in the 2D XRF image most closely correlates to the true 3D positioning (Fig. 5.1).

Figure 5.1: Left (Image modified from [18]): CT volume renderings of an aorta with different virtual XRF projection views. The view in the middle aligns the inferior portions of all 3 coronary cusps, indicating that the annulus is parallel to the source-detector axis, as opposed to the view on the left. Right (Image modified from [19]): An example XRF image during an aortogram, where all three cusps are aligned and therefore at the optimal XRF projection angle.

A few methods have been proposed for obtaining the optimal projection angles, which are outlined below:

Aortography The most common method for obtaining the optimal XRF projection for TAVR is through the use of aortography. This method is based on the fact that the aortic annulus is approximately parallel to the plane containing the inferior tips of the three aortic valve cusps. It follows that, when the inferior tips of the three valve cusps are co-linear in the XRF image (Fig. 5.1), the aortic annulus is parallel to the source-detector axis. The main problem with this method is that it requires repeated aortography, which in turn increases the dose of x-ray contrast agent to the patient. Due to the anatomical characteristics of the patient population, even experienced physicians can have trouble identifying the cusps and aligning them as a result of abnormal aortic valve anatomy. For example, one of the cusps may be misshaped or contain extensive calcium, affecting its visibility and rendering the procedure unreliable. Nevertheless, this method is considered the standard for obtaining or at least confirming the optimal XRF projection angles.

Preoperative CT Preoperative CT can be used to estimate the optimal XRF projection by segmenting the cusps in the volumetric CT image, and finding the angle that projects the cusps onto a straight line in a virtual fluoroscopy system [19]. The underlying assumption behind this technique is that the patient anatomy during the CT scan is in the same position and orientation relative to the XRF imaging system used during the TAVR procedure. This may not be the case, as the patient may be positioned differently or internal

anatomy may have shifted, especially when the CT scan is taken weeks before the procedure. In [19], it was shown that using CT to obtain the optimal XRF projection decreased the number of suboptimal PHV deployments from 35% to 10% in two 20 patient groups. Despite the fact that this method may not reflect the exact anatomical pose of the aortic root during the TAVR procedure, it typically gives a good starting position for the aortography method, decreasing the number of aortograms needed to find the optimal projection.

Intraoperative CBCT Intraoperative CBCT can be used obtain the optimal projection angles at the beginning of the TAVR procedure using the same principles as preoperative CT, with the main advantage being that the images more accurately represent the intraprocedural anatomical state of the patient. In [18], it was shown that the projection predicted by CBCT (Siemens DynaCT) was closer to the final pose of the deployed valve than the projection predicted by preoperative CT. Furthermore, it was noted in [18] that CBCT required less x-ray contrast that preoperative CT due to direct injection into the aortic root rather than intravenous injection. However, in most cases, preoperative CT is performed anyway due to the need for procedural planning, and therefore in practice CBCT further increases the dose of x-ray contrast and ionizing radiation. Another important problem with CBCT is that the acquisition is difficult to perform due to space constraints and the presence of auxiliary equipment in the cath lab, which often results in non-trivial workflow disruptions.

Echocardiography Using echo to determine the optimal XRF projection has also been explored recently. In [69], the authors developed and validated an ad hoc technique for finding the optimal projection angle using TTE. First, the TTE probe was placed on the chest such that the aortic annulus was in the exact center of the 2D echo slice. Next the TTE probe was kept still, and a ruler was placed on the chest next to and parallel to the probe. Finally, the XRF gantry was rotated so that the detector was perpendicular to the ruler. While this technique was suboptimal in many ways, it essentially worked using the same principles as the method presented in this chapter.

In the introduction it was noted that XRF/TEE fusion has already been implemented as a clinical product (EchoNavigator, Philips Healthcare). Recently, a one page letter was published in the Journal of the American College of Cardiology [70] that presented the idea of using EchoNavigator for optimal gantry alignment. However, methods and results have not yet been published.

Study Purpose In this study, the optimal projection angle obtained using multiple methods was compared. As the standard method used at UW hospital, the deployment projection angles obtained via aortography were used as a baseline for comparison with:

- The projection angles obtained via preoperative CT (CT method).
- The projection angles obtained via XRF/TEE registration (echo method).

It was hypothesized that the echo method would agree with the aortographic method to within a mean of 5°, and that the echo method would outperform the CT method in terms of comparison with aortography.

5.1 Methods

Rather than computing differences between the RAO/LAO and CRA/CAU angles reported by each method as done in [18], the error (θ_{error}) was defined as the angle between vectors perpendicular to the annular plane as measured by each modality.

The aortic annulus was modeled as a circular set of points A, with a center c and a unit vector $\vec{v}$ perpendicular to A. In order to specify the modality that the annulus was detected in, as well as the modality that the annulus was registered to, the following notation is adopted: $\vec{v}^{(originalmodality)}_{currentmodality}$. For example, the proposed method required the annulus to be segmented in echo and then registered to XRF. The vector representing the annular plane from echo following registration to XRF would be would be $\vec{v}^{(echo)}_{xrf}$. When the annulus is optimally aligned with the XRF system, $\vec{v}_{xrf}$ is perpendicular to the source detector axis, which is defined by the vector $[0, 0, 1]$. Therefore, $\theta^{(echo)}_{error}$ is defined as the angular deviation from perpendicularity, which is computed as:

$$\theta^{(echo)}_{error} = \frac{\pi}{2} - cos^{-1}(\vec{v}^{(echo)}_{xrf} \cdot [0, 0, 1]^T) \tag{5.1}$$

By definition, $\vec{v}^{(aortography)}_{xrf}$ is assumed perpendicular to the source-detector axis. In order to compute $\vec{v}^{(echo)}_{xrf}$, the following equation is used.

$$\vec{v}^{(echo)}_{xrf} = {}^{xrf}\, T_{probe}\, {}^{probe}T_{echo}\, \vec{v}^{(echo)}_{echo} \tag{5.2}$$

The methods for determining ${}^{xrf}T_{probe}$ and ${}^{probe}T_{echo}$ are detailed in chapter 2 other and Appendix B.

A different calculation was used to obtain the difference between the projection angle determined by preoperative CT and aortography. The preoperative CT projection is provided as a set of estimated RAO/LAO and CRA/CAU angles, which can be compared to the RAO/LAO and CRA/CAU angles used for PHV deployment (available in the DICOM headers). This allows for the creation of two vectors, one representing the annulus plane estimated via aortography and the other the annulus plane estimated via CT. The angle between these two vectors is used to determine $\theta_{error}^{(CT)}$:

$$\vec{v}^{(deployment)} = R_{CC}^{(deployment)} R_{RAO}^{(deployment)} \cdot [0,0,1]^T \tag{5.3}$$

$$\vec{v}^{(CT)} = R_{CC}^{(CT)} R_{RAO}^{(CT)} \cdot [0,0,1]^T \tag{5.4}$$

$$R_{CC} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_{CC} & -s_{CC} \\ 0 & s_{CC} & c_{CC} \end{bmatrix} \tag{5.5}$$

$$R_{RAO} = \begin{bmatrix} c_{RAO} & 0 & -s_{RAO} \\ 0 & 1 & 0 \\ s_{RAO} & 0 & c_{RAO} \end{bmatrix} \tag{5.6}$$

$$c_a = \cos(\theta_a), s_a = \sin(\theta_a) \tag{5.7}$$

$$\theta_{error}^{(CT)} = cos^{-1}(\vec{v}_{deployment} \cdot \vec{v}_{CT}) \tag{5.8}$$

Finally, post-deployment PHV angulation was also measured. Similar methods were used as alternative ground truth measures in [18] and [19]. By analyzing the post-deployment XRF image, the out-of-plane rotation (tilt) of the PHV can be estimated. Because the aortic annulus is aligned with the long-axis of the deployed prosthetic valve, the tilt is used as a measure of how well the final annulus orientation matched the orientation obtained from aortography.

The $\theta_{error}^{(PHV)}$ is computed by examining the struts of the deployed PHV, which are arrange in a circular pattern. If the deployed PHV is perfectly perpendicular to the source-detector axis, the struts will form a line segment in the XRF image. Otherwise, they will form an ellipses with major axes of length a and minor axes of length b (Fig. 5.2). $\theta_{error}^{(PHV)}$ is calculated as:

$$\theta_{error}^{(PHV)} = sin^{-1}(\frac{b}{a}) \tag{5.9}$$

Because there are two pairs of struts on both sides of the PHV, two values of $\theta_{error}^{(PHV)}$ are computed, which may not be exactly the same. The final reported value is the mean of these two values.

The underlying assumptions behind this method are that the annulus orientation isn't affected by the deployment of the valve, and that the PHV struts are circular after deployment. One limitation of this method is that only the absolute value of $\theta_{error}^{(PHV)}$ can be computed, as it is not clear from the

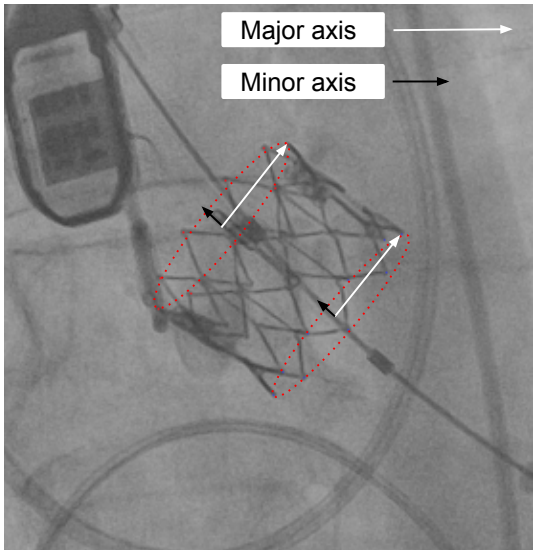

Figure 5.2: Methodology for computing $\theta_{error}^{(PHV)}$ for the deployed PHV.

2D projection image which direction the PHV tilts.

5.2 Experimental design

Ten patients undergoing TAVR at UW Hospital were enrolled in the study. All patients gave written consent. Patients underwent preoperative CT imaging 14-28 days prior to the procedure. The preoperative CT was used to estimate the initial XRF projection via segmentation of the aortic valve cusps by a trained radiologist. During the procedure, the initial CT based projection angles were refined using aortography. If it seemed as if the TEE probe would not obscure the aortic root during aortography, a 3D echo was recorded simultaneously with the final aortogram. Otherwise, a second XRF sequence was recorded following aortography with simultaneous 3D echo imaging.

5.3 Data Processing

All data was processed offline using custom software (Appendix B). TEE was registered to XRF using the patch gradient correlation (PGC) method from (see chapter 3). The transformation between the TEE probe and echo volume was computed using the method specified in Appendix B. Annulus segmentation was performed by a trained cardiologist, who adjusted the segmentation until it conformed to the example specified in [71]. Segmentation was always performed during systolic frames as this most closely matches the anatomical position of the aortic root during rapid pacing and hence PHV deployment. Segmentation results for every case are shown in Appendix D.

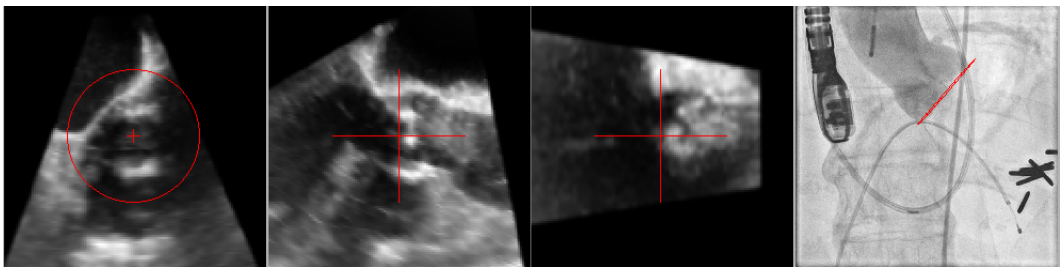

Figure 5.3: Example segmentation of the aortic annulus during simultaneous aortography and echo imaging, with the segmented annulus projected onto the XRF image.

5.4 Results

Results for each case are shown in Fig. 5.4. The mean and standard deviation for $\theta_{error}^{(echo)}$ was $3.46° \pm 3.28°$, while for $\theta_{error}^{(CT)}$ it was $9.35° \pm 6.50°$. A paired t-test was used to determine the p-value between the echo and CT error measure-

ments, which was found to be 0.042. The mean difference between the echo and CT methods was $5.88° \pm 7.30°$.

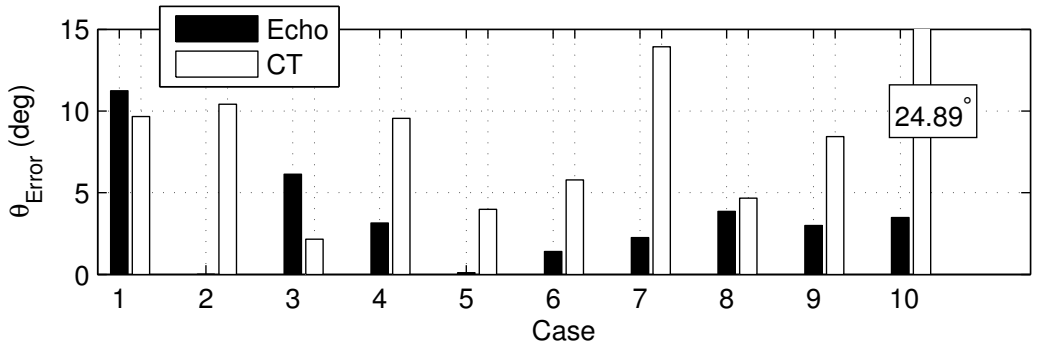

Figure 5.4: $\theta_{error}^{(echo)}$ and $\theta_{error}^{(CT)}$ for each case ($p = 0.042$).

Fig. 5.5 shows the tilt angle $\theta_{error}^{(PHV)}$ of the PHV following deployment for each case. The mean $\pm$ std error was $7.01° \pm 2.78°$. This indicates that optimal gantry alignment via aortography may be suboptimal as a ground truth measure, with the caveat that the tilt of the deployed valve may not exactly represent the tilt of the aortic annulus prior to deployment.

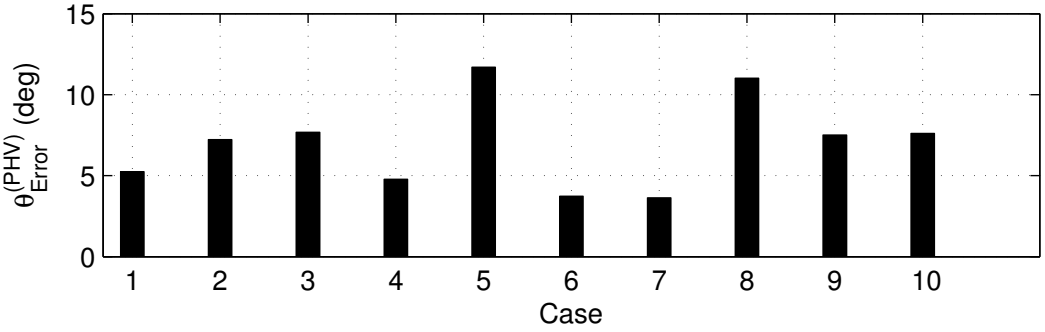

Figure 5.5: $\theta_{error}^{(PHV)}$ for each case.

5.5 Discussion

In this study, a novel echo-based method for estimating the optimal projection for TAVR, as well as a prior method utilizing preoperative CT, were compared against the standard clinical method using aortography. It was shown that the optimal projection angle estimated using the echo method was, on average, more accurate in relation to aortography than preoperative CT ($p = 0.042$). This is likely because CT is usually performed under different patient positioning and often weeks prior to the procedure, resulting in differing positions of the aortic root relative to the imaging system. The advantage of the XRF/TEE registration based method is that it can be performed during the procedure without disrupting the clinical workflow or adding additional x-ray contrast dose.

The main limitation of this study is that the ground truth was obtained via aortography, which both the XRF/TEE registration and CT methods were compared against. However, as shown by our analysis of post-deployment PHV tilt, aortography cannot always perfectly predict the proper gantry angle. It could be argued that a better ground truth would be intraoperative CBCT, as this method is able to reconstruct the full aortic root anatomy in the middle of the procedure, and is intrinsically registered to the XRF imaging system. However, this was not possible in our study, as it was determined at our institution that the benefits of intraoperative CBCT did not justify its disruption of the clinical workflow and addition of x-ray contrast dose.

Other limitations include:

- Although prosthetic valve deployment is performed under rapid pacing, the 3D echo images were not recorded during rapid pacing except for in one case. It is unknown if segmenting the annulus in systole is a close enough substitute for comparison, or if this introduced systematic segmentation errors.

- A few of the 3D echo scans were recorded at low resolution and/or with the aortic annulus partially out of the field-of-view, which may have resulted in segmentation errors.

- This study was conducted using expert segmentations. Therefore, intra and inter-rater variability in the segmentation should be examined to determine upper and lower bounds on the accuracy of the method.

- Manual segmentation of the aortic annulus plane in echo is currently required, and it is unclear how much time is needed to complete this step in the middle of a real procedure. Automatic methods would resolve this issue, but accurate algorithms need to be developed and validated.

In future work, echo imaging parameters such as field-of-view, depth, and image resolution will be standardized. Another interesting topic is the use of echo-contrast agents to aid with the segmentation of the aortic annulus, as this may improve accuracy in cases where echo image quality is poor. Finally, the workflows for intraprocedural use of XRF/TEE registration, such as shown in Fig. 1.10 of the introduction, will be implemented and validated.

6. Summary and Future Work

Thesis Summary In this thesis, novel real-time image processing algorithms and applications aimed at improving TAVR were presented and validated. More specifically, the work focused on designing improved algorithms for image registration between x-ray fluoroscopy and transesophageal echocardiography, which were determined to be necessary for enabling advanced TAVR clinical workflows.

In the first aim, a GPU-accelerated implementation of the Hough Forest object detection algorithm was developed and applied to the problem of in-plane localization of devices in XRF images. The purpose was to minimize the amount of manual interaction required for 1) XRF/TEE registration, which was one of the main topics of this thesis and, 2) PHV tracking, which can be used to automatically selected device centered views in 3D echo images. It was shown that the TEE probe could be successfully localized in 95.8% of 1077 images and that the PHV could successfully be localized in 90.1% of 388 images.

Next, novel XRF/TEE registration algorithms were introduced that, for the first time, demonstrated real-time, accurate registration at fluoroscopic frame rates. These algorithms can be used to enable new clinical workflows, where moving anatomical targets in echocardiographic images can accurately be registered to fluoroscopic images.

Finally, a method for determining the optimal XRF projection angles for prosthetic valve deployment using XRF/TEE registration was presented and validated in 10 patients. Results indicated that the proposed method was accurate to within $3.46° \pm 3.28°$ of the clinical gold standard (aortography), and

was more accurate than projection angles obtained from CT. A major advantage compared to other methods is that the novel method does not require the use of iodinated contrast agent and can be quickly performed at any time during the procedure without disrupting the clinical workflow.

Future Work On the topic of automatic object detection related to the work presented in chapter 2, the following areas are the focus of future work:

- Following detection of the TEE probe, estimation of the initial out-of-plane parameters is necessary, otherwise the 2D/3D registration algorithm may not converge properly. A fast, GPU-based template matching approach, similar to the one proposed in [33], is a possible solution. Another possible solution is to use the Hough Forest algorithm to regress out-of-plane parameters along with in-plane parameters.

- For the PHV, localization of the undeployed PHV may be useful for pre-deployment positioning and automatic selection of device centered echo views. However, the ability to not only detect the PHV position but also the degree of deployment would enable the ability to visualize valve deployment in echo, and future work will focus on estimating the valve deployment process from 1 or 2 XRF projections.

- Because the Hough forest algorithm is generic in the sense that in can be applied to any type of image, an interesting application would be automatic detection of aortic root landmarks, which would enable the method presented in chapter 5 to be automated.

For the real-time 2D/3D registration presented in chapters 3-4, future work will focus on the following:

- The presented algorithms will also be validated on a bi-plane XRF system, which have been shown to be more accurate but also slower using prior methods [30].

- Because it is possible to compute the DSC cost function Jacobian analytically, implementation of the DSC method using optimizers that require gradient information (e.g. LBFGS) may result in even faster computations, and will be investigated.

- Application of the proposed algorithms to other 2D/3D registration problems, such as lung tumor tracking or other types of devices, will be investigated

Finally, for the work from chapter 5, the following topics will be addressed

- The work presented in this thesis was a from a small initial patient population, and more robust statistics would likely result from a larger study.

- Intra and inter-operator variability is an important issue with regard to obtaining the optimal XRF projection angles using either aortography, echo, or CT, and future clinical studies should be conducted that measure not only the mean accuracy of each method, but also the variability associated with each method.

In addition, the image processing workflow envisioned in the introduction depended not only upon real-time XRF/TEE registration, but also on real-time echo segmentation and/or CT/echo registration. These topics are challenging, particularly due to the large image volumes and low signal-to-noise ratios present in echo images, and are also a topic of future work.

A. Philips 3D Echo DICOM

Accessing the 3D echo volumes created by the Philips iE33 ultrasound system is a non-trivial task. This section outlines how access to the image data and metadata was obtained so that it could be analyzed. A listing of important proprietary DICOM tags and their meanings are included.

A.1 Getting Echo Data From the Philips iE33

The process of obtaining usable echo data from the iE33 machine is as follows:

1. On the iE33 machine, select the desired 3D volumes and press the "Save Media As.." and export to DVD.

2. The iE33 will export what we will call "Native" DICOM files to DVD. However, when these files are loaded into `MATLAB`, the images will be flattened 2D views of the 3D volume. The 3D data is stored in the file a proprietary format that cannot be interpreted without special software.

3. In order to generate a usable image volume, the data needs to be pro-

cessed with a special version of Philips' US image analysis software called "Qlab". The special version of Qlab has a "Cartesian" export option, which converts the Native DICOM to a Cartesian DICOM.

4. Once the Cartesian DICOM is exported from Qlab, it has to be read into `MATLAB` and exported as a .mhd file. Because the volume size of the image is not stored using standard DICOM tags in the header, a special trick must be used to read the data properly:

   - In the `MATLAB` file `dicomread.m`, a break-point should be set line 648.
   - Called the `dicomread` function with the Cartesian DICOM file name as the first argument.
   - When the script stops at the breakpoint, type:
     `metadata.SamplesPerPixel = N`
     where `N` is the size of the z-dimension of the volume.
   - Press F5 to continue.

5. Once the data volume has been read into `MATLAB` as a matrix, we it first re-sampled to have isotropic resolution, which makes data processing and segmentation easier. It is then written to disk as a .mhd file using custom `MATLAB` functions.

A.2 Decoding the echo image volume spatial transformation from the DICOM tags

For the clinical study conducted in Chapter 5, the spatial relationship between the echo image volume and the echo probe, ${}^{probe}T_{echo}$, was needed to transform voxels from the echo image to the XRF image space. It was not feasible to perform echo volume calibration for every sequence. Fortunately, this matrix could be estimated from the DICOM header information. This was done by finding the origin of the echo signal, which is depicted in Fig. A.1, as well as the rotation of the echo volume about the z-axis of the TEE probe. ${}^{probe}T_{echo}$ is computed using equation A.1:

$$ {}^{probe}T_{echo} = R_z C S \tag{A.1}$$

where:

$$ R_z = \begin{bmatrix} cos(\theta_{seek}) & -sin(\theta_{seek}) & 0 & 0 \\ sin(\theta_{seek}) & cos(\theta_{seek}) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{A.2}$$

$$ C = \begin{bmatrix} 1 & 0 & 0 & -c_x \\ 0 & 1 & 0 & -c_y \\ 0 & 0 & 1 & c_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{A.3}$$

$$S = \begin{bmatrix} s & 0 & 0 & 0 \\ 0 & s & 0 & 0 \\ 0 & 0 & s & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{A.4}$$

In these equations, s is the voxel spacing ($\frac{mm}{voxel}$), which is used to convert voxel coordinates to physical coordinates, c_x and c_y refer to the x and y spatial coordinates of a virtual ultrasonic point source that emanates from the TEE probe face (Fig. A), c_z is the vertical distance of the z-axis of the volume from the z-axis of the TEE probe, and θ_{seek} is the angle representing the rotation of the echo volume about the z-axis. Fig. A.1 shows how c_x and c_y are computed from steering and elevation angles $\theta_1 - \theta 4$, both of which are provided in the Native DICOM header.

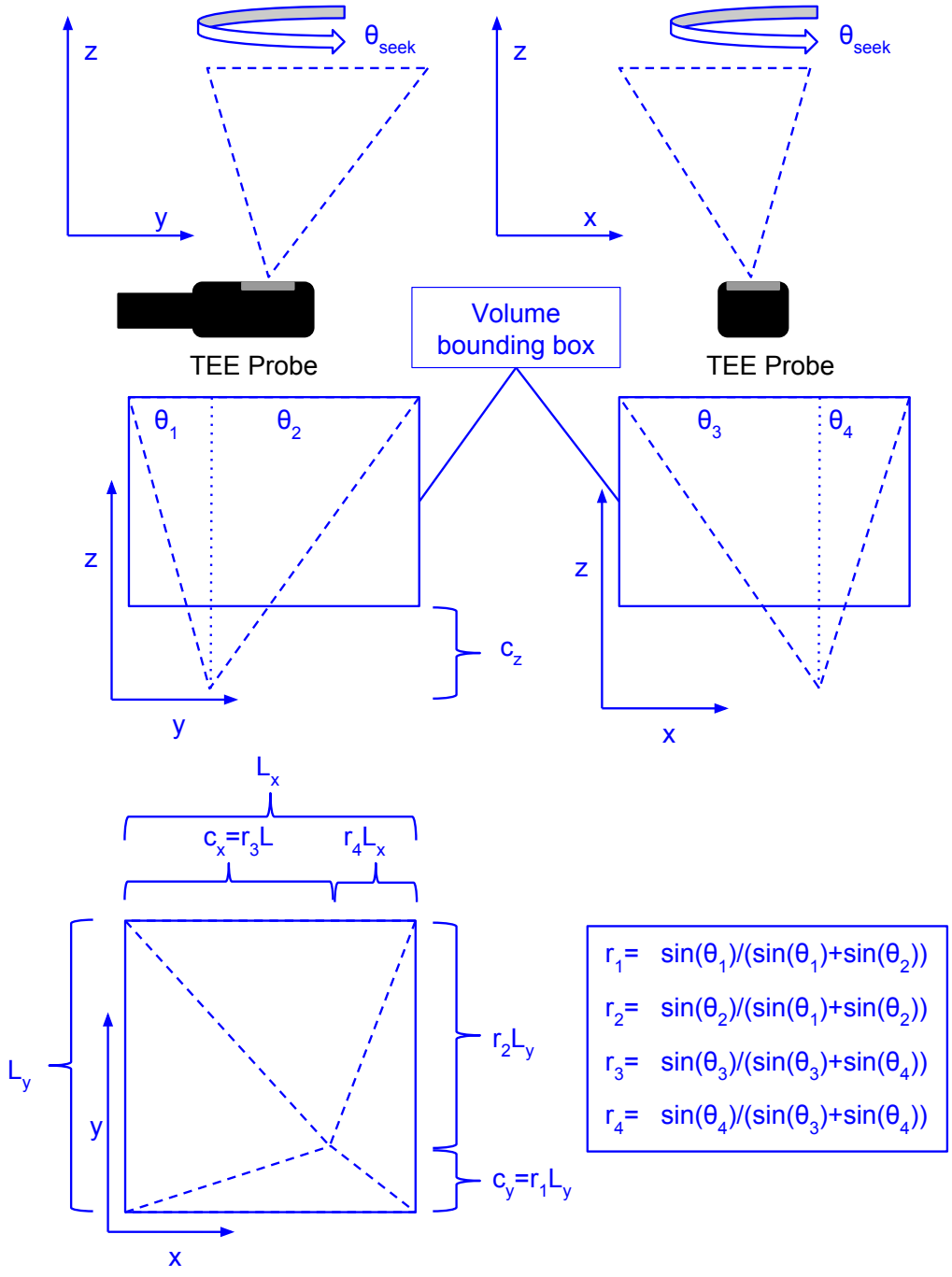

Figure A.1: The 3D echo coordinate system in relation to the TEE probe, with spatial transformation parameters.

A.3 DICOM tags

The following DICOM tags are necessary for either reading the 3D echo volumes correctly or for generating the ${}^{probe}T_{echo}$ transform.

- Native DICOM. See Fig. A for drawings that describe each of these parameters.
  - `Private_200d_3126`: This tag represents the seek angle, θ_{seek}.
  - `Private_200d_3104`: This tag represents the steering angle, θ_3.
  - `Private_200d_3105`: This tag represents the steering angle, θ_4.
  - `Private_200d_3203`: This tag represents the elevation angle, θ_1.
  - `Private_200d_3204`: This tag represents the elevation angle, θ_2.
  - `Private_200d_3102`: This tag represents the physical coordinate of the volume z-axis, c_z.
- Cartesian DICOM
  - `Columns`: The x-dimension size of the volume.
  - `Rows`: The y-dimension size of the volume.
  - `Private_3001_1001`: The z-dimension size of the volume.
  - `NumberOfFrames`: The number of temporal frames.
  - `PhysicalDeltaX`: The pixel spacing in the x-direction.
  - `PhysicalDeltaY`: The pixel spacing in the y-direction.
  - `Private_3001_1003`: The pixel spacing in the z-direction.

- `FrameTime`: The temporal resolution.

B. Custom software

B.1 Aortic Annulus Segmentation Software

For chapter 5, it was necessary to design a GUI for manual aortic annulus segmentation, as it was not possible to perform this operation using commercial software and export it to a usable format.

The software, written in `MATLAB` using `GUIDE`, worked by allowing the user to manually change the pose parameters of the aortic annulus, as well as the temporal frames of the echo and XRF images, using slider controls (Fig. B.1). This allowed to user to view the cut-planes that were parallel and perpendicular to the aortic annulus.

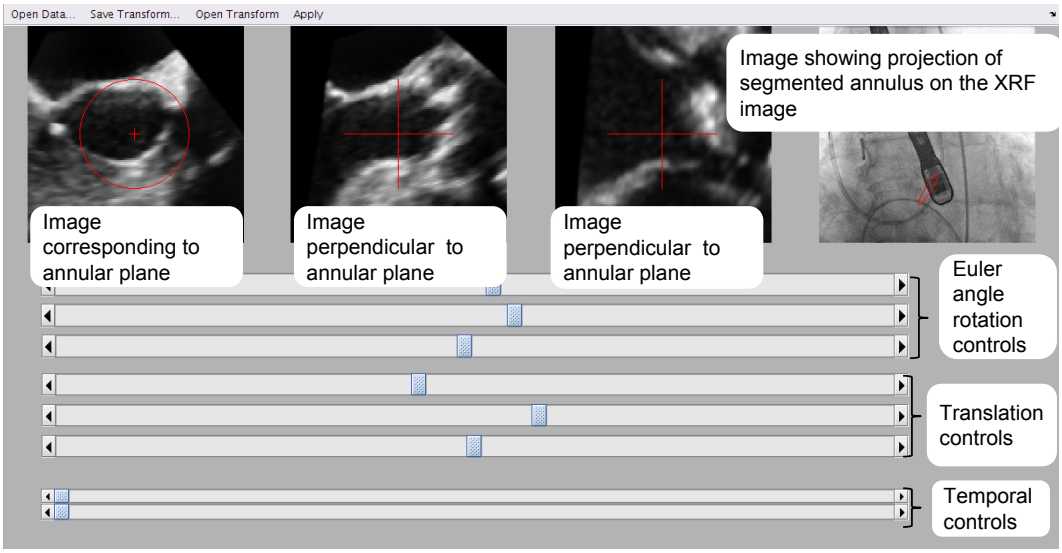

Figure B.1: The annulus segmentation GUI.

B.2 TEE/XRF Manual Pose Estimation Software

For the clinical dataset validation in chapters 3 and 4, it was necessary to manually register the TEE probe to the XRF image in order to generate pseudo-ground truth measures. In this case, manual registration refers to manual initialization followed by optimization-based registration refinement. Therefore, a GUI was designed, again using `GUIDE` in `MATLAB`, for performing XRF/TEE registration. The user was able to quickly change the TEE pose parameters using hot keys and the mouse-wheel. Immediate visualization of the registration was provided by projecting a "ghost" of the probe model on top of the XRF image (Fig. B.2). Three options for registration refinement following initialization were provided: direct splat correlation, gradient correlation, and

patch gradient correlation.

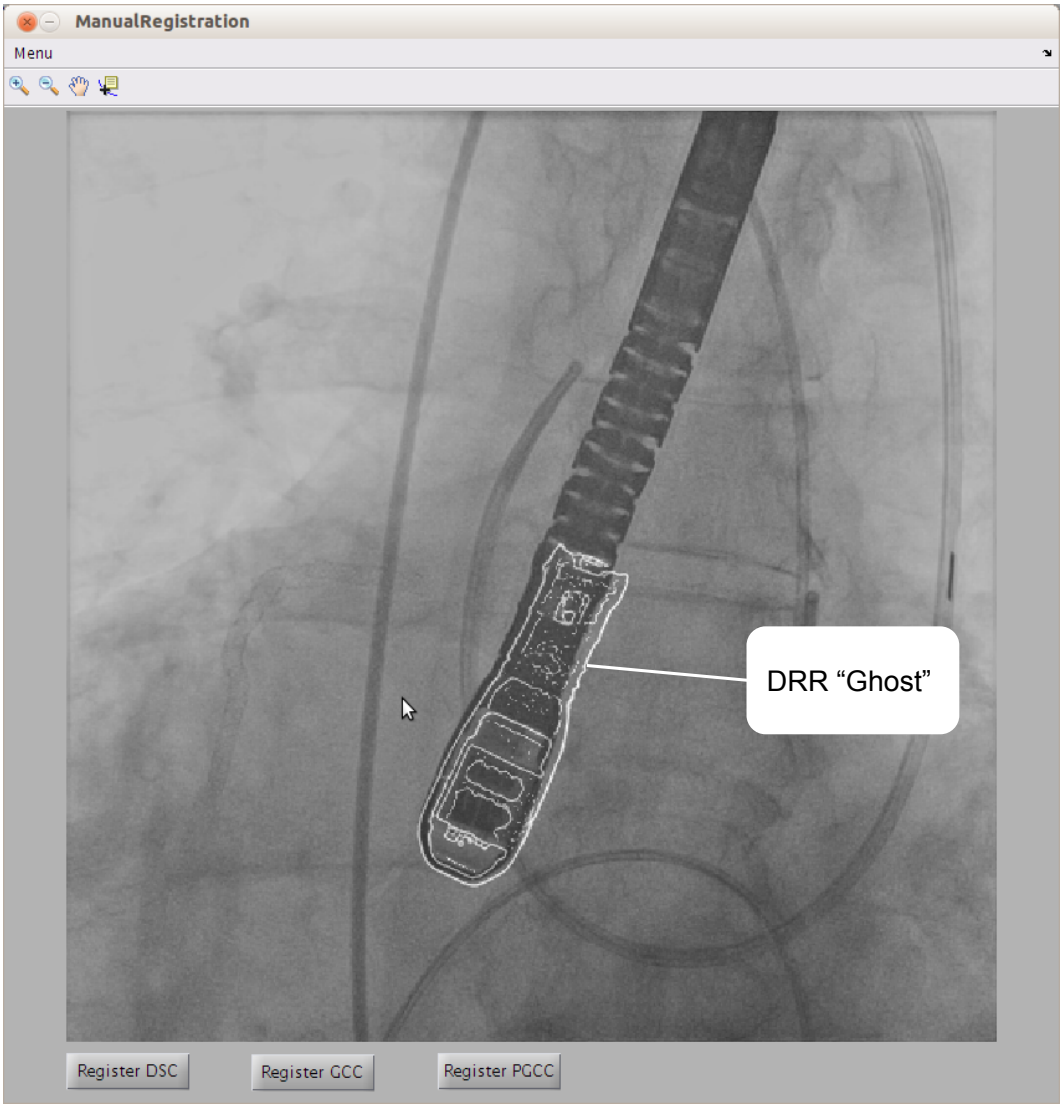

Figure B.2: XRF/TEE manual registration GUI

C. Cost Function Plots

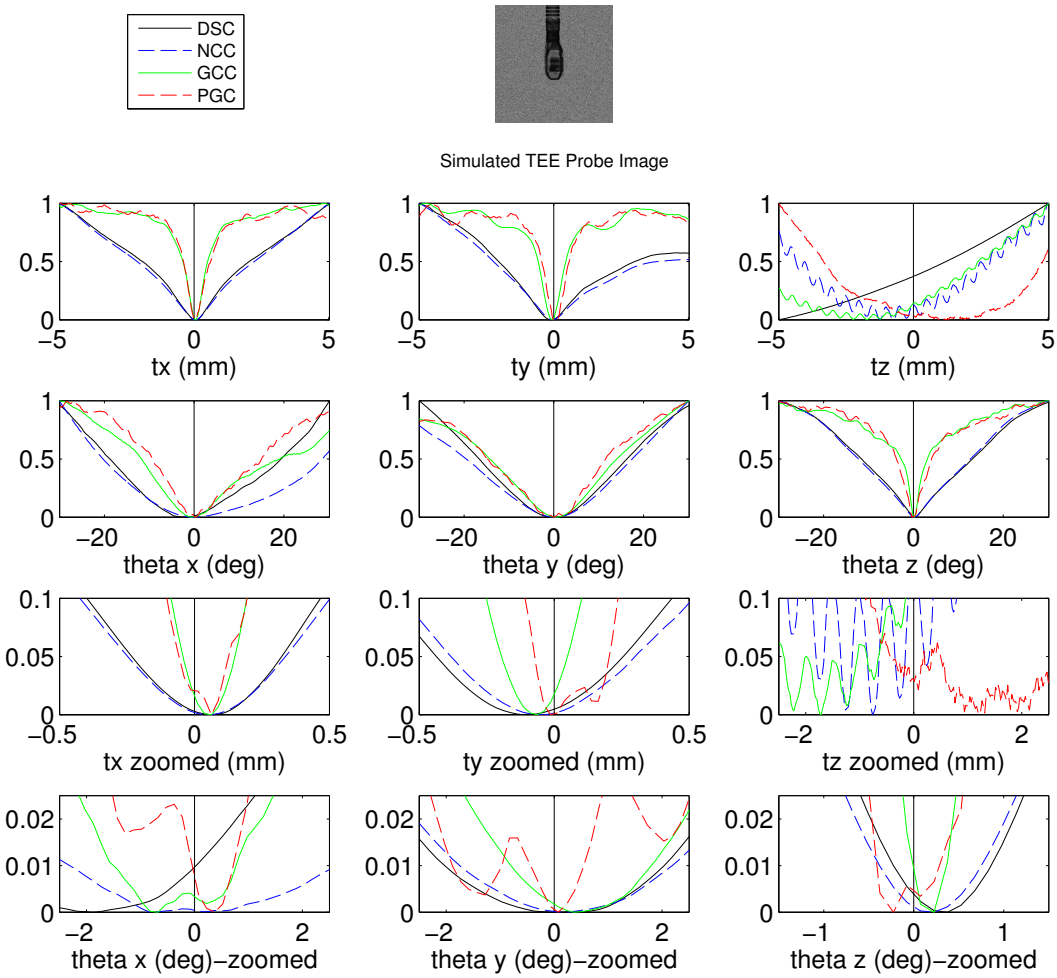

Figure C.1: Cost function curves for a simulated image ($\theta_x = 0°$, $\theta_y = 30°$)

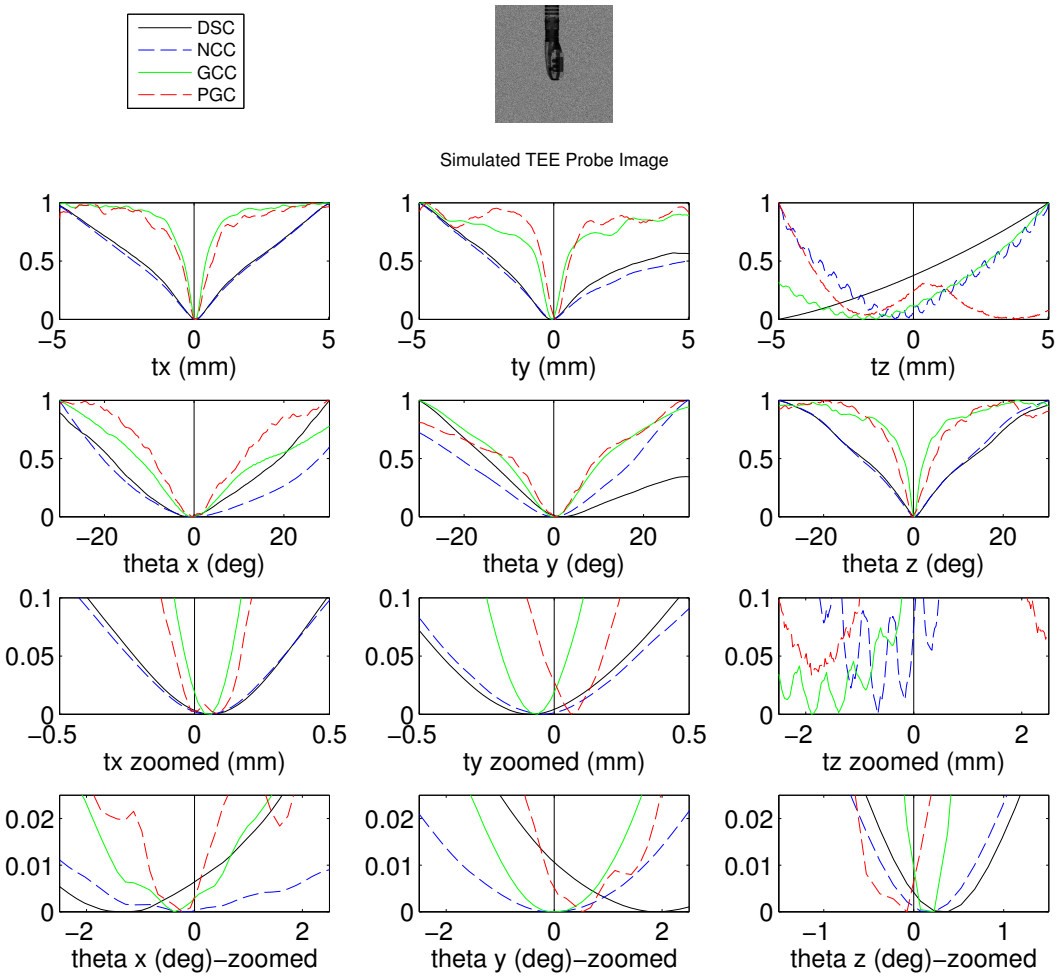

Figure C.2: Cost function curves for a simulated image ($\theta_x = 0°$, $\theta_y = 60°$)

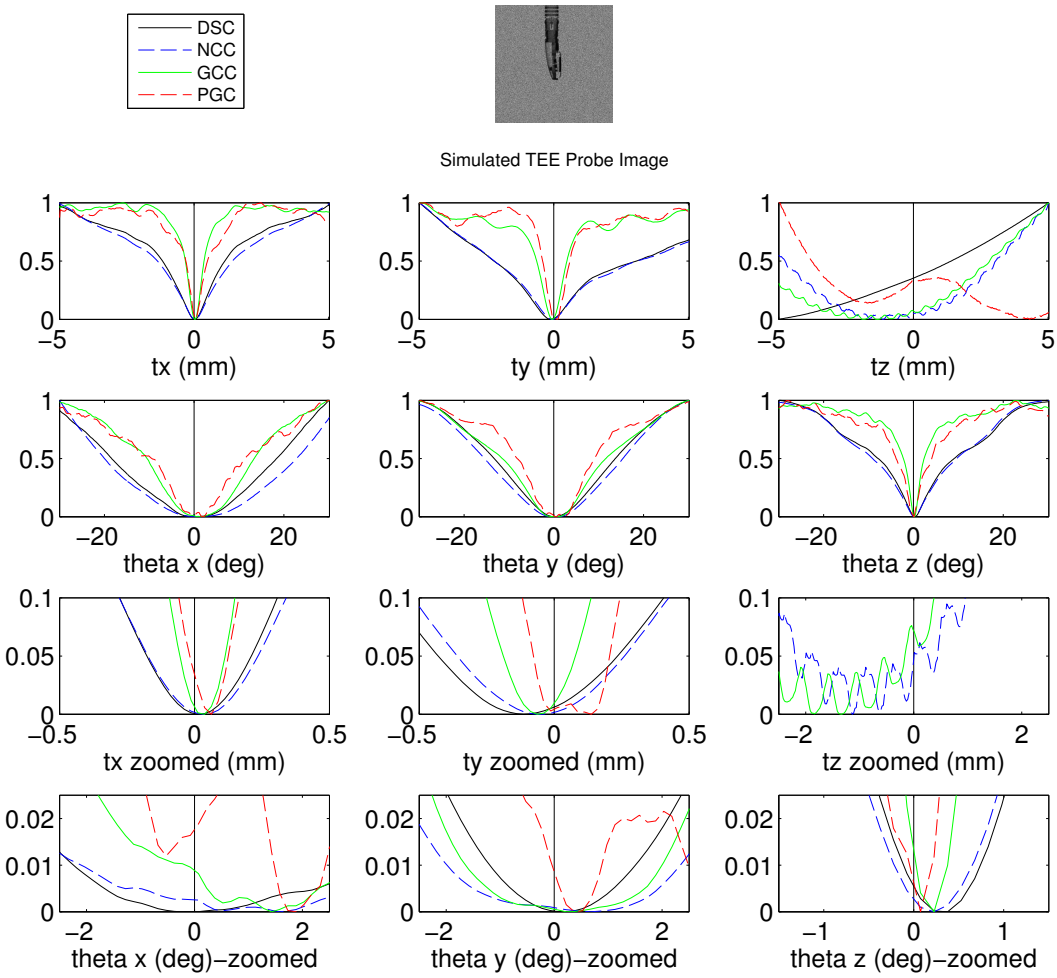

Figure C.3: Cost function curves for a simulated image ($\theta_x = 0°$, $\theta_y = 90°$)

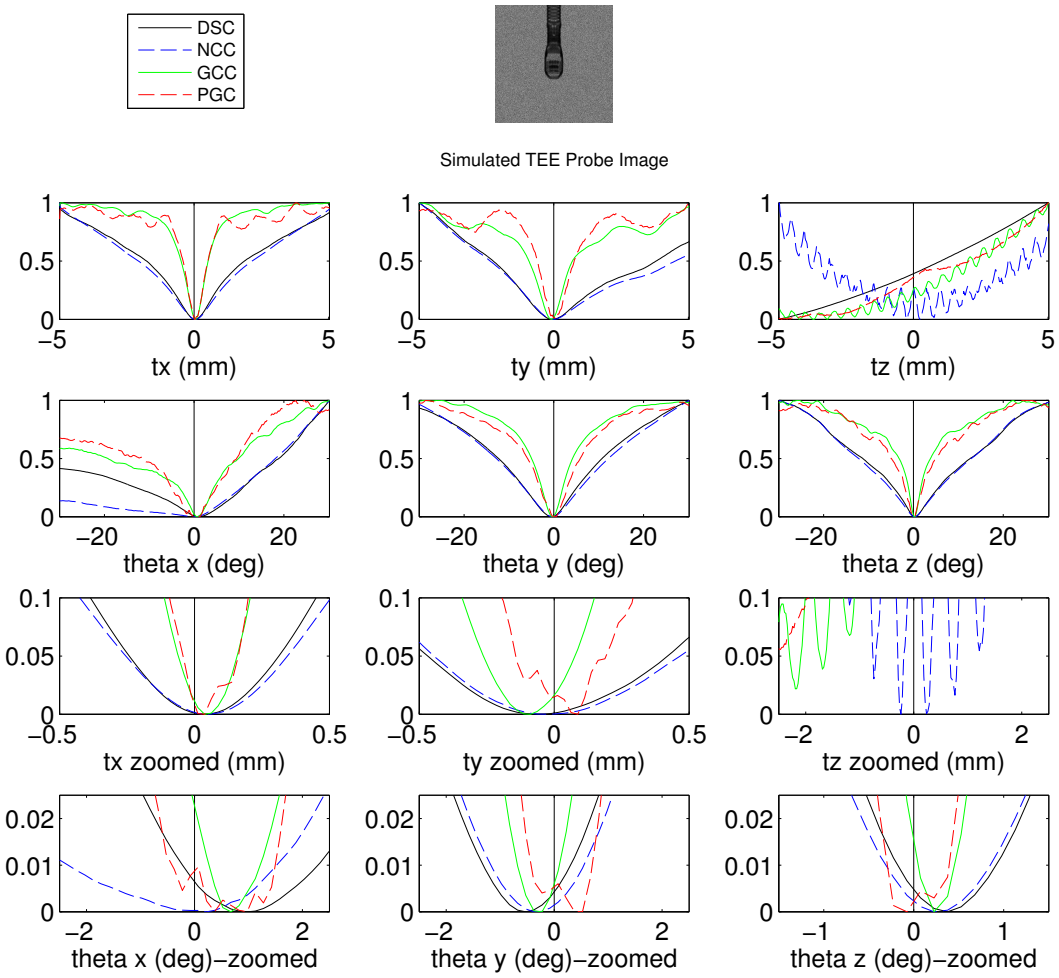

Figure C.4: Cost function curves for a simulated image ($\theta_x = 30°$, $\theta_y = 0°$)

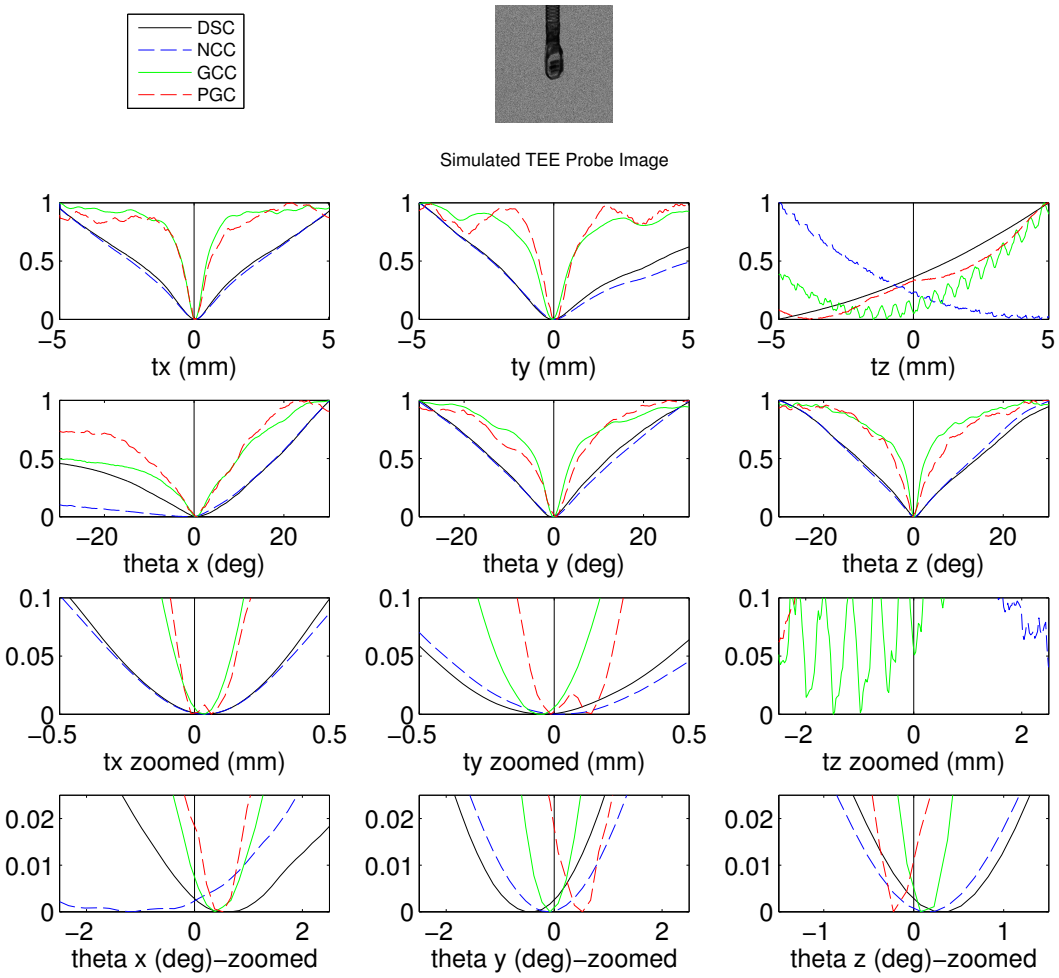

Figure C.5: Cost function curves for a simulated image ($\theta_x = 30°$, $\theta_y = 30°$)

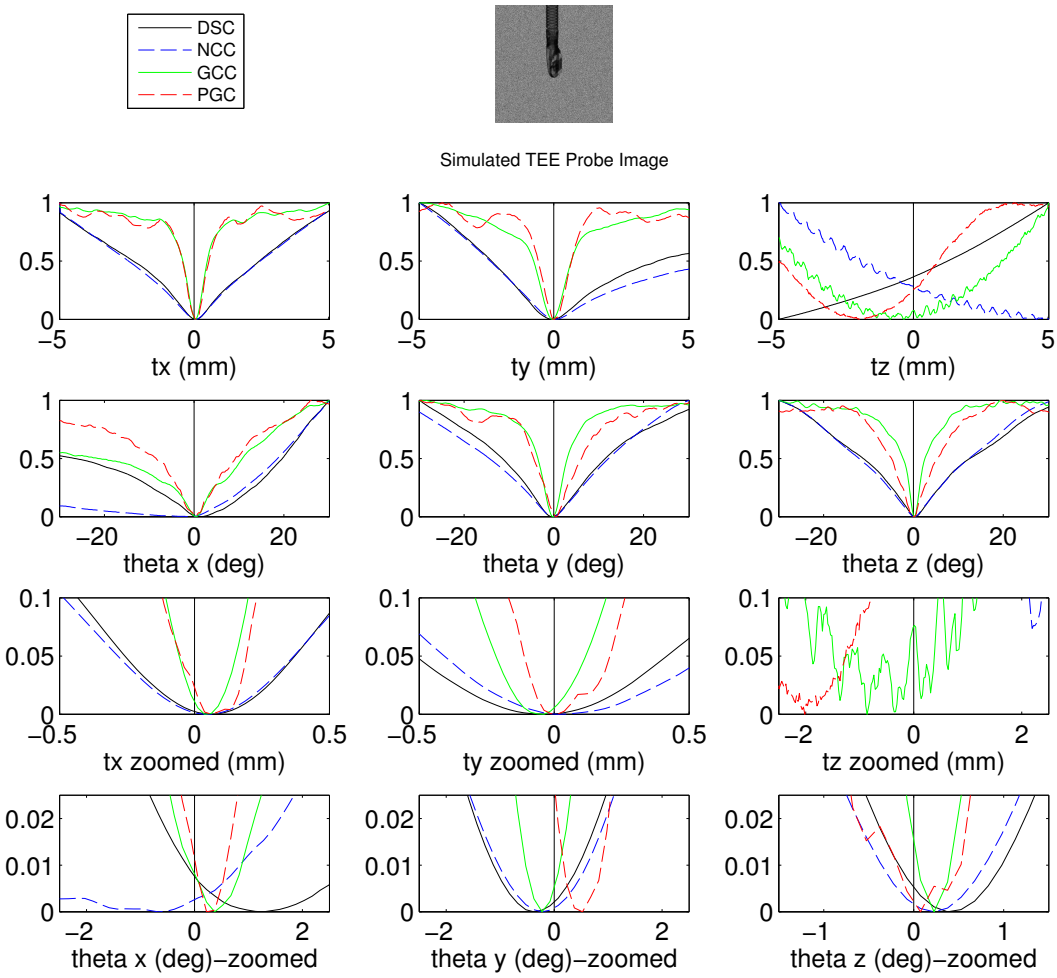

Figure C.6: Cost function curves for a simulated image ($\theta_x = 30°$, $\theta_y = 60°$)

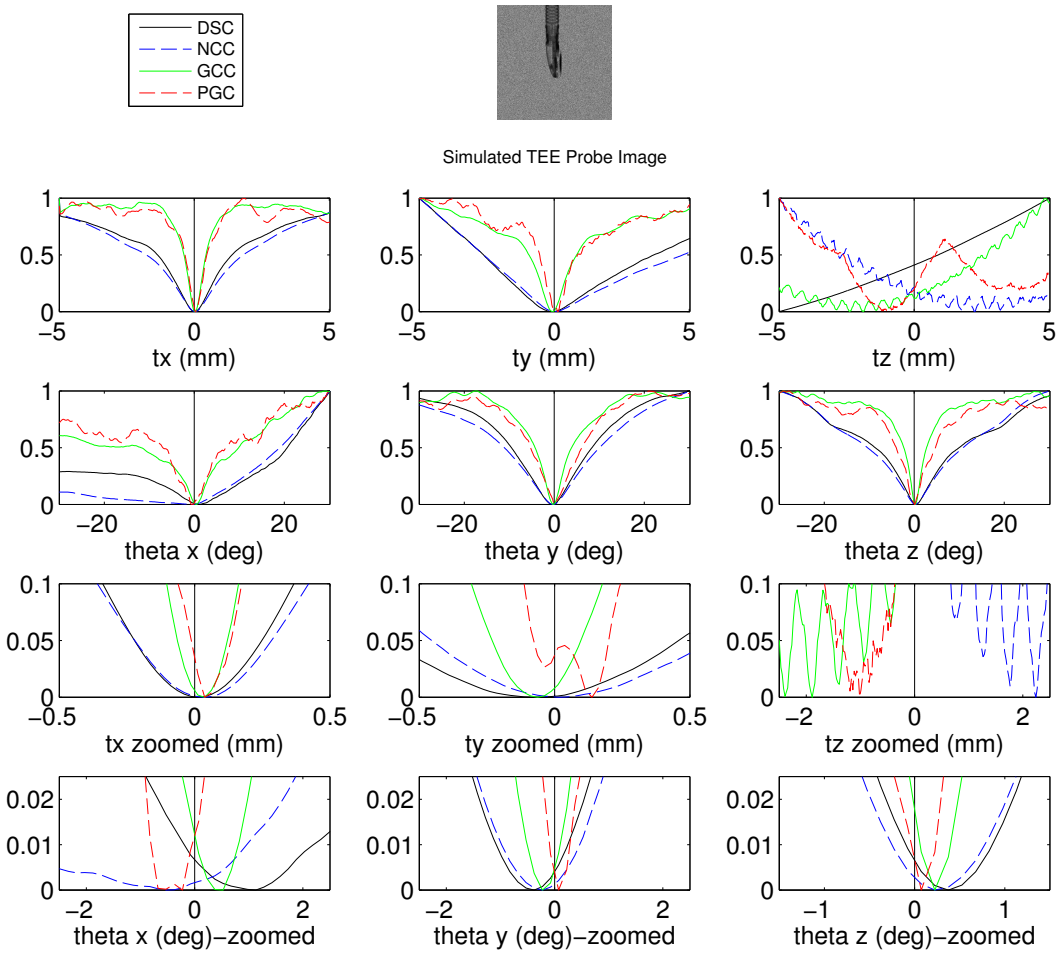

Figure C.7: Cost function curves for a simulated image ($\theta_x = 30°$, $\theta_y = 90°$)

D. Aortic Annulus Segmentation

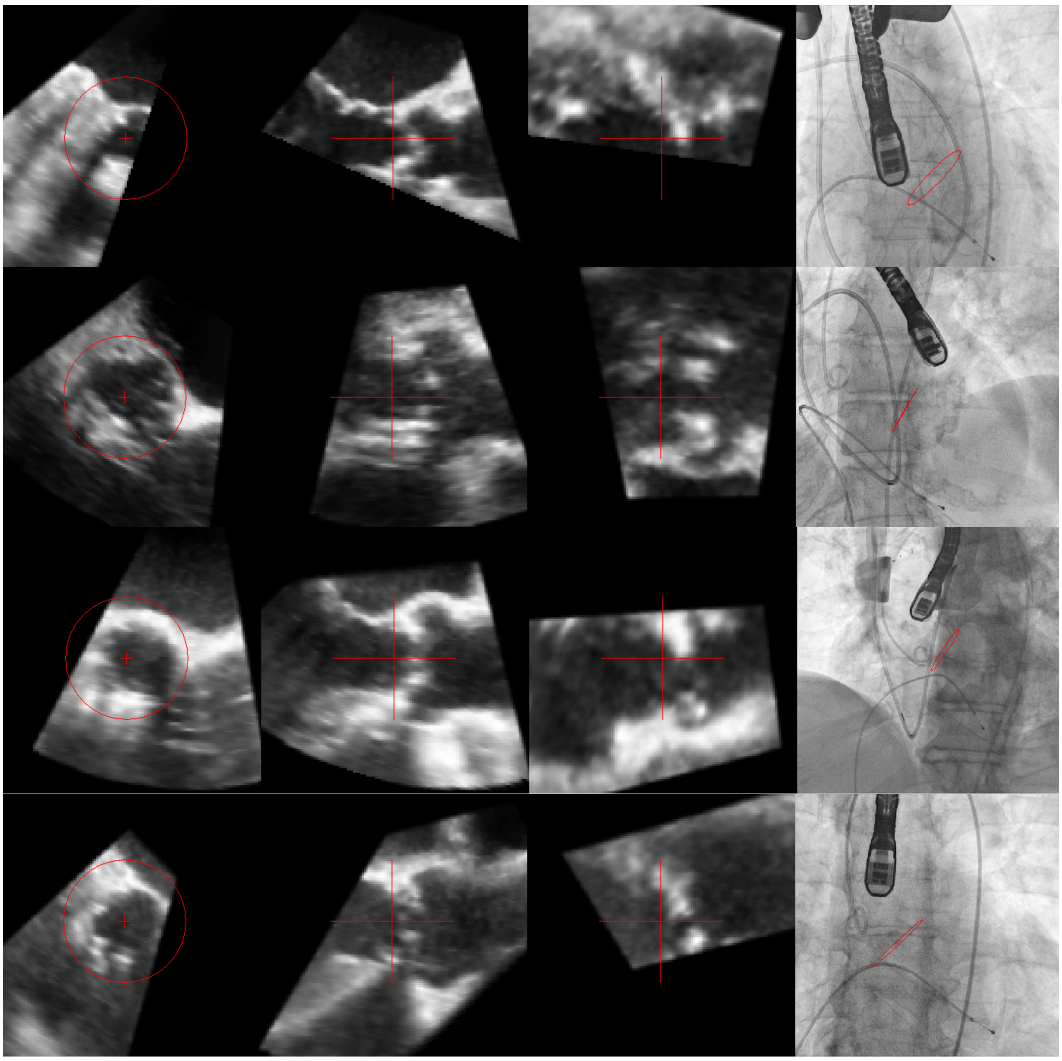

Figure D.1: Aortic annulus segmentation and XRF projection, cases 1-4.

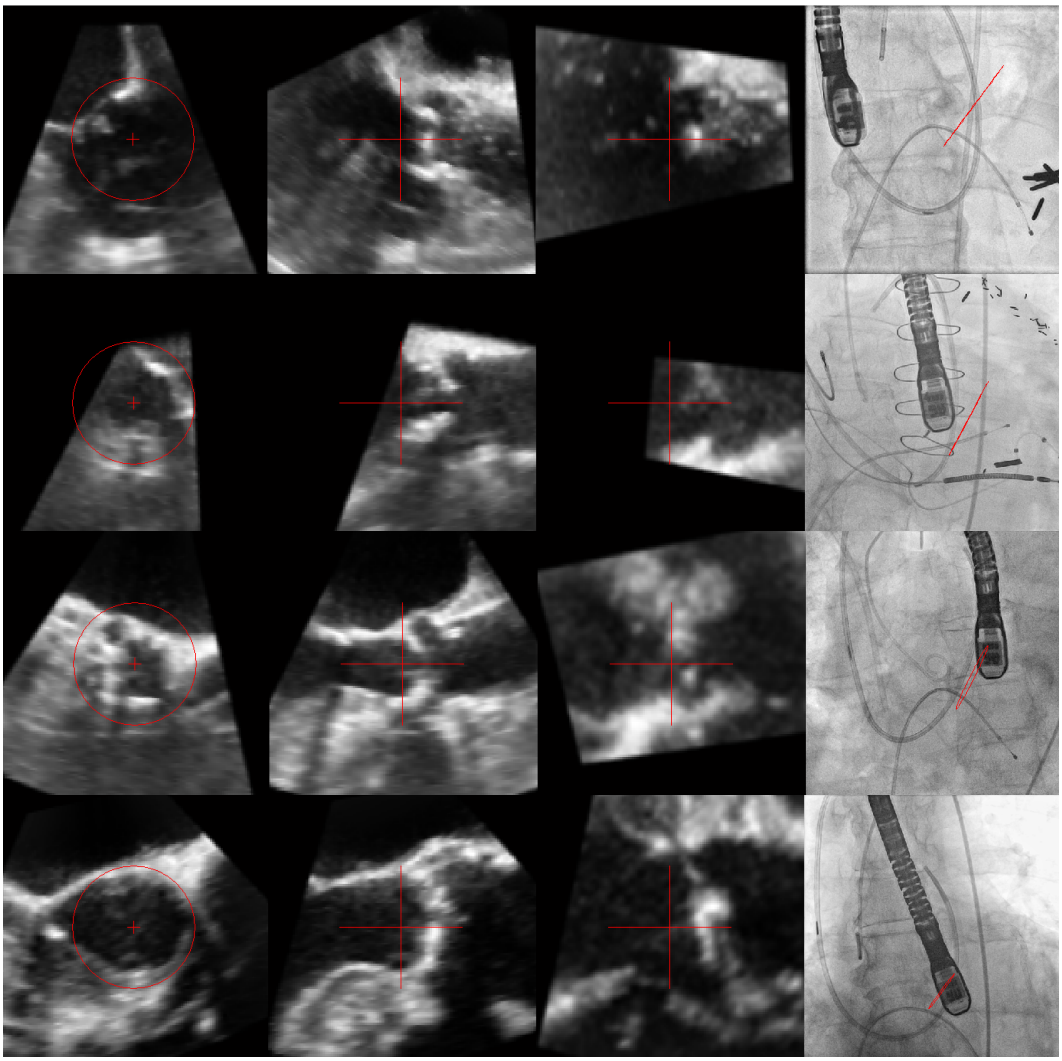

Figure D.2: Aortic annulus segmentation and XRF projection, cases 5-8.

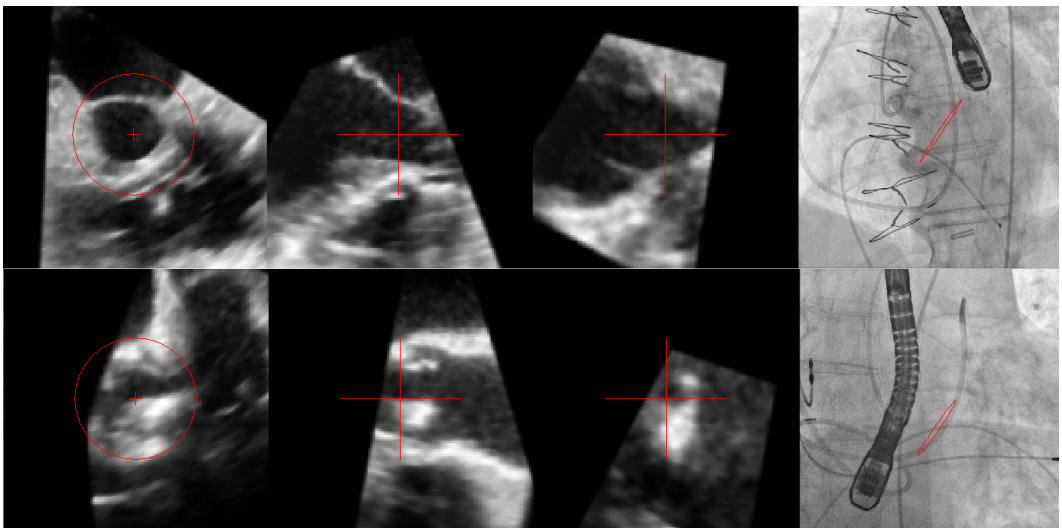

Figure D.3: Aortic annulus segmentation and XRF projection, cases 9-10.

Bibliography

[1] WHO. Global health statistics. Technical report, 2012.

[2] Derek Chin. Echocardiography for transcatheter aortic valve implantation. *European Journal of Echocardiography*, 10(1):i21–i29, 2009.

[3] Richard L Mueller and Timothy A Sanborn. The history of interventional cardiology: cardiac catheterization, angioplasty, and related interventions. *American heart journal*, 129(1):146–172, 1995.

[4] Elliot R McVeigh, Michael A Guttman, Peter Kellman, Amish N Raval, and Robert J Lederman. Real-time, interactive mri for cardiovascular interventions. *Academic radiology*, 12(9):1121–1127, 2005.

[5] Dara L Kraitchman, Wesley D Gilson, and Christine H Lorenz. Stem cell therapy: MRI guidance and monitoring. *Journal of Magnetic Resonance Imaging*, 27(2):299–310, 2008.

[6] Michael T Tomkowiak, Andrew J Klein, Karl K Vigen, Timothy A Hacker, Michael A Speidel, Michael S VanLysel, and Amish N Raval. Targeted transendocardial therapeutic delivery guided by MRIx-ray image fusion. *Catheterization and Cardiovascular Interventions*, 78(3):468–478, 2011.

[7] Charles R Hatt, Ameet K Jain, Vijay Parthasarathy, Andrew Lang, and Amish N Raval. MRI3D ultrasoundX-ray image fusion with electromagnetic tracking for transendocardial therapeutic injections: In-vitro validation and in-vivo feasibility. *Computerized Medical Imaging and Graphics*, 37(2):162–173, 2013.

[8] Ranil De Silva, Luis F Gutiérrez, Amish N Raval, Elliot R McVeigh, Cengizhan Ozturk, and Robert J Lederman. X-ray fused with magnetic resonance imaging (XFM) to target endomyocardial injections validation in a swine model of myocardial infarction. *Circulation*, 114(22):2342–2350, 2006.

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

[10] Cristian A Linte, John Moore, Chris Wedlake, Daniel Bainbridge, Gérard M Guiraudon, Douglas L Jones, and Terry M Peters. Inside the beating heart: An in vivo feasibility study on fusing pre-and intra-operative imaging for minimally invasive therapy. *International journal of computer assisted radiology and surgery*, 4(2):113–123, 2009.

[11] Rui Liao, Shun Miao, and Yefeng Zheng. Automatic and efficient contrast-based 2-D/3-D fusion for trans-catheter aortic valve implantation (TAVI). *Computerized Medical Imaging and Graphics*, 37(2):150–161, 2013.

[12] Zhe Luo, Junfeng Cai, Terry M Peters, and Lixu Gu. Intra-operative 2-D ultrasound and dynamic 3-D aortic model registration for magnetic navigation of transcatheter aortic valve implantation. *Medical Imaging, IEEE Transactions on*, 32(11):2152–2165, 2013.

[13] 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.

[14] Li-Wei Lo and Shih-Ann Chen. Developments and recent advances in catheter ablation of paroxysmal atrial fibrillation. *Future cardiology*, 5(6):557–565, 2009.

[15] Alain Cribier, Helene Eltchaninoff, Assaf Bash, Nicolas Borenstein, Christophe Tron, Fabrice Bauer, Genevieve Derumeaux, Frederic Anselme, François Laborde, and Martin B Leon. Percutaneous transcatheter implantation of an aortic valve prosthesis for calcific aortic stenosis first human case description. *Circulation*, 106(24):3006–3008, 2002.

[16] Philipp Blanke, U Joseph Schoepf, and Jonathon A Leipsic. CT in transcatheter aortic valve replacement. *Radiology*, 269(3):650–669, 2013.

[17] Anson Cheung and Kevin M Lichtenstein. Illustrated techniques for transapical aortic valve implantation. *Annals of Cardiothoracic Surgery*, 1(2):231, 2012.

[18] Ronald K Binder, Jonathon Leipsic, David Wood, Teri Moore, Stefan Toggweiler, Alex Willson, Ronen Gurvitch, Melanie Freeman, and John G Webb. Prediction of optimal deployment projection for transcatheter aortic valve replacement angiographic 3-dimensional reconstruction of the aortic root versus multidetector computed tomography. *Circulation: Cardiovascular Interventions*, 5(2):247–252, 2012.

[19] Ronen Gurvitch, David A Wood, Jonathon Leipsic, Edgar Tay, Mark Johnson, Jian Ye, Fabian Nietlispach, Namal Wijesinghe, Anson Cheung, and John G Webb. Multislice computed tomography for prediction of optimal angiographic deployment projections during transcatheter aortic valve implantation. *JACC: Cardiovascular Interventions*, 3(11):1157–1165, 2010.

[20] Martin B Leon, Craig R Smith, Michael Mack, D Craig Miller, Jeffrey W Moses, Lars G Svensson, E Murat Tuzcu, John G Webb, Gregory P Fontana, Raj R Makkar, et al. Transcatheter aortic-valve implantation for aortic stenosis in patients who cannot undergo surgery. *New England Journal of Medicine*, 363(17):1597–1607, 2010.

[21] Craig R Smith, Martin B Leon, Michael J Mack, D Craig Miller, Jeffrey W Moses, Lars G Svensson, E Murat Tuzcu, John G Webb, Gregory P Fontana, Raj R Makkar, et al. Transcatheter versus surgical aortic-valve replacement in high-risk patients. *New England Journal of Medicine*, 364(23):2187–2198, 2011.

[22] Amir-Ali Fassa, Dominique Himbert, and Alec Vahanian. Mechanisms and management of TAVR-related complications. *Nature Reviews Cardiology*, 10(12):685–695, 2013.

[23] Joanne Shannon, Marco Mussardo, Azeem Latib, Kensuke Takagi, Alaide Chieffo, Matteo Montorfano, and Antonio Colombo. Expert review of cardiovascular therapy. *JACC: Cardiovascular Interventions*, 2011.

[24] Jan-Malte Sinning, Alexander Ghanem, Hannah Steinhäuser, Viktoria Adenauer, Christoph Hammerstingl, Georg Nickenig, and Nikos Werner. Renal function as predictor of mortality in patients after percutaneous transcatheter aortic valve implantation. *JACC: Cardiovascular Interventions*, 3(11):1141–1149, 2010.

[25] Masanori Yamamoto, Kentaro Hayashida, Gauthier Mouillet, Bernard Chevalier, Kentaro Meguro, Yusuke Watanabe, Jean-Luc Dubois-Rande, Marie-Claude Morice, Thierry Lefèvre, and Emmanuel Teiger. Renal function–based contrast dosing predicts acute kidney injury following transcatheter aortic valve implantation. *JACC: Cardiovascular Interventions*, 6(5):479–486, 2013.

[26] Navid Madershahian, Maximilian Scherner, Oliver Liakopoulos, Parwis Rahmanian, Elmar Kuhn, Martin Hellmich, Jochen Mueller-Ehmsen, and Thorsten Wahlers. Renal impairment and transapical aortic valve implantation: impact of contrast medium dose on kidney function and survival. *European Journal of Cardio-Thoracic Surgery*, 41(6):1225–1232, 2012.

[27] Susheel K Kodali, Mathew R Williams, Craig R Smith, Lars G Svensson, John G Webb, Raj R Makkar, Gregory P Fontana, Todd M Dewey, Vinod H Thourani, Augusto D Pichard, et al. Two-year outcomes after transcatheter or surgical aortic-valve replacement. *New England Journal of Medicine*, 366(18):1686–1695, 2012.

[28] Philippe Généreux, Stuart J Head, Rebecca Hahn, Benoit Daneault, Susheel Kodali, Mathew R Williams, Nicolas M Van Mieghem, Maria C Alu, Patrick W Serruys, A Pieter Kappetein, et al. Paravalvular leak after transcatheter aortic valve replacement: the new Achilles' heel? a comprehensive review of the literature. *Journal of the American College of Cardiology*, 61(11):1125–1136, 2013.

[29] Itzhak Kronzon, Vladimir Jelnin, Carlos E Ruiz, Muhamed Saric, Mathew Russell Williams, Albert M Kasel, Anupama Shivaraju, Antonio Colombo, and Adnan Kastrati. Optimal imaging for guiding TAVR: Transesophageal or transthoracic echocardiography, or just fluoroscopy? *JACC: Cardiovascular Imaging*, 8(3):361–370, 2015.

[30] Gang Gao, Graeme Penney, Yingliang Ma, Nicolas Gogin, Pascal Cathier, Aruna Arujuna, Geraint Morton, Dennis Caulfield, Jaswinder Gill, C Aldo Rinaldi, et al. Registration of 3D trans-esophageal echocardiography to X-ray fluoroscopy using image-based probe tracking. *Medical image analysis*, 16(1):38–49, 2012.

[31] R James Housden, Aruna Arujuna, Y Ma, N Nijhof, Geert Gijsbers, Roland Bullens, Mark ONeill, Michael Cooklin, C Aldo Rinaldi, J Gill,

et al. Evaluation of a real-time hybrid three-dimensional echo and x-ray imaging system for guidance of cardiac catheterisation procedures. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*, pages 25–32. Springer, 2012.

[32] Pencilla Lang, Petar Seslija, Michael WA Chu, Daniel Bainbridge, Gerard M Guiraudon, Douglas L Jones, and Terry M Peters. US–fluoroscopy registration for transcatheter aortic valve implantation. *Biomedical Engineering, IEEE Transactions on*, 59(5):1444–1453, 2012.

[33] Peter Mountney, Razvan Ionasec, Markus Kaizer, Sina Mamaghani, Wen Wu, Terrence Chen, Matthias John, Jan Boese, and Dorin Comaniciu. Ultrasound and fluoroscopic images fusion by autonomous ultrasound probe detection. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*, pages 544–551. Springer, 2012.

[34] Charles R Hatt, Michael A Speidel, and Amish N Raval. Efficient feature-based 2D/3D registration of transesophageal echocardiography to x-ray fluoroscopy for cardiac interventions. In *SPIE Medical Imaging*, pages 90361J–90361J. International Society for Optics and Photonics, 2014.

[35] Zhuowen Tu. Probabilistic boosting-tree: Learning discriminative models for classification, recognition, and clustering. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 2, pages 1589–1596. IEEE, 2005.

[36] Yefeng Zheng, Bogdan Georgescu, Haibin Ling, Shaohua Kevin Zhou, Michael Scheuering, and Dorin Comaniciu. Constrained marginal space learning for efficient 3D anatomical structure detection in medical images. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 194–201. IEEE, 2009.

[37] Gustavo Carneiro, Bogdan Georgescu, Sara Good, and Dorin Comaniciu. Detection and measurement of fetal anatomies from ultrasound images using a constrained probabilistic boosting tree. *Medical Imaging, IEEE Transactions on*, 27(9):1342–1355, 2008.

[38] Yefeng Zheng, Adrian Barbu, Bogdan Georgescu, Michael Scheuering, and Dorin Comaniciu. Fast automatic heart chamber segmentation from 3D CT data using marginal space learning and steerable features. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8. IEEE, 2007.

[39] Yefeng Zheng, Adrian Barbu, Bogdan Georgescu, Michael Scheuering, and Dorin Comaniciu. Four-chamber heart modeling and automatic segmentation for 3-D cardiac CT volumes using marginal space learning and steerable features. *Medical Imaging, IEEE Transactions on*, 27(11):1668–1681, 2008.

[40] Tobias Heimann, Peter Mountney, Matthias John, and Razvan Ionasec. Learning without labeling: Domain adaptation for ultrasound transducer localization. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, pages 49–56. Springer, 2013.

[41] Tobias Heimann, Peter Mountney, Matthias John, and Razvan Ionasec. Real-time ultrasound transducer localization in fluoroscopy images by transfer learning from synthetic training data. *Medical image analysis*, 18(8):1320–1328, 2014.

[42] Mohamed Esmail Karar, DR Merk, C Chalopin, Thomas Walther, Volkmar Falk, and Oliver Burgert. Aortic valve prosthesis tracking for transapical aortic valve implantation. *International journal of computer assisted radiology and surgery*, 6(5):583–590, 2011.

[43] Alexander Brost, Rui Liao, Norbert Strobel, and Joachim Hornegger. Respiratory motion compensation by model-based catheter tracking during EP procedures. *Medical Image Analysis*, 14(5):695–706, 2010.

[44] 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 *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009*, pages 394–401. Springer, 2009.

[45] YingLiang Ma, Nicolas Gogin, Pascal Cathier, R James Housden, Geert Gijsbers, Michael Cooklin, Mark O'Neill, Jaswinder Gill, C Aldo Rinaldi, Reza Razavi, et al. Real-time x-ray fluoroscopy-based catheter detection and tracking for cardiac electrophysiology interventions. *Medical physics*, 40(7):071902, 2013.

[46] Fausto Milletari, Vasileios Belagiannis, Nassir Navab, and Pascal Fallavollita. Fully automatic catheter localization in c-arm images using 1-sparse coding. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*, pages 570–577. Springer, 2014.

[47] Wen Wu, Terrence Chen, Adrian Barbu, Peng Wang, Norbert Strobel, Shaohua Kevin Zhou, and Dorin Comaniciu. Learning-based hypothesis fusion for robust catheter tracking in 2D x-ray fluoroscopy. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 1097–1104. IEEE, 2011.

[48] Juergen Gall and Victor Lempitsky. Class-specific hough forests for object detection. In *Decision Forests for Computer Vision and Medical Image Analysis*, pages 143–157. Springer, 2013.

[49] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[50] Paul VC Hough. Method and means for recognizing complex patterns. Technical report, 1962.

[51] Dana H Ballard. Generalizing the hough transform to detect arbitrary shapes. *Pattern recognition*, 13(2):111–122, 1981.

[52] A Criminisi, J Shotton, and E Konukoglu. Decision forests for classification, regression, density estimation, manifold learning and semi-supervised learning. *Microsoft Research Cambridge, Tech. Rep. MSRTR-2011-114*, 5(6):12, 2011.

[53] Ameet Jain, Luis Gutierrez, and Douglas Stanton. 3D TEE registration with x-ray fluoroscopy for interventional cardiac applications. In *Functional Imaging and Modeling of the Heart*, pages 321–329. Springer, 2009.

[54] Gang Gao, Graeme Penney, Nicolas Gogin, Pascal Cathier, Aruna Arujuna, Matt Wright, Dennis Caulfield, Aldo Rinaldi, Reza Razavi, and Kawal Rhode. Rapid image registration of three-dimensional transesophageal echocardiography and x-ray fluoroscopy for the guidance of cardiac interventions. In *Information Processing in Computer-Assisted Interventions*, pages 124–134. Springer, 2010.

[55] Pencilla Lang, Petar Seslija, Damiaan F Habets, Michael WA Chu, David W Holdsworth, and Terry M Peters. Three-dimensional ultrasound probe pose estimation from single-perspective x-rays for image-guided interventions. In *Medical Imaging and Augmented Reality*, pages 344–352. Springer, 2010.

[56] Pencilla Lang, Petar Seslija, Daniel Bainbridge, Gerard M Guiraudon, Doug L Jones, Michael W Chu, David W Holdsworth, and Terry M Peters. Accuracy assessment of fluoroscopy-transesophageal echocardiography registration. In *SPIE Medical Imaging*, pages 79641Y–79641Y. International Society for Optics and Photonics, 2011.

[57] Markus Kaiser, Matthias John, Anja Borsdorf, Peter Mountney, Razvan Ionasec, Alois Nöttling, Philipp Kiefer, Jörg Seeburger, and Thomas Neumuth. Significant acceleration of 2D-3D registration-based fusion of ultrasound and x-ray images by mesh-based DRR rendering. In *SPIE*

Medical Imaging, pages 867111–867111. International Society for Optics and Photonics, 2013.

[58] Markus Kaiser, Matthias John, Tobias Heimann, Alexander Brost, Thomas Neumuth, and Georg Rose. 2D/3D registration of TEE probe from two non-orthogonal c-arm directions. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*, pages 283–290. Springer, 2014.

[59] R James Housden, YingLiang Ma, Aruna Arujuna, Niels Nijhof, Pascal Cathier, Geert Gijsbers, Roland Bullens, Jaswinder Gill, C Aldo Rinaldi, Victoria Parish, et al. Extended-field-of-view three-dimensional transesophageal echocardiography using image-based x-ray probe tracking. *Ultrasound in medicine & biology*, 39(6):993–1005, 2013.

[60] Xianliang Wu, James Housden, Yingliang Ma, Daniel Rueckert, and Kawal S Rhode. Real-time catheter extraction from 2D x-ray fluoroscopic and 3D echocardiographic images for cardiac interventions. In *Statistical Atlases and Computational Models of the Heart. Imaging and Modelling Challenges*, pages 198–206. Springer, 2013.

[61] Primoz Markelj, D Tomaževič, Bostjan Likar, and F Pernuš. A review of 3D/2D registration methods for image-guided interventions. *Medical image analysis*, 16(3):642–661, 2012.

[62] Mark Harris et al. Optimizing parallel reduction in cuda. *NVIDIA Developer Technology*, 2(4), 2007.

[63] Charles Hatt. A tutorial on combining nonlinear optimization with cuda. 2015.

[64] Graeme P Penney, Jürgen Weese, John A Little, Paul Desmedt, Derek LG Hill, and David J Hawkes. A comparison of similarity measures for use in 2-D-3-D medical image registration. *Medical Imaging, IEEE Transactions on*, 17(4):586–595, 1998.

[65] Damiaan F Habets, Steven I Pollmann, Xunhua Yuan, Terry M Peters, and David W Holdsworth. Error analysis of marker-based object localization using a single-plane XRII. *Medical physics*, 36(1):190–200, 2009.

[66] Charles R Hatt, Amish N Raval, and Michael A Speidel. A fiducial apparatus for 6DOF pose estimation of an external echo probe from a single x-ray projection: Initial simulation studies on design requirements. *World congress on Medical Physics and Biomedical Engineering*, page In Press, 2015.

[67] Christelle Gendrin, Hugo Furtado, Christoph Weber, Christoph Bloch, Michael Figl, Supriyanto Ardjo Pawiro, Helmar Bergmann, Markus Stock, Gabor Fichtinger, Dietmar Georg, et al. Monitoring tumor motion by real time 2d/3d registration during radiotherapy. *Radiotherapy and oncology*, 102(2):274–280, 2012.

[68] Kristf Ralovich, Matthias John, Estelle Camus, Nassir Navab, and Tobias Heimann. 6DoF Catheter Detection, Application to Intracardiac Echocardiography. In Polina Golland, Nobuhiko Hata, Christian Barillot,

Joachim Hornegger, and Robert Howe, editors, *Medical Image Computing and Computer-Assisted Intervention MICCAI 2014*, volume 8674 of *Lecture Notes in Computer Science*, pages 635–642. Springer International Publishing, 2014.

[69] Jacek Baranowski, Henrik Ahn, Wolfgang Freter, Niels-Erik Nielsen, Eva Nylander, Birgitta Janerot-Sjoberg, Michael Sandborg, and Lars Wallby. Echo-guided presentation of the aortic valve minimises contrast exposure in transcatheter valve recipients. *Catheterization and cardiovascular interventions*, 77(2):272–275, 2011.

[70] Michael S Kim, John Bracken, Niels Nijhof, Ernesto E Salcedo, Robert A Quaife, John C Messenger, and John D Carroll. Integrated 3D echo-x-ray navigation to predict optimal angiographic deployment projections for TAVR. *JACC: Cardiovascular Imaging*, 7(8):847–848, 2014.

[71] Rebecca T Hahn, Stephen H Little, Mark J Monaghan, Susheel K Kodali, Mathew Williams, Martin B Leon, and Linda D Gillam. Recommendations for comprehensive intraprocedural echocardiographic imaging during TAVR. *JACC: Cardiovascular Imaging*, 8(3):261–287, 2015.